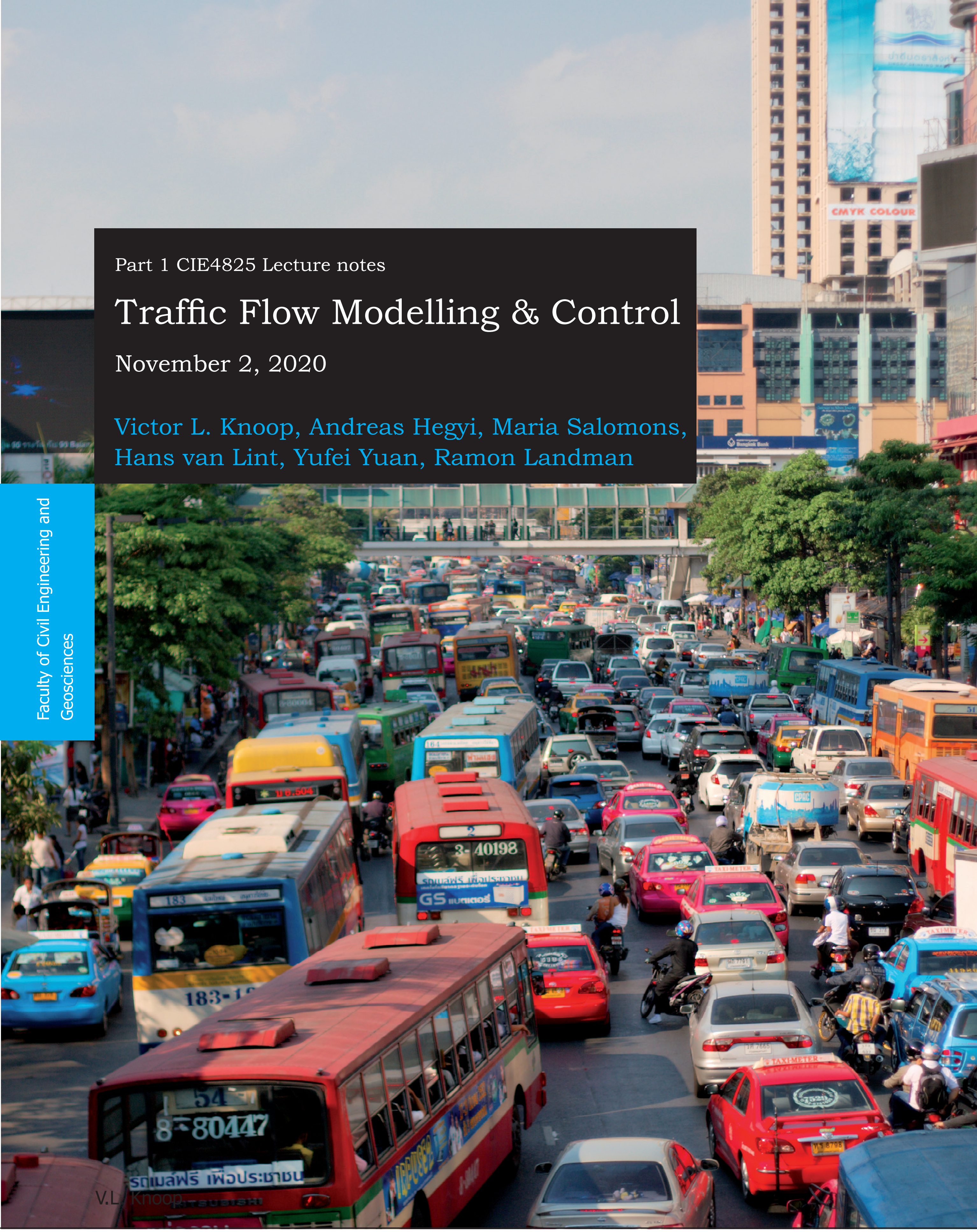
Part 1 CIE4825 Lecture notes

Traffic Flow Modelling & Control

November 2, 2020

Victor L. Knoop, Andreas Hegyi, Maria Salomons,
Hans van Lint, Yufei Yuan, Ramon Landman

Faculty of Civil Engineering and
Geosciences



TU Delft
Delft
University of
Technology

Challenge the future

The authors thank Christian Haugen for making the picture on the cover page available. The picture is licensed according to attribution 2.0 Generic (CC BY 2.0).

Version

2018 November 28	Initial release.
2019 February 10	Chapter “Traffic state estimation” added.
	Chapter “Local intersection control – advanced” added.
2019 February 20	Chapter “An introduction to Node Models” added.
2019 March 2	Chapter “COSCAL – A cooperative speed control algorithm to resolve jams” added.
2019 March 14	Chapter “Coordinated ramp metering” added.
	Chapter “A design methodology for traffic control systems” added.
2019 November 8	Minor changes
2020 Summer	revision overall; addition of part III

Preface

Traffic processes cause several problems in the world. Traffic delay, pollution are some of it. They can be solved with the right road design or traffic management (control) measure. Before implementing these designs of measures, though, their effect could be tested. To this end, knowledge of traffic flow theory is needed, and to be combined with knowledge on control theory to in the end be applied to a traffic system.

This book accompanies the courses CIE4825 and CIE5821, which are given at Delft University of Technology. This course has been introduced in the year 2018. Before that, there were separate courses: Traffic Flow Theory and Simulations, Traffic Management and Control and Innovations in Dynamic Traffic Management. However, we believe that the modelling and control aspects in these courses should be considered together. Therefore, since academic year 2018-2019, there are two *integrated* courses: Traffic Flow Modelling and Control Part I (course code 4825) and Part II (course code 5821), which reflects our wish to integrate the modelling and control aspects.

This book contains 4 parts. Part 1 and part 2 are what we consider the "basics" which are being discussed in course 4825. Part 3 and part 4 are considered more advanced, and discussed in course 5821. Note this distinction is arbitrary and partially based on the amount of study load in each of the courses. There is another separation: two parts of this book discuss traffic flow modelling (part 1 and part 3), and two parts control (part 2 and part 4). This reader is to a large extent based on – or taken from – material developed for the previous courses: Introduction to Traffic Flow Theory: An introduction with exercises. Second edition, ISBN 978-94-6366-062-4 and the course reader for the course 4822.

To learn an engineering discipline, practicing is essential. The courses 4825 and 5821 are new per academic year 2018-2019. As help for students, for some topics questions (with answers) are included. These are based on exams of previous years for the course Traffic Flow Theory and Simulation. Also, the 2018-2019 exam of the 4825 course is included. In the relevant chapters, a reference is made to the questions that students should be able to answer.

The book is, like science, not finished. The plan is to improve this book over the years. If you have remarks – errors, additional request, things which are unclear – please let us know at `v.l.knoop@tudelft.nl`, `a.hegyi@tudelft.nl`, or `a.m.salomons@tudelft.nl`. Comments of earlier years have been included, and we would like to mention the comments of Attila Borsos in particular, who has pointed at many minor mistakes, that were corrected in this version.

The book is meant as introduction to the field of traffic flow theory. Only basic calculus is assumed as base knowledge. For different approaches, the reader can continue in other books, including:

- May, A.D. Traffic flow fundamentals. 1990.
- Leutzbach, W. Introduction to the theory of traffic flow. Vol. 47. Berlin: Springer-Verlag, 1988.
- Daganzo, C.F. Fundamentals of transportation and traffic operations. Vol. 30. Oxford: Pergamon, 1997.

- Treiber, M., and A. Kesting. "Traffic flow dynamics." Traffic Flow Dynamics: Data, Models and Simulation, Springer-Verlag Berlin Heidelberg (2013).
- Elefteriadou, L. An introduction to traffic flow theory. Vol. 84. New York, NY, USA: Springer, 2014.
- Ni, D. Traffic Flow Theory: Characteristics, Experimental Methods, and Numerical Techniques. Butterworth-Heinemann, 2015.
- Knoop, V.L. Introduction to Traffic Flow Theory: An introduction with exercises. Second edition, ISBN 978-94-6366-062-4, 2018
- Kessels, F., Traffic Flow Modelling. Introduction to Traffic Flow Theory Through a Genealogy of Models. Springer, 2018

Of course, there also is a vast, and ever expanding, body of scientific literature which the reader can use as follow-up material. The list of references in this reader provides a good starting point.

Victor Knoop, Andreas Hegyi, Maria Salomons, Hans van Lint, Yufei Yuan

2018 November 13

Table of Contents

Part I

Traffic Flow Modelling and Control Part I – Modelling

Chapter 1

Variables

After this chapter, the student is able to:

- Use the right terms for level of descriptions, and stationarity and homogeneity
- Give an interpretation to variables $\{x, n, t, v, q, k\}$
- Analyse and explain the differences between observation methods and averaging methods (time mean vs space mean)
- Compute and explain Edie's generalized variables of traffic

This chapter describes the main variables which are used in traffic flow theory. Section 1-1 will show the different levels (microscopic, macroscopic and other levels) at which traffic is generally described. Section 1-2 will describe different principles (local, instantaneous and spatio-temporal) to measure the traffic flow. The last section (1-3) describes traffic flow characteristics.

1-1 Levels of description

This section will show the different levels at which traffic is generally described. Sections 1-1-1 and 1-1-2 will discuss the variables in the microscopic and macroscopic descriptions in more detail.

In a microscopic traffic description, every vehicle-driver combination is described. The smallest element in the description is the vehicle-driver combination. The other often used level of traffic flow description is the macroscopic traffic description. Different from the microscopic description, this level does not consider individual vehicles. Instead, the traffic variables are aggregated over several vehicles or, most commonly, a road stretch. Typical characteristics of the traffic flow on a road stretch are the average speed, vehicle density or flow (see section 1-1-2).Other levels of description can also be used, these are described in the last section(see section 1-1-3).

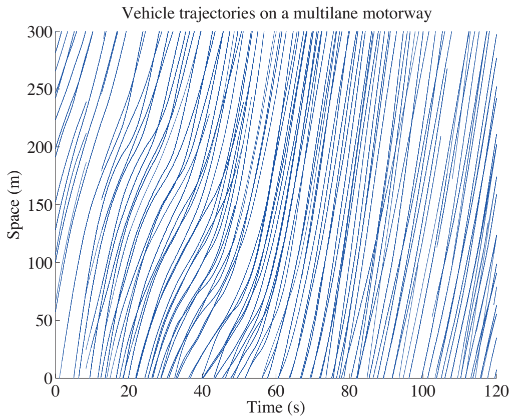


Figure 1-1: Vehicle trajectories on a multilane motorway

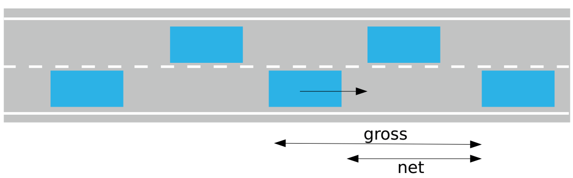


Figure 1-2: The difference between gross and net spacing (or headway)

1-1-1 Microscopic

In a microscopic traffic description, the vehicle-driver combinations (often referred to as "vehicles", which we will do from now on) are described individually. Full information of a vehicle is given in its trajectory, i.e. the specification of the position of the vehicle at all times. To have full information on these, the positions of all vehicles at all times have to be specified. A graphical representation of vehicle trajectories is given in figure 1-1

The trajectories are drawn in a space time plot, with time on the horizontal axis. Note that vehicle trajectories can never go back in time. Trajectories might move back in space if the vehicles are going in the opposite direction, for instance on a two-lane bidirectional rural road. This is not expected on motorways. The slope of the line is the speed of the vehicles. Therefore, the trajectories cannot be vertical – that would mean an infinite speed. Horizontal trajectories are possible at speed zero.

Basic variables in the microscopic representation are speed, headway, and space headway. The speed is the amount of distance a vehicle covers in a unit of time, which is indicated by v. Sometimes, the inverse of speed is a useful measure, the amount of time a vehicle needs to cover a unit of space; this is called the pace p. Furthermore, there is the space headway or spacing (s) of the vehicle. The net space headway is the distance between the vehicle and its leader. This is also called the gap. The gross space headway of a (following) vehicle is the distance *including* the length of the vehicle, so the distance from the rear bumper of the leading vehicle to the rear bumper of the following vehicle. Similarly, we can define the time it takes for a follower to get to reach (with its front bumper) the position of its leader's rear bumper. This is called the net time headway. If we also add the time it costs to cover the distance of a vehicle length, we get the gross time headway. See also figure 1-2. The symbol

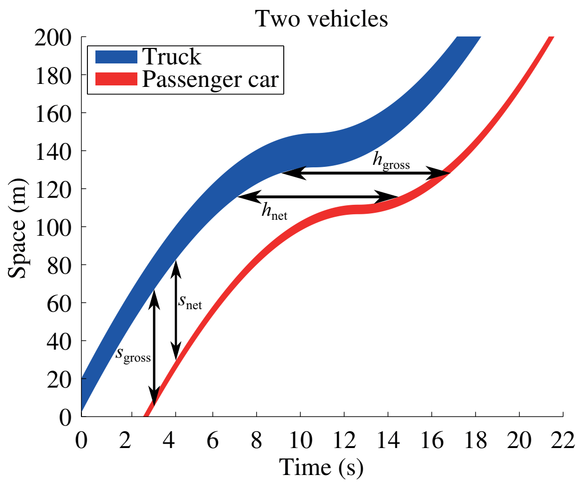


Figure 1-3: The microscopic variables explained based on two vehicles

used to indicate the headway is h.

From now on, in this reader we will use the following conventions:

- Unless specified otherwise, headway means time headway
- Unless specified otherwise, headways and spacing are given as gross values

Figure 1-3 shows the variables graphically. The figure shows two vehicles, a longer vehicle and a shorter vehicle. Note that the length of the vehicles remains unchanged, so the difference between the gross and net spacing is the same, namely the vehicle length. The lines hence have the same thickness (i.e., vertical extension), being the vehicle length. They seem thicker as the line is more horizontal; this is a perception error since the reduce in thickness orthogonal to the direction of the line (which has no physical meaning). Whereas the difference between gross and net space headway is constant (namely the vehicle length), the difference between the gross and net time headway changes based on the vehicle speed.

In a trajectory plot, the slope of the line is the speed. If this slope changes, the vehicle accelerates or decelerates. So, the curvature of the lines in a trajectory plot shows the acceleration or deceleration of the vehicle. If the slope increases, the vehicle accelerates, if it decreases, it decelerates.

1-1-2 Marcoscopic

In a macroscopic traffic description, one does not describe individual vehicles. Rather, one describes for each road section the aggregated variables. That is, one can specify the density k, i.e. how close in space vehicles are together. Furthermore, one can specify the flow q i.e. the number of vehicles passing a reference point per unit of time. Finally, one can describe the average speed u of the vehicles on a road section. Other words for flow are throughput, volume or intensity; we will strictly adhere to the term flow to indicate this concept.

All of the mentioned macroscopic variables have their microscopic counterpart. This is summarized in table 1-1. The density is calculated as one divided by the average spacing, and

Table 1-1: Overview of the microscopic and macroscopic variables and their relationship; the pointy brackets indicate the mean.

Microscopic	symbol	unit	Macroscopic	symbol	unit	relation
Headway	h	s	Flow	q	vtg/h	$q = \dfrac{3600}{\langle h \rangle}$
Spacing	s	m	Density	k	vtg/km	$k = \dfrac{1000}{\langle s \rangle}$
Speed	v	m/s	Average speed	u	km/u	$u = 3.6 \langle v \rangle$

is calculated over a certain road stretch. For instance, if vehicles have a spacing of 100 meters, there are 1/100 vehicles per meter, or 1000/100=10 veh/km. The flow is the number of vehicles that pass a point per unit of time. It can be directly calculated from the headways by dividing one over the average headway. For instance, if all vehicles have a headway of 4 seconds, there are 1/4 vehicles per second. That means there are 3600(s/h)/4(s/veh)= 900 veh/h. In table 1-1 units are provided and in the conversion from one quantity to the other, one needs to pay attention. Note that the provided units are not obligatory: one can present individual speed in km/h, or density in veh/hm. However, always pay attention to the units before converting or calculating.

Relation to the microscopic level

The average speed is calculated as an average of the speeds of vehicles at a certain road stretch. This speed differs from the average speed obtained by averaging speed of all vehicles passing a certain point. The next section explains the different measuring principles. The full explanation of the differences between the two speeds and how one can approximate the (space) average speed by speeds of vehicles passing a certain location is presented in section 3-4.

Another concept for a traffic flow, in particular in relation to a detector (see also section 1-2), is the occupancy o. This indicates which fraction of a time a detector embedded in the roadway is occupied, i.e. whether there is a vehicle on top of the detector. Suppose a detector has a length L_{det} and a vehicle a length of L_i. The occupancy is defined as the time the detector is occupied, τ_{occupied} divided by all time, i.e. the time it is occupied and time is is not occupied $\tau_{\text{not occupied}}$

$$o = \frac{\tau_{\text{occupied}}}{\tau_{\text{occupied}} + \tau_{\text{not occupied}}} \tag{1-1}$$

The occupation time can be derived from the distances and the speed. The distance the vehicle has to cover from the moment it starts occupying the detector up to the time it leaves the detector is its own length plus the length of the detector. Hence, the occupancy time is

$$\tau_{\text{occupied}} = \frac{L_i + L_{\text{det}}}{v} \tag{1-2}$$

Once the first vehicle drives off the detector, the distance for the following vehicle to reach the detector is the gap (i.e., the spacing minus the length of the vehicle) between the vehicles minus the length of the detector. The amount of time this takes is

$$\tau_{\text{not occupied}} = \frac{s - L_i - L_{\text{det}}}{v} \tag{1-3}$$

Substituting the expressions for the occupancy time and the non-occupancy time into equation 1-1 and rearranging the terms, we get

$$o = \frac{L_i + L_{\text{det}}}{s} \tag{1-4}$$

In practice, the detector length is known for a certain road configuration (usually, there are country specific standards). So assuming a vehicle length, one can calculate the spacing, and hence the density, from the occupancy.

1-1-3 Other levels

Apart from the macroscopic and microscopic traffic descriptions, there are three other levels to describe traffic. They are are less common, and are therefore not discussed in detail. The levels mentioned here are mainly used in computer simulation models.

Mesoscopic

The term mesoscopic is used for any description of traffic flow which is in-between macroscopic and microscopic. It can also be a term for simulation models which calculate some elements macroscopically and some microscopically. For instance, Dynasmart (2003), uses such a mesoscopic description.

Submicroscopic

In a submicroscopic description the total system state is determined by the sub levels of a vehicle and/or driver. Processes which influence the speed of a vehicle, like for instance mechanically throttle position and engine response, or psychologically speed perception, are explicitly modelled. This allows to explicitly model the (change in) reaction on inputs. For instance, what influence would cars with a stronger engine have on the traffic flow.

Network level

A relatively new way of describing the traffic state is the network level. This has recently gained attention after the publication by Geroliminis and Daganzo (2008). Instead of describing a part of a road as smallest element, one can take an area (e.g. a city center) and consider this as one unit.

1-2 Measuring principles

Whereas the previous sections described which variables are used to describe traffic flow, this section will introduce three principles of measuring the traffic flow. These principles are local, instantaneous and spatio-temporal.

1-2-1 Local

With local measurements one observes traffic at one location. This can be for instance a position at the roadway. To measure motorway traffic, often inductive loops are used. These are coils embedded in the pavement in which a electrical current produces a (vertical) magnetic field. If a car enters or leaves this magnetic field, this can be measured in the current of the coil. Thus, one knows how long a loop is occupied. In the US, usually single loops are used, giving the occupancy of the loop. Using equation 1-4, this can be translated into density. The detectors also measure the flow. As will be explained later in section 3-1, this suffices to completely characterise the traffic flow.

This determination of density builds upon the assumption of the vehicle length being known. One can also measure the length of a vehicle for passing vehicles, using dual loop detectors. These are inductive loops which are placed a known short distance (order of 1 m) from each other. If one carefully measures the time between the moment the vehicle starts occupying the first loop and the moment it starts occupying the second loop, one can measure its speed. If its speed is known, as well as the time it occupies one loop, the length of the vehicle can also be determined.

1-2-2 Instantaneous

Contrary to local measurements, there are instantaneous measurements. These are measurements which are taken at one moment in time, most likely over a certain road stretch. An example of such a measurement is an areal photograph. In such a measurement, one can clearly distinguish spatial characteristics, as for instance the density. However, measuring the temporal component (flow) is not possible.

1-2-3 Spatio-temporal measurements

Apart from local or instantaneous measurements, one can use measurements which stretch over a period of time *and* a stretch of road. For instance, the trajectories in figure 1-1 are an example thereof. This section will introduce Edie's definitions of flow, density and speed for an area in space and time.

A combination of instantaneous measurements and local measurements can be found in remote sensing observations. These are observations which stretch in both space and time. For instance, the trajectories presented in figure 1-1 can be observed using a camera mounted on a high point or a helicopter. One can see a road stretch, and observe it for a period of time.

Measuring average speed by definition requires an observation which stretches over time and space. At one location, one cannot determine speed, nor at one moment. One needs at least two locations close by (several meters) or two time instances close by. Ignoring these short distances one can calculate a local mean speed based on speeds of the vehicles passing by location. Ignoring the short times, one can calculate the time mean speed from the speed of the vehicles currently at the road. At this moment, we suffice by mentioning these average speeds are different. Section 3-4 will show how the space mean speed can be approximated from local measurements.

Figure 1-4 shows the same trajectories as figure 1-1, but in figure1-4 an area is selected. Trajectories within this area in space and time are coloured red. Note that an selected area is not necessarily square. It is even possible to have a convex area, or boundaries moving backwards and forwards in time. The definitions as introduced here will hold for all types of areas, regardless of their shape in space-time.

Let us consider the area X. We indicate its size by W_X, which is expressed in km-h, or any other unit of space times time. For all vehicles, we consider the distance they drive in area X, which we call $d_{X,i}$. Adding these for all vehicles i gives the total distance covered in area X, indicated by TD:

$$TD = \sum_{\text{all vehicles } i} d_{X,i} \tag{1-5}$$

For a rectangular area in space and time, the distance covered might be the distance from the upstream end to the downstream end, but the trajectory can also begin and/or end at the side of the area, at a certain time. In that case, the distance is less than the full distance.

Similarly, we can define the time a vehicle spends in area X, $t_{X,i}$, which we can sum for all vehicles i to get the total time spent in area X, indicated by TT.

$$TT = \sum_{\text{all vehicles } i} t_{X,i} \tag{1-6}$$

Obviously, both quantities grow in principle with the area size. Therefore, the traffic flow is best characterised by the quantities TD/W_X and TT/W_X. This gives the flow and the density respectively:

$$q = \frac{TD}{W_X} \tag{1-7}$$

$$k = \frac{TT}{W_X} \tag{1-8}$$

Intuitively, the relationship is best understood reasoning from the known relations of density and flow. Starting with a situation of 1000 veh/h at a cross section, and an area of 1 h and 2 km. In 1 hour, 1000 vehicles pass by, which all travel 2 kilometres in the area. (There the vehicles which cannot cover the 2 km because the time runs out, but there are just as many which are in the section when the time window starts). So the total distance is the flow times the size of the area: $TD = qW_X$. This can be simply rewritten to equation 1-8.

A similar relation is constructed for the density, considering again the rectangular area of 1 hour times 2 kilometres. Starting with a density of 10 veh/km, there are 20 vehicles in the area, which we all follow for one hour. The total time spent, is hence 10*2*1, or $TT = kW_X$. This can be rewritten to equation 1-8.

The average speed is defined as the total distance divided by the total time, so

$$u = \frac{TD}{TT} \tag{1-9}$$

The average travel time over a distance l can be found as the average of the time a vehicle travels over a distance l. In an equation, we find:

$$\langle tt \rangle = \left\langle \frac{l}{v} \right\rangle = l \left\langle \frac{1}{v} \right\rangle \tag{1-10}$$

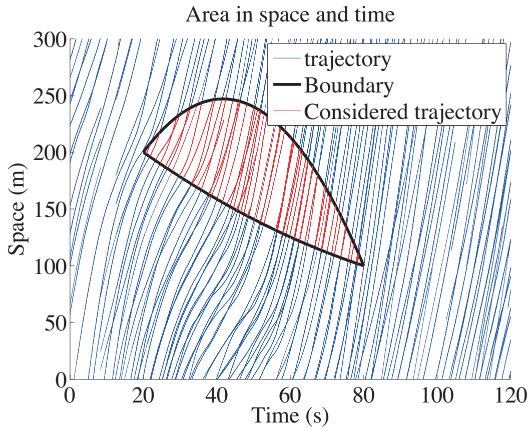


Figure 1-4: Vehicle trajectories and the selection of an area in space and time

In this equation, tt indicates the travel time and the pointy brackets indicate the mean. This can be measured for all vehicles passing a road stretch, for instance at a local detector. Note that the mean travel time is *not* equal to the distance divided by the mean speed:

$$\langle tt \rangle = l \left\langle \frac{1}{v} \right\rangle \neq l \frac{1}{\langle v \rangle} \tag{1-11}$$

In fact, it can be proven that in case speeds of vehicles are not the same, the average travel time is underestimated if the mean speed is used.

$$\langle tt \rangle = l \left\langle \frac{1}{v} \right\rangle \leq l \frac{1}{\langle v \rangle} \tag{1-12}$$

The harmonically averaged speed (i.e., 1 divided by the average of 1 divided by the speed) does provide a good basis for the travel time estimation. In an equation, we best first define the pace, p_i:

$$p_i = \frac{1}{v_i} \tag{1-13}$$

The harmonically averaged speed now is

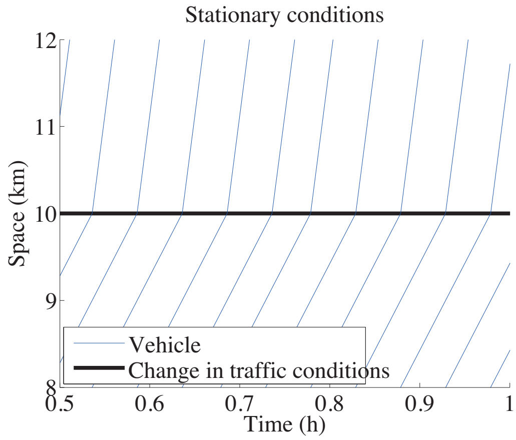
$$\langle v \rangle_{\text{harmonically}} = \frac{1}{\langle p \rangle} = \frac{1}{\left\langle \frac{1}{v_i} \right\rangle} \tag{1-14}$$

The same quantity is required to find the space mean speed. Section 3-4 shows the difference qualitatively. In short, differences can be several tens of percents.

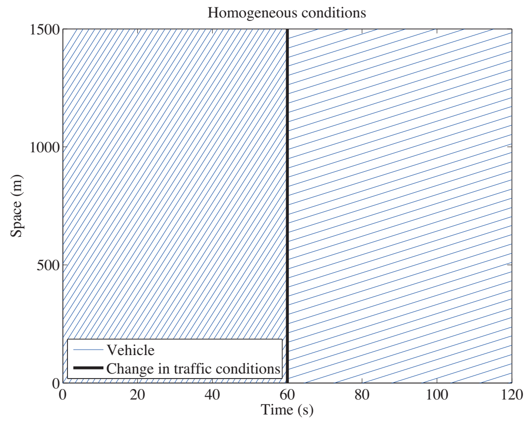
1-3 Stationarity and homogeneity

Traffic characteristics can vary over time and/or over space. There are dedicated names for traffic if the state does not change.

Traffic is called stationary if the traffic flow does not change over time (but it can change over space). An example can be for instance two different road sections with different characteristics. An example is given in figure 1-5a, where there first is a low speed, then the speed of the vehicles is high.



(a) Stationary conditions



(b) Homogeneous conditions

Traffic is called homogeneous if the traffic flow does not change over space (but it can change over time). An example is given in figure 1-5b, where at time 60 the speed decreases at the whole road section. This is much less common than the stationary conditions. For this type of situation to occur, the traffic regulations have to change externally. For instance, the speed limits might change at a certain moment in time (lower speeds at night).

Selected problems

For this chapter, consider problems: 5, 6, 162, 172, 173, 174, 175, 196, 217, 232, A-10-3, A-11-3

Chapter 2

Cumulative curves

After this chapter, the student is able to:

- Construct (slanted/oblique) cumulative curves in practice or from theoretical problem
- Interpret these and calculate: delays, travel times, density, flow

This chapter discusses cumulative curves, also known as cumulative flow curves. The chapter first defines the cumulative curves (section 2-1), then it is shown how traffic characteristics can be derived from these (section 2-2). Section 2-4 shows the application of slanted cumulative curves.

2-1 Defenition

The function $N_x(t)$ is defined as the number of vehicles that have passed a point x at time t and is only used for traffic into one direction. Hence, this function only increases over time. Strictly speaking, this function is a step function increasing by one every time a vehicle passes. However, for larger time spans and higher flow rates, the function is often smoothed into a continuous differentiable function.

The increase rate of this function equals the flow:

$$\frac{dN}{dt} = q \tag{2-1}$$

Hence from the flow, we can construct the cumulative curve:

$$N = \int q dt \tag{2-2}$$

This gives one degree of freedom, the initial value. This can be chosen freely, or should be adapted to cumulative curves for other locations.

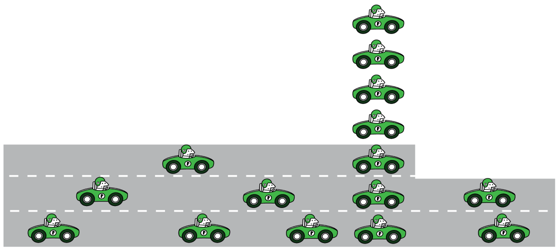


Figure 2-1: Illustration of a vertical queue

2-2 Vertical queuing model

A vertical queuing model is a model which assumes an unlimited inflow and and an outflow which is restricted to capacity. The vehicles which cannot pass the bottleneck are stacked "vertically" and do not occupy any space. Figure 2-1 illustrates this principle.

Let us now study the dynamics of such a queue. We discretize time in steps of duration Δt, referred to by index t. The demand is externally given, and indicated by D. At time steps t we compute the flow into and out of the stack (the number of vehicles in the stack indicated as S). In between the time steps, indicated here as $t + 1/2$, the number of vehicles in the stack is updated based on the flows q. Then, the stack provides the basis for the flows in the next time step.

The stack starts at zero. Then, for each time step first the inflow to the stack is computed.

$$q_{\text{in},t} = D \tag{2-3}$$

and the stack is updated accordingly, going to an intermediate state at time step t+1/2. This intermediate step is the number of vehicles in the queue if there was no outflow, so the original queue plus the inflow:

$$S_{t+1/2} = S_t + q_{\text{in}}\Delta t \tag{2-4}$$

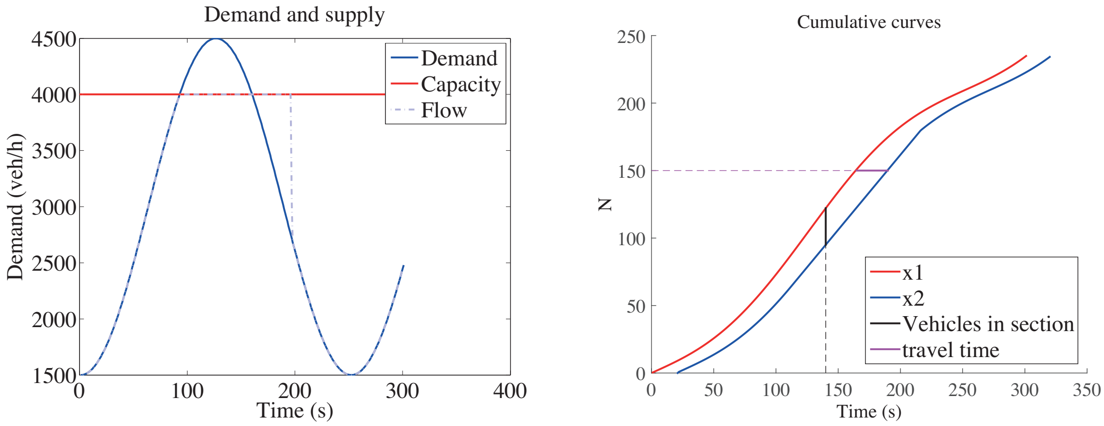
Then, the outflow out of the stack (q_{out}) is the minimum of the number of vehicles in this intermediate queue and the maximum outflow determined by the capacity C:

$$q_{\text{out}} = \min\{C\Delta t, S_{i+1/2}\} \tag{2-5}$$

The stack after the time step is then computed as follows

$$S_{i+1} = S_{i+1/2} - q_{\text{out}}\Delta t = S_i + (q_{\text{in},i} - q_{\text{out},i})\,\Delta t \tag{2-6}$$

Let us consider a situation as depicted in figure 2-1, and we are interested in the delays due to the bottleneck with a constant capacity of 4000 veh/h. The demand curve is plotted in figure 2-2a. The flows are determined using the vertical queuing model. The flows are also shown in figure 2-2a. Note that the area between the flow and demand curve where the demand is higher than the flow (between approximately 90 to 160 seconds), is the same as the area between the curves where the flow is higher than the demand (between approximately 160 and 200 seconds). The reasoning is that the area represents a number of vehicles (a flow times a time). From 90 to 160 seconds the demand is higher than the flow, i.e., the inflow



(a) Demand and supply

(b) Cumulative curves

Figure 2-2: Demand and cumulative curves

is higher than the outflow. The area represents the number of vehicles that cannot pass the bottleneck, and hence the number of queued vehicles. From 160 seconds, the outflow of the queue is larger than the inflow. That area represents the number of vehicles that has left the queue, and cannot be larger than the number of vehicles queued. Moreover, the flow remains at capacity until the stack is empty, so both areas must be equal.

2-3 Travel times, densities and delays

This section explains how travel times and delays. Delay can be computed using cumulative curves. Note that this methodology does not take spillback effects into account. If one requires this to be accounted for, please refer to shockwave theory (chapter 4).

2-3-1 Construction of cumulative curves

The cumulative curves for the above situation is shown in figure 2-2b. The curves show the flows as determined by the vertical queuing model. For the inflow we hence use equation 2-3 and for the outflow we use 2-5; for both, the cumulative curves are constructed using equation 2-2.

2-3-2 Travel times, number of vehicles in the section

A black line is drawn at $t = 140s$ in figure 2-2b. The figure shows by intersection of this line with the graphs how many vehicles have passed the upstream point $x1$ and how many vehicles have passed the downstream point $x2$. Consequently, it can be determined how many vehicles are in the section between $x1$ and $x2$. This number can also be found in the graph, by taking the difference between the inflow and the outflow at that moment. This is indicated in the graph by the bold vertical black line.

Similarly, we can take a horizontal line; consider for instance the line at $N = 150$. The intersection with the inflow line shows when the 150th vehicle enters the section, and the intersection with the outflow line shows when this vehicle leaves the section. So, the horizontal distance between the two lines is the travel time of the 150th vehicle. At times where the demand is lower than the capacity, the vehicles have a free flow travel time. So without congestion, the outflow curve is the inflow curve which is translated to the right by the free flow travel time.

The vertical distance is the number of vehicles in the section (ΔN) at a moment t. In a time period dt this adds $\Delta N dt$ to the total travel time (each vehicle contributes dt). To get the total travel time, we integrate over all infinitesimal intervals dt:

$$tt = \int \Delta N dt \tag{2-7}$$

The horizontal distance between the two lines is the travel time for one vehicle, and vertically we find the number of vehicles. Adding up the travel times for all vehicles gives the total travel time:

$$tt = \sum_i tt_i \tag{2-8}$$

In a continuous approach, this changes into

$$tt = \int tt_i di \tag{2-9}$$

Both calculation methods lead to the same interpretation: *the total time spent can be determined by the area between the inflow and outflow curve.*

2-3-3 Delays

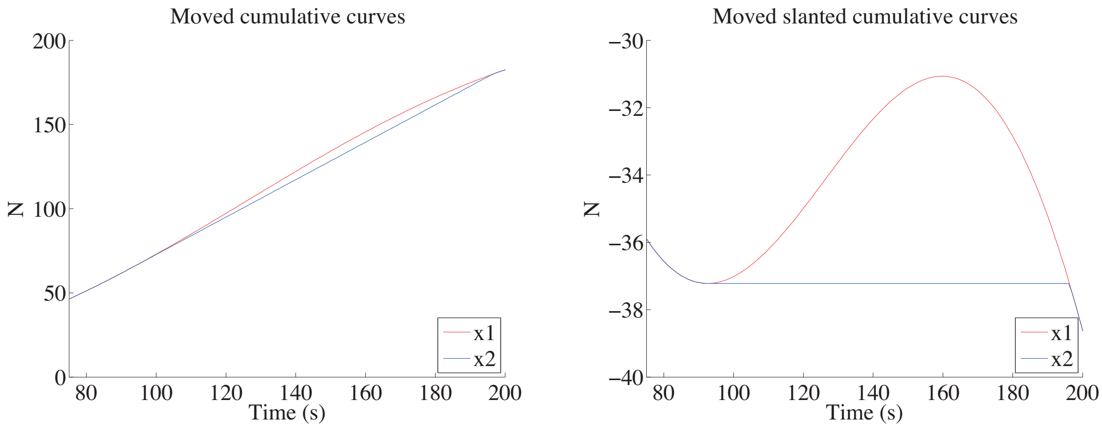
Delays for a vehicle are the extra time it needs compared to the free flow travel time; so, to calculate delay, one subtracts the free flow travel time from the actual travel time. To subtract the free flow travel time from the travel time, we can graphically move the outflow curve to the left, as is shown in figure 2-3a. For illustration purposes, the figure is zoomed at figure 2-3b. The figure shows that if the travel time equals the free flow travel time, both curves are the same, leading to 0 delay.

Similar to how the cumulative curves can be used to determine the travel time, the moved cumulative curves can be used to determine the delay. The delay for an individual vehicle can be found by the horizontal distance between the two lines. The vertical distance between the two lines can be interpreted as the number of vehicles queuing. The total delay is the area between the two lines:

$$\mathcal{D} = \int tt_i - tt_{\text{free flow}} di \tag{2-10}$$

This is the area between the two lines. If we define N_{queue} as the number of vehicles in the queue at moment t, we can also rewrite the total delay as

$$\mathcal{D} = \int N_{queue}(t) dt \tag{2-11}$$



(a) Moved cumulative curves

(b) Slanted cumulative curves

Figure 2-3: Determining the delay and the flows from cumulative curves

2-4 Slanted cumulative curves

Slanted cumulative curves or oblique cumulative curves is a very powerful yet simple tool to analyse traffic streams. These are cumulative curves which are off set by a constant flow:

$$\tilde{N} = \int q - q_0 dt - \int q_0 dt = \int q dt - \int q_0 dt \tag{2-12}$$

This means that differences with the freely chosen reference flow q_0 are amplified: in fact, only the difference with the reference flow are counted. The best choice for the reference flow q_0 is a capacity flow.

Figure 2-3b shows the slanted cumulative curves for the same situation as in figure 2-3a. The figure is off set by $q_0 = 4000$ veh/h. Because the demand is initially lower than the capacity, $\tilde{N}$ reaches a negative value. From the moment outflow equals capacity, the slanted cumulative outflow curve is constant. Since the demand is higher than the capacity, this increases. At the moment both curves intersect again, the queue is dissolved.

The vertical distance between the two lines still shows the length of the queue, N_{queue}. That means that equation 2-10 still can be applied in the same way for the slanted cumulative curves, and the delay is the area between the two lines.

Slanted cumulative curves are also particularly useful to determine capacity, and to study changes of capacity, for instance the capacity drop (see section 5-2). In that case, for one detector the slanted cumulative curves are drawn. By a change of the slope of the line a change of capacity is detected. In appendix C a Matlab code is provided by which cumulative curves can be made, and which includes the computation of several key performance indices.

2-5 Practical limitations

Cumulative curves are very useful for models where the blocking of traffic does not play a role. For calculating the delay in practise, the method is not very suitable due to failing detectors.

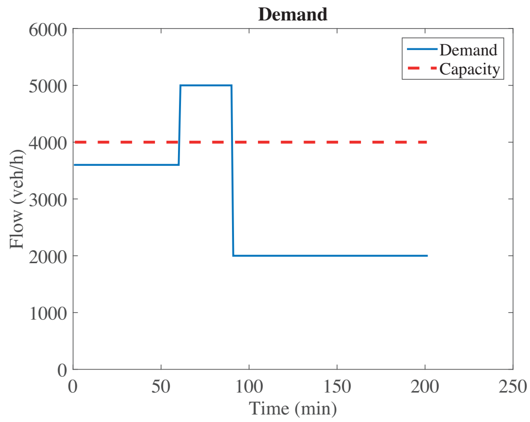


Figure 2-4: Demand and capacity

Any error in the detection (a missed or double counted observation), will change one of the curves and will offset the cumulative flow, and this is never corrected; this is called cumulative drift. Recently, an algorithm has been proposed to check the offsets by cross checking the cumulative curves with observed travel times (Van Lint et al., 2014). This is work under development. Moreover, some types of detectors will systematically miscount vehicles, which makes the above-mentioned error larger.

Apart from their use in models, slanted cumulative curves are very powerful to show changes in capacity in practise.

2-6 Example application

Consider a road with a demand of:

$$q_{\text{in}} = \begin{cases} 3600\text{v/h} & \text{for } t < 1\text{h} \\ 5000\text{v/h} & \text{for } 1\text{h} < t < 1.5\text{h} \\ 2000\text{v/h} & \text{for t} > 1.5\text{h} \end{cases} \tag{2-13}$$

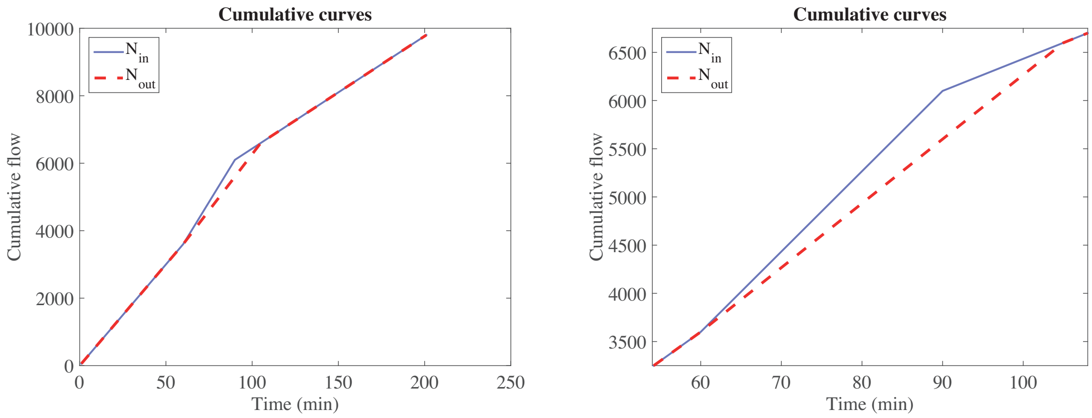
The capacity of the road is 4000 veh/h. A graph of the demand and capacity is shown in figure 2-4.

1. Construct the (translated=moved) cumulative curves
2. Calculate the first vehicle which encounters delay (N)
3. Calculate the time at which the delay is largest
4. Calculate the maximum number of vehicles in the queue
5. Calculate the vehicle number (N) with the largest delay
6. Calculate the delay this vehicle encounters (in h, or mins)
7. Calculate the time the queue is solved

8. Calculate the last vehicle (N) which encounters delay
9. Calculate the total delay (veh-h)
10. Calculate the average delay of the vehicles which are delayed (h)

This can be answered by the following:

1. For the cumulative curves, an inflow and an outflow curve needs to be constructed; both increase. For the inflow curve, the slope is equal to the demand. For the outflow curve, the slope is restricted to the capacity. During the first hour, the demand is lower than the capacity, hence *the outflow is equal to the demand*. From t=1h, the inflow exceeds the capacity and the outflow will be equal to the demand. The cumulative curve hence increases with a slope equal to the capacity. As long as there remains a queue, i.e. the cumulative inflow is higher than the outflow, the outflow remains at capacity. The outflow remains hence increasing with a slope equal to the capacity until it intersects with the cumulative inflow. Then, the outflow follows the inflow: see figure 2-5a and for a more detailed figure 2-5b.
2. The first vehicle which encounters delay (N) Delays as soon as q>C: so after 1h at 3600 v/h = 3600 vehicles.
3. The time at which the delay is largest: A queue builds up as long as q>C, so up to 1.5 h. At that moment, the delay is largest
4. The maximum number of vehicles in the queue: 0.5 h after the start of the queue, 0.5*5000=2500 veh entered the queue, and 0.5*4000=2000 left: so 500 vehicles are in the queue at t=0.5h
5. The vehicle number (N) with the largest delay: N(1.5h)=3600+0.5*5000 = 6100
6. The delay this vehicle encounters (in h, or mins): It is the 2500th vehicle after t=1h. The delay is the horizontal delay between the entry and exit curve. It takes at capacity 2500/4000 = 37.5 mins to serve 2500 vehicles. It entered 0.5 hours = 30 mins after t=1, so the delay is 7.5 mins
7. The time the queue is solved: This is the time point that the inflow and outflow curves intersect again. 500 vehicles is the maximum queue length, and it reduces with 4000-2000=2000veh/h. So 500/2000=15 minutes after the time that q<C the queue is solved, i.e. 1:45h after the start.
8. The last vehicle (N) which encounters delay This is the vehicle number at the moment the inflow and outflow curves meet again. 15 minutes after the vehicle number with the largest delay: 6100+0.25*2000 = 6600 veh
9. The total delay. This is the area of the triangle between inflow and outflow curve. This area is computed by 0.5 * height * base = 0.5 * 500 * (30+15)/60 = 187,5 veh-h. Note that here we use a generalised equation for the area of a triangle. Indeed, we transform the triangle to a triangle with a base that has the same width, and the hight which is the same for all times (i.e., we skew it). The hight of this triangle is 500 vehicles (the largest distance between the lines) and the width is 45 minutes.



(a) Overview of the cumulative curves

(b) More detailed view of the cumulative curves

Figure 2-5: Cumulative curves for the example

10. The average delay of the vehicles which are delayed (h) 187,5 veh-h/ (6600-3600) veh = 0,0625 h = 3,75 min

Selected problems

For this chapter, consider problems: 3, 4, 68, 69, 70, 71, 100, 142, 143, 144, A-7-2, 205, 205, 234, 235, 260, A-11-2

Chapter 3

Relationships of traffic variables

After this chapter, the student is able to:

- Comment on the restrictions of the fundamental diagram
- Translate the fundamental diagram in three planes
- Interpret the shape of the fundamental diagram in terms of driving behavior

Chapter 1 defined the variables and their definition. This chapter will discuss the relationship between these variables. First of all the mathematically required relationships are shown (section 3-1), then typical properties of traffic in equilibrium are discussed (section 3-2). Section 3-3 discusses these relationships in the light of drivers, and expands this to non-equilibrium conditions. Finally, section 3-4 gives attention to the moving observer.

3-1 Fundamental relationship

In microscopic view, it is obvious that the headway (h), the spacing (s) and the speed (v) are related. The headway times the speed will give the distance covered in this time, which is the spacing. It thus suffices to know two of the three basic variables to calculate the third one.

$$s = hv \tag{3-1}$$

Since headways and spacings have macroscopic counterparts, there is a macroscopic equivalent for this relationship. After reordering, equation 3-1 reads

$$\frac{1}{h} = \frac{1}{s}v \tag{3-2}$$

The macroscopic equivalent of this relationship is the average of this equation. Remembering that $q = \frac{1}{\langle h \rangle}$ and $k = \frac{1}{\langle s \rangle}$, we get:

$$\boxed{q = ku} \tag{3-3}$$

Table 3-1: The basic traffic variables and their relationship

Microscopic	Macroscopic
s	$k = \frac{1}{\langle s \rangle}$
h	$q = \frac{1}{\langle h \rangle}$
v	$u = \frac{1}{\langle v \rangle}$
$s = hv$	$q = ku$

(a) One extreme: an empty road

(b) One extreme: a jammed road

(c) Basic shape of the fundamental diagram

Figure 3-1: The extreme situations and an idea for the fundamental diagram

This equation shows that the flow q is proportional with both the speed u and the density k. Intuitively, this makes sense because when the whole traffic stream moves twice as fast if the flow doubles. Similarly, if – at original speed – the density doubles, the flow doubles as well.

Table 3-1 summarizes the variables and their relationships.

3-2 Fundamental diagram

If two of the three macroscopic traffic flow variables are known, the third one can be calculated. This section will show that there is another relationship. In fact, there is an equilibrium relationship between the speed and the density. First, a qualitative understanding will be given, after that the effect will be shown for various couples of variables. Also, different shapes of the supposed relationship will be shown (section 3-2-5).

3-2-1 Qualitative understanding of the shape

Let us, for the sake of argument, consider the relationship between density and flow. And let us furthermore start considering the most extreme cases. First, the case that there is no vehicle on the road. Since the density is 0, the flow is 0, according to equation 3-3. In the other extreme case the density on the road is very high, and the speed is 0. Using again equation 3-3 we find also for this case a flow of 0. In between, there are traffic states for which the traffic flow is larger than zero. Assuming a continuous relationship between the speed and the density (which is not necessarily true, as will be discussed section 5-2) there will be a curve relating the two points at flow 0. This is indicated in figure 3-1c.

This relationship is being observed in traffic. However, it is important to note that this is *not* a causal relationship. One might argue that due to the low speed, drivers will drive closer together. Alternatively, one might argue that due to the close spacing, drivers need to slow down.

3-2-2 Traffic state

We can define a traffic state by its density, flow and speed. Using equation 3-3, we only need to specify two of the variables. Furthermore, using the fundamental diagram, one can be sufficient. It is required that the specified variable then has a unique relationship to the others. For instance, judged by figure 3-1c, specifying the density will lead to a unique flow, and a unique speed (using equation 3-3, and thus a unique traffic state). However, specifying the flow (at any value between 0 and the capacity) will lead to two possible densities, two possible speeds, and hence two possible traffic states.

The speed of the traffic can be derived using the equation 3-3:

$$u = \frac{q}{k} \tag{3-4}$$

For a traffic state in the flow density plane, we can draw a line from the traffic state to the origin. The slope of this line is q/k. So the speed of the traffic can be found by the slope of a line connecting the origin to the traffic state in the flow density plane. The free flow speed can be found by the slope of the fundamental diagram at k=0, i.e. the derivative of the fundamental diagram in the origin.

3-2-3 Important points

The most important aspect of the fundamental diagram for practitioners is the capacity. This is the maximum flow which can be maintained for a while at a road. The same word is also used for the traffic state at which maximum flow is obtained. This point is found at the top of the fundamental diagram. Since we know that the flow can be determined from the headway, we can estimate a value for the capacity if we consider the minimum headway. For drivers on a motorway, the minimum headway is approximately 1.5 to 2 seconds, so we find a typical capacity value of $\frac{1}{2}$ to $\frac{1}{1.5}$ vehicles per second. If we convert this to vehicles per hour, we find (there are 3600 seconds in an hour) $\frac{3600}{2} = 1800$ to $\frac{3600}{1.5} = 2400$ vehicles per hour.

The density for this point is called the critical density, and the related speed the critical speed. The capacity is found when the average headway is shortest, which is when a large part of the vehicles is in car-following mode. This happens at speeds of typically 80 km/h; this then is the critical speed. From the capacity and the critical speed, the critical density can be calculated using equation 3-3. This varies from typically 20 veh/km/lane to 28 veh/km/lane.

For densities lower then the critical density, traffic is in an uncongested state; for higher densities, traffic is in a congested state. In the uncongested part, the traffic flow increases with an increase of density. In the congested branch, the traffic flow decreases with an increase of density. The part of the fundamental diagram of uncongested traffic states is called the uncongested branch of the diagram. Similarly, the congested branch gives the points for which the traffic state is congested.

The free flow speed is the speed of the vehicles at zero density. At the other end, we find the density at which the vehicles come to a complete stop, which is called the jam density. For the jam density, we can also make an estimation based on the length of the vehicles and the distance they keep at standstill. A vehicle is approximately 5 meters long, and they keep some distance even at standstill (2-3 meters), which means the jam density is $\frac{1}{5+3}$ to $\frac{1}{5+2}$ veh/m, or $\frac{1000}{5+3}$=125veh/km to $\frac{1000}{5+2}$=142 veh/km.

3-2-4 Fundamental diagram in different planes

So far, the fundamental diagram has only be presented in the flow density plane. However, since the fundamental equation (equation 3-3) relates the three variables to each other, any function relating two of the three variables to each other will have the same effect. Stated otherwise, the fundamental relationship can be presented as flow-density relationship, but also as speed-density relationship or speed-flow relationship. Figure 3-2 shows all three representations of the fundamental diagram for a variety of functional forms.

In the speed-density plane, one can observe the high speeds for low densities, and the speed gradually decreasing with increasing density. In the speed-flow diagram, one sees two branches: the congested branch with high speeds and high flows, and also a congested branch with a low speed and lower flows.

3-2-5 Shapes of the fundamental diagram

There are many shapes proposed for the fundamental diagram. The data are quite scattered, so different approaches have been taken: very simple functions, functions with mathematically useful properties, or functions derived from a microscopic point of view. Even today, new shapes are proposed. In the remainder of this section, we will show some elementary shapes; the graphs are shown in figure 3-2.

Greenshields

Greenshields was the first to observe traffic flows and publish on this in 1934 (Greenshields, 1934). He observed a platoon of vehicles and checked the density of the platoon and their speed. He assumed this relationship to be linear:

$$v = v_0 - ck \tag{3-5}$$

Note that for $k = \frac{v}{c}$ the speed equals 0, hence the flow equals zero, so the jam density equals $\frac{c}{v_0}$.

Triangular

The Greenshields diagram is not completely realistic since for a range of low densities, drivers keep the same speed, possibly limited by the current speed limit. The fundamental diagram which is often used in academia is the triangular fundamental diagram, reffering to the triangular shape in the flow-density plane. The equation is as follows:

$$q = \begin{cases} v_0 k & \text{if } k < k_c \\ q_c - \dfrac{k - k_c}{k_j - k_c} q_c & \text{if } k \geq k_c \end{cases} \tag{3-6}$$

Truncated triangular

Daganzo (1997) shows a truncated triangular fundamental diagram. That means that the flow is constant and maximized for a certain range of densities. The equation is as follows:

$$q = \begin{cases} v_0 k & \text{if } k < k_1 \\ v_0 k_1 & \text{if } k_1 < k < k_c \\ q_c - \dfrac{k - k_c}{k_j - k_c} q_c & \text{if } k \geq k_c \end{cases} \tag{3-7}$$

Smulders

Smulders (1989b) proposed a fundamental diagram in which the speed decreases linearly with the density for the free flow branch. In the congested branch the flow decreases linearly with density.

Drake

Drake et al. (1967) proposes a continuous fundamental diagram where the speed is an exponentially decreasing function of the density:

$$v = v_0 \exp\left(-\frac{1}{2}\left(\frac{k}{k_c}\right)^2\right) \tag{3-8}$$

Inverse lambda

The capacity drop (see section 5-2) is not present in the fundamental diagrams presented above. Koshi et al. (1981) introduced an 'inverse lambda'fundamental diagram. This means the traffic has a free speed up to a capacity point. The congested branch however, does not start at capacity but connects a bit lower at the free flow branch. It is assumed that traffic remains in the free flow branch and after congestion has set in, will move to the congested branch. Only after the congestion has solved, passing a density lower than the density where the congested branch connects to the free flow branch, traffic flows can grow again to higher values. The description is as follows:

$$q = \begin{cases} v_0 k & \text{if } k < k_c \\ v_0 k_1 - \dfrac{k - k_1}{k_j - k_1} v_0 k_1 & \text{if } k \geq k_1 \text{and traffic is congested} \end{cases} \tag{3-9}$$

This shape of the fundamental diagram allows for two traffic states with similar densities but different flows. This can yield unrealistic solutions to the kinematic wave model (see section 4).

Wu

An addition to the inverse-lambda fundamental diagram is made by Wu (2002). He assumes the speed in the free flow branch to decrease with increasing density. The shape of the free flow branch is determined by the overtaking opportunities, which in turn depend on the number of lanes, l. The equation for the speed is

$$q = \begin{cases} k\left(1-\left(\frac{k}{k_1}\right)^{l-1} * v_0 + \left(\frac{k}{k_1}\right)^{l-1} v_p\right) & \text{if } v > u_p \\ v_p k_1 - \frac{k-k_1}{k_j-k_1} v_0 k_1 & \text{if } k \geq k_1 \text{and traffic is congested} \end{cases} \tag{3-10}$$

Kerner

Kerner (2004b) has proposed a different theory on traffic flow, the so-called three phase traffic flow theory. This will be described in more detail in section 5-4. For here, it is important to note that the congested branch in the three phase traffic flow theory is not a line, but an area.

3-3 Microscopic behaviour

Section 3-2 showed the equilibrium relationships observed in traffic. This is a result of behavior, which can be described at the level of individual drivers as well. This section does so. First, the equilibrium behaviour is described in section 3-3-1. Then, section 3-3-2 discusses hysteresis, i.e. structural off-equilibrium behaviour under certain conditions.

3-3-1 Equilibrium behaviour

The fundamental diagram describes traffic in equilibrium conditions. That can happen if all drivers are driving in equilibrium conditions, i.e. all drivers are driving at a headway which matches a speed. Using the relationships in table 3-1 one can change a fundamental diagram on an aggregated level to a fundamental diagram on an individual level. This way, one can relate individual headways to individual speeds.

The fundamental diagram gives the average distance drivers keep. However, there is a large variation in drivers' behaviour. Some keep a larger headway, and some drivers keep a smaller headway for the same speed. These effects average out in a fundamental diagram, since the average headway for a certain speed is used. On an individual basis, there is a much larger spread in behaviour.

3-3-2 Hysteresis

Apart from the variation between drivers, there is also a variation within a driver for a distance it keeps at a certain speed (which we assume in this section as representative of the fundamental diagram). These can be random variations, but there are also some structural

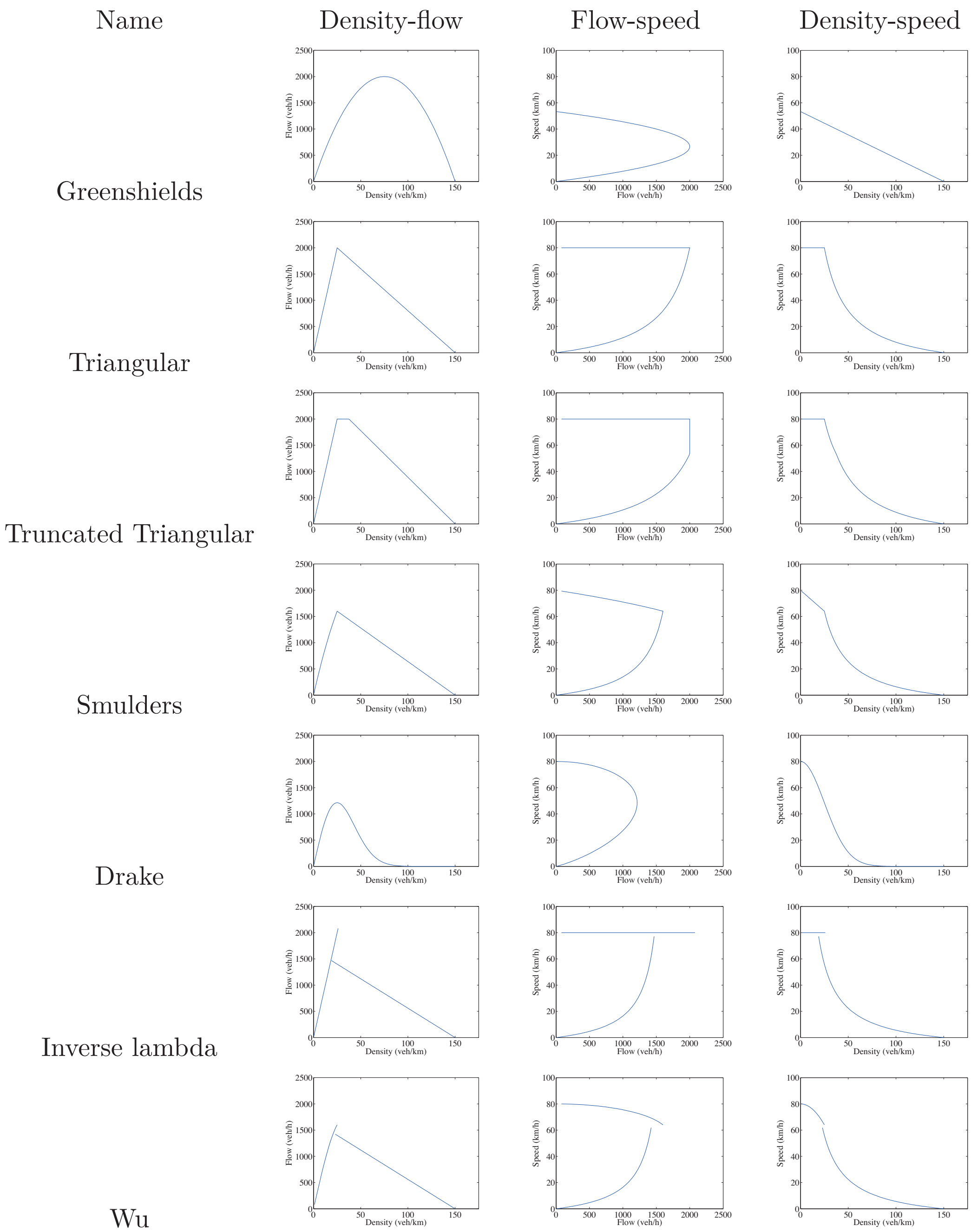
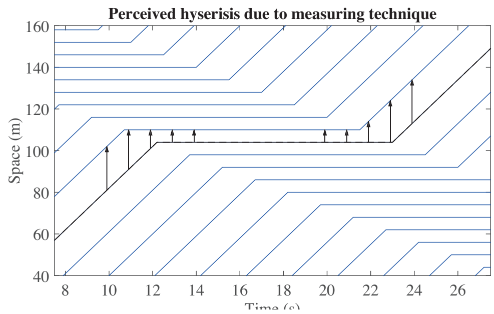
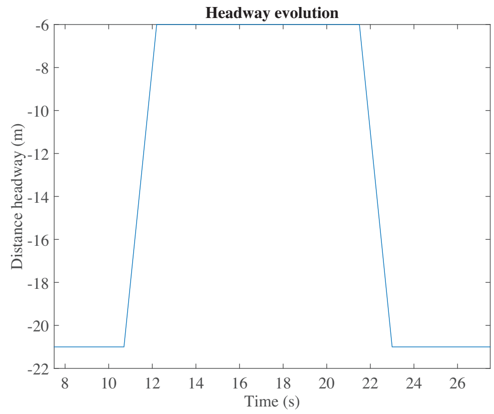


Figure 3-2: Different shapes of the fundamental diagram



(a) Trajectories



(b) Headways

Figure 3-3: Mistakes if making the headway analysis at one moment in time

variations. Usually the term Hysteresis is used to indicate that the driving behaviour (i.e. the distance) is different for drivers before they enter the congestion compared to after they come out of congestion. That is, the distance *at the same speed* is different in each of these conditions.

Two phenomena might play a role here:

1. Delayed reaction to a change of speed
2. Anticipation of a change in speed

Zhang (1999) provides an excellent introduction to hysteresis. The simplified reasoning is as follows.

Let's first discuss case 1, drivers have a delayed reaction to a change of speed. That means that when driving at a speed, first the speed of the leader reduces, then the distance reduces. So during the deceleration process, the headway is shorter than the equilibrium headway. When the congestion solves, first the leader will accelerate, and the driver will react late on that. That means that the leader will shy away from the considered car, and the distance will be larger than the equilibrium distance.

In case 2, if the driver anticipates the change in speed, the exact opposite happens. Before the deceleration actually happens, a driver will already decrease speed (by definition in anticipation), leading to a larger headway than the equilibrium headway for a certain speed. Under acceleration, the opposite happens, and a driver can already accelerate before that would be suitable in case of equilibrium conditions. Hence, in the acceleration phase, the driver has a shorter headway than in equilibrium conditions for the same speed.

In traffic, we expect drivers to have a reaction time. In fact, the reaction time can be derived from the fundamental diagram, as section 6-1 will show later on. It will also show that the best way to analyse car-following behaviour is not comparing the distance-speed relationship for one pair *at one moment in time*, as shown in figure 3-3. It shows that the drivers have no hysteresis – they copy the movement of the leader perfectly – but still the gap changes with a constant speed. Instead, one should make the analysis of car-following behavior *along the axis parallel to the wave speed.* Laval (2011) provides a very good insight in the differences one can obtain using this correct technique or using the (erroneous) comparison of instantaneous headways (as in the arrows in figure 3-3b).

3-4 Moving observer

An observer will only observe what is in the observation range. Many observations are taken at a point in space . This point might move with time, for instance a driver might check the number of trucks he is overtaking. In this case, the driver is moving and observing, this is called a moving observer. This section discusses the effects of the speed of the moving observer, and also discusses the effect of observed subjects passing a stationary observer with different speeds.

Basically, the movement at speed υ has no effect on density, but the relative speed of traffic changes. Therefore, applying equation 3-3, the relative flow changes. Written down explicitly, one obtains a relative flow q_{rel}:

$$q_{\text{rel}} = k\,(v - \upsilon) \tag{3-11}$$

If the observer moves with the speed of the traffic, the observed relative flow becomes zero. In practice if not all vehicles drive at the same speed, the flow needs to be divided into classes with the same speed and the partial flows for each of these classes should be calculated.

$$q_{\text{rel}} = \sum_{v} \left(k_{\text{Vehicles at speed } v}\,(v - \upsilon)\right) \tag{3-12}$$

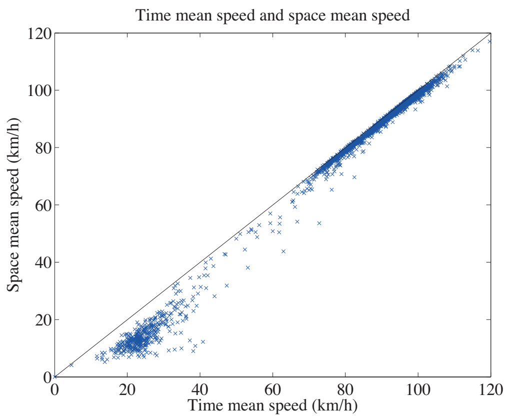
Local measurements

Suppose there is a local detector located at location x_{detector}. Now we reconstruct which vehicles will pass in the time of one aggregation period. For this to happen, the vehicle i must be closer to the detector than the distance it travels in the aggregation time:

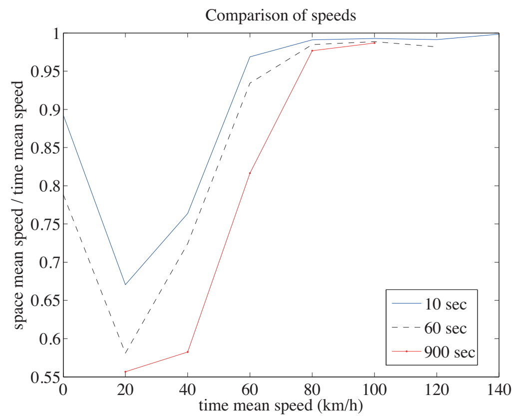
$$x_{\text{detector}} - x_i \leq t_{\text{agg}} v_i \tag{3-13}$$

In this formula, x is the position on the road and t_{agg} the aggregation time. For faster vehicles, this distance is larger. Therefore, if one takes the local (arithmetic) mean, one overestimates the influence of the faster vehicles. If the influence of the faster vehicles on speeds is overestimated, the average speed u_t is overestimated (compared to the space-mean speed u_s).

The discussion above might be conceived as academic. However, if we look at empirical data, then the differences between the time-mean speeds and space-mean speeds become apparent. Figure 3-4a shows an example where the time-mean speed and space-mean speed have been computed from motorway individual vehicle data collected on the A9 motorway near Amsterdam, The Netherlands. Figure 3-4b shows that the time mean speed can be twice as high as the space mean speed. Also note that the space-mean speeds are always lower than the time-mean speeds.



(a) The arithmetic mean and the harmonic mean of the speeds of the vehicles passing a cross section of a motorway



(b) The arithmetic mean and the harmonic mean of the speeds of the vehicles passing a cross section of a motorway

Figure 3-4: The effect of inhomogeneities in speeds: the difference in arithmetic mean and harmonic mean speed; based on Knoop et al. (2007)

In countries where inductive loops are used to monitor traffic flow operations and arithmetic mean speeds are computed and stored, average speeds are overestimated, affecting travel time estimations. Namely, to estimate the average travel time, the average of $TT = L/v$ is needed, in which L is the length of the road stretch and v the travel speed. Since L is constant, the average travel time can be expressed as

$$\langle TT \rangle = L \left\langle \frac{1}{v} \right\rangle \tag{3-14}$$

From this, it follows that the harmonically averaged speed is required for estimating the mean travel time.

Furthermore, since $q = ku$ (equation 3-3) can only be used for space-mean speeds, we cannot determine the density k from the local speed and flow measurements, complicating the use of the collected data for, e.g. traffic information and traffic management purposes. As figure 3-4b already shows, largest relative deviation is found at the lower speeds. In absolute terms, this is not too much, so one might argue this is not important. However, a low speed means a high density or a large travel time. For estimating the densities, this speed averaging has therefore a large effect, as can be seen in figure 3-5.

Selected problems

For this chapter, consider problems: 1, 66, 90, A-5-3, 140, 141, 155, 156, 176, 177, 178, 188, 195, 194, 197, 215, 223, 224, 245, 255, A-10-5, 258, 285, 286, 290

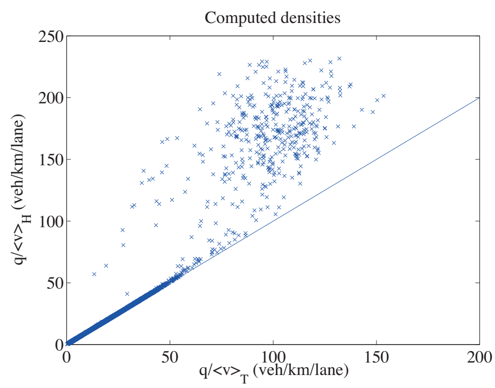


Figure 3-5: Densities computed using arithmic mean speeds or harmonic mean speeds; based on Knoop et al. (2007).

Chapter 4

Shock wave theory

After this chapter, the student is able to:

- Construct the traffic dynamics in space-time for a given demand profile and one (or more) stationary bottleneck(s), or give properties of traffic given the traffic in space-time

This chapter describes shock wave theory to capture queuing dynamics. This differs from cumulative curves in the way that the spatial extent of the queue is considered. In this section, fixed bottlenecks are discussed. At fixed bottleneck some lanes of a highway are (temporarily) blocked. Firstly the theory and derivation of equations will be discussed (section 4-1). Thereafter, two examples are given (Section 4-2 and 4-3). Then section 4-4 describes the stop-and-go waves in this modelling framework. This framework can also be applied to moving observers, which is presented in chapter 16.

4-1 Theory and derivation of equations

Let us consider a situation with two different states: state A downstream, with a matching q_A, k_A and v_A, and state B upstream (q_B, k_B and v_B). The states are plotted in the space-time diagram 4-1. We choose the axis in such a way that the shockwave moves through the point t=0 at x=0. We will now derive the equation to get the speed of this wave.

In the derivation, we base the reasoning on figure 4-1. The boundary between is called a shock wave. This wave indicates where the speed of the vehicles changes. It is important to note that there are no vehicles captured in the wave itself: the wave itself does not have a physical length. Thus the assumption is that vehicles change speed instantaneously.

Because there are no vehicles in the wave, the number of vehicles entering the wave must be equal to the number of vehicles exiting the wave, or we can state that rate of vehicles entering the wave must be equal to the rate of vehicles exiting the wave. This principle, in combination

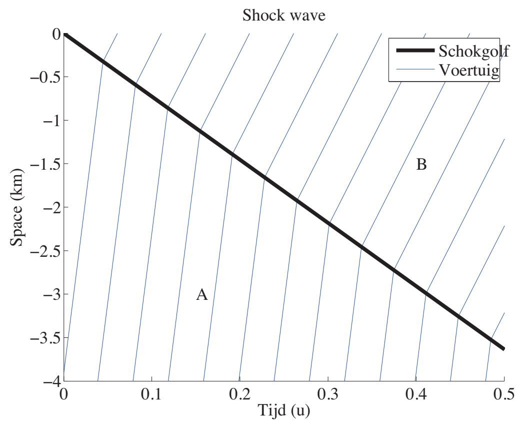


Figure 4-1: A shockwave where traffic speed changes from high to low.

with the following equation (already presented in equation 3-3) is used to calculate the speed of the wave:

$$q = kv \tag{4-1}$$

The speed of the wave is indicated by w. Note that we can apply this equation to moving frames of reference as well. In that case, the flow changes, as does the speed. The density is invariant under a change of reference speed.

To determine the attachment and exit rate, we will move with the speed of the wave (in the frame of a moving observer). At the downstream end, the density is k_B. The speed in the moving frame of reference is $v_B - w$. The exit rate in the moving frame of reference is calculated using equation 4-1.

$$q_{\text{exit}} = k_B \left(v_B - w\right) \tag{4-2}$$

In the same way, the attachment rate can be determined. The upstream density is k_A. The speed of the vehicles in the frame of reference moving with the wave speed w is $v_A + w$. Using equation 4-1 again, we find the attachment rate in this moving frame of reference:

$$q_{\text{attachment}} = k_A \left(v_A - w\right) \tag{4-3}$$

Since these rates have to be equal, we find:

$$q_{\text{exit}} = q_{\text{attachment}} \tag{4-4}$$
$$k_A \left(v_A - w\right) = k_B \left(v_B - w\right) \tag{4-5}$$

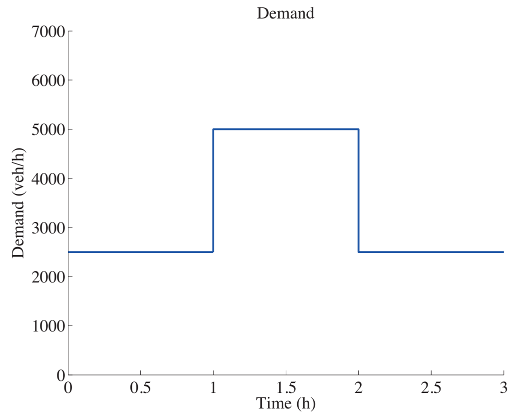
This can be rewritten as:

$$k_A v_A - k_A w = k_B v_B - k_B w \tag{4-6}$$
$$q_A - k_A w = q_B - k_B w \tag{4-7}$$

In the last step, the speeds have been substituted using equation 4-1. We can solve this equation for the shock wave speed w. We find

$$q_A - q_B = \left(-k_B + k_A\right) w \tag{4-8}$$

(a) The road



(b) Demand pattern

Figure 4-2: Situation

And isolating w gives the wave speed equation:

$$\boxed{w = \frac{q_A - q_B}{k_A - k_B} = \frac{\Delta q}{\Delta k}} \tag{4-9}$$

Note that in a space-time plot, the speed w is the slope of the shock wave between A and B. The right hand side is the ratio between the difference in flow and the difference in density of states A and B. This is also the slope of a line segment between A and B in the flow-density plot. This becomes very useful when constructing the traffic states.

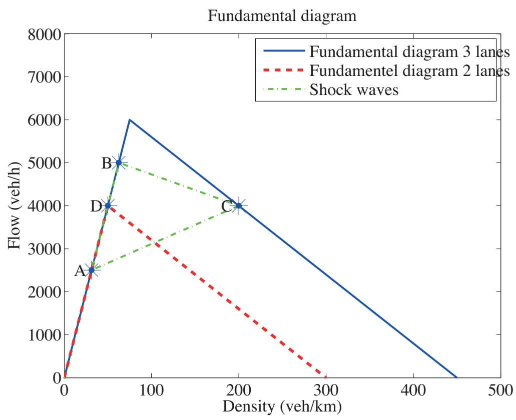
The above reasoning holds for the speed of any shock wave, moving backward or forward. The following section shows an example for both.

4-2 Example: Temporal increase in demand at a road with a lane drop

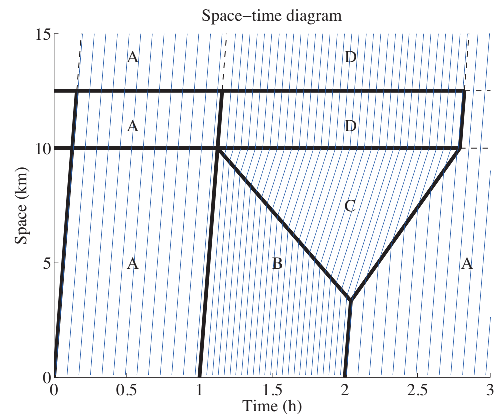
Let's consider a 3 lane road with a reduction to 2 lanes over a 1 km section between x=10 and x=12.5 (see figure 4-2b). For the road, we assume lanes with equal characteristics, described by a triangular fundamental diagram with a free speed of 80 km/h, a capacity of 2000 veh/h/lane and a jam density of 150 veh/km/lane. At the start of the road, there is a demand of 2500 veh/h which temporarily increases to 5000 veh/h between t=1h and t=2h (see the demand profile in figure 4-2b). The question we will answer in this example is: *What are the resulting traffic conditions?*

The final answer to the question are the traffic states which are shown in table 4-1, and shown on the fundamental diagram in figure 4-3a. The speed of the shock waves is given in table 4-2, and the resulting traffic situation is shown in figure 4-3b. We will now explain how this solution can be found.

The inflow to the system is given, being 2500 veh/h (state A) and 5000 veh/h (state B) on a three lane road. The matching densities can be computed from the fundamental diagram for



(a) Position of traffic states at the fundamental diagram



(b) Shock waves in space-time

Figure 4-3: The situation

Table 4-1: The states on the road

State	Flow (ven/h)	Density (veh/km)	Speed (km/h)
A	2500	31.25	80
B	5000	62.5	80
C	4000	200	20
D	4000	50	80

Table 4-2: The shock waves present on the road

State 1	State 2	shock wave speed w (km/h)
A	B	80
B	C	-7.3
A	C	8.9

a three lane road. This gives densities of (equation 4-1) $k_A = \frac{q_A}{u_A} = \frac{2500}{80} = 31.25$veh/km and $k_B = \frac{q_B}{u_B} = \frac{5000}{80} = 62.5$veh/km. The separation with the empty road moves forward with a speed of 80 km/h, i.e. the speed of the vehicles. This can also be found by the shock wave equation, equation 4-9. The difference in flow is 2500 veh/h and the difference in density is 31.25 veh/km. The shock wave then moves with:

$$w_{0A} = \frac{q_0 - q_A}{k_0 - k_A} = \frac{-2500}{-31.25} = 80km/h \tag{4-10}$$

Note that it does not matter whether the speed of shock wave w_{0A} or shock wave w_{A0} is calculated. This remark holds for every combination of states.

Also the wave between state A and B moves forward with 80 km/h, calculated by equation 4-9:

$$w_{AB} = \frac{q_A - q_B}{k_A - k_B} = \frac{2500 - 5000}{31.25 - 62.5} = 80km/h \tag{4-11}$$

Note that this equals the free flow speed. This is because we use a triangular fundamental diagram. The speed w_{ab} is the direction of the line segment AB in figure 4-3a. This has the same slope as the slope of the free flow speed (i.e. the slope of the fundamental diagram at the origin) because its shape is triangular.

Now this wave hits the two lane section. The flow is higher then the capacity of the two lane section. That means that downstream, the road will operate at capacity, and upstream a queue will form (i.e. we have a congested state). The capacity of the downstream part is, according to the fundamental diagram, 4000 veh/h (state D). The speed follows from the (triangular) fundamental diagram, and is 80 km/h. The density hence is:

$$k_D = \frac{q_D}{u_D} = \frac{4000}{80} = 50veh/km \tag{4-12}$$

If 4000 veh/h drive onto the two lane segment, 4000 veh/h have to drive off the three lane segment (no vehicles can be lost or created at the transition from three to two lanes). That means that upstream of the transition, we have a congested state with a flow of 4000 veh/h. The density is derived from the fundamental diagram:

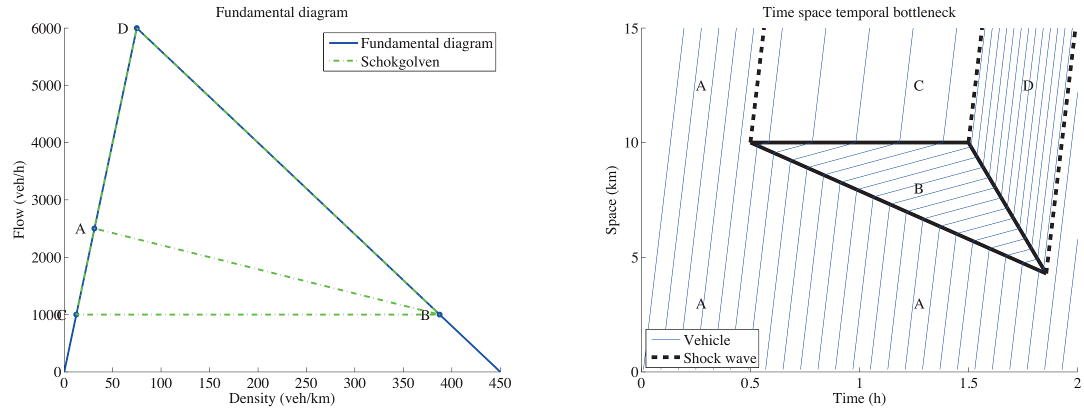
$$q_{\text{cong}} = q_{\text{capacity}} - q_{\text{capacity}} \frac{k - k_c}{k_j - k_c} = 4000 \tag{4-13}$$

Substituting the parameters for the fundamental diagram, and realising that the capacity is found by $q_{\text{capacity}} = l u_{\text{capacity}} k_c$ (l is the number of lanes) we calculate the density at point C, 200 veh/km. Graphically, we can find point C on the fundamental diagram by the intersection of the congested branch and a line at a constant flow value of 4000 veh/h.

The speed at which the tail of the queue moves backwards, i.e., the speed of the boundary between B and C, is calculated by the shock wave equation (equation 4-9)

$$w_{BC} = \frac{q_B - q_C}{k_B - k_C} = \frac{5000 - 4000}{62.5 - 200} = -7.3\text{km/h} \tag{4-14}$$

Note that this speed can also be derived graphically from the fundamental diagram, i.e. the slope of the line segment BC. That the slope in figure 4-3a and figure 4-3b is graphically not the same, which is a consequence of different axis scales in the figures.



(a) Position of traffic states at the fundamental diagram **(b)** Shock waves in space-time

Figure 4-4: The situation

Then the demand reduces again and the inflow state returns to A. When lower demand hits the tail of the jam, the queue can solve from the tail. At the head of the queue, this change has no influence yet, since drivers are still waiting to get out onto the smaller roadway segment. The boundary between C and A moves with a speed of

$$w_{CA} = \frac{q_C - q_A}{k_C - k_A} = \frac{4000 - 2500}{200 - 31.25} = 8.9\text{km/h} \tag{4-15}$$

Here again, the speed could also be derived graphically, by the slope of the line segment AC.

Then, this wave arrives at the transition of the two lane road to a three lane road, and the congestion state C is dissolved. In the two lane part, the flow is now the same as the demand (the state in A, 2500 veh/h). The boundary between state D and A moves forward with a speed of

$$w_{DA} = \frac{q_D - q_A}{k_D - k_A} = \frac{4000 - 2500}{50 - 31.25} = 80\text{km/h} \tag{4-16}$$

4-3 Example: Temporal capacity reduction

Another typical situation is a road with a temporal local reduction of capacity, for instance due to an accident. This section explains which traffic situation will result from that.

The case is as follows. Consider a three-lane freeway, with a triangular fundamental diagram. The free flow speed is 80 km/h, the capacity is 2000 veh/h/lane and the jam density is 150 veh/km/lane. The demand is constant at 2500 veh/h. From t=0.5h to t=1.5h, an incident occurs at x=10, limiting the capacity to 1000 veh/h. Calculate the traffic states and the shock waves, and draw them in the space-time diagram. Also draw several vehicle trajectories.

For referring to certain states, we will first show the resulting states, and then explain how these states are constructed. Figure 4-4a shows the fundamental diagram and the occurring states, 4-4b shows how the states move in space and time. The details of the states can be found in table 4-3, and the details of the shock waves can be found in table 4-4.

Table 4-3: The states on the road with a temporal bottleneck

Number	Flow	Density	Speed
A	2500	31.25	80
B	1000	387.5	
C	1000	12.5	12.5
D	6000	75	75

Table 4-4: The shock waves present on the road with a temporal bottleneck

State 1	State 2	shock wave speed w (km/h)
A	C	80
B	C	0
A	B	-4.2
B	D	-16

At the start, there are free flow conditions (state A) at an inflow of 2500 veh/h. For the assumed triangular fundamental diagram, the speed for uncongested conditions is equal to the free flow speed, so 80 km/h. The matching density can be found by applying equation 4-1: $k_A = \frac{q}{u} = \frac{2500}{80} = 31.25$ veh/km.

From t=0.5h to t=1.5h, a flow limiting condition is introduced. We draw this in the space time diagram. The flow is too high to pass the bottleneck, so the moment the bottleneck occurs, a congested state (B) will form upstream. Downstream of the bottleneck we find uncongested conditions (once the vehicles have passed the bottleneck, there is no further restriction in their progress): state C. For state C, the flow equals the flow that can pass the bottleneck, which is given at 1000 veh/h. The speed is the free flow speed of 80 km/h, so the matching density can be found by applying equation 4-1: $k_C = \frac{q}{u} = \frac{1000}{80} = 12.5$ veh/km. The speed of the shock wave between state A and C can be calculated using the shock wave equation, 4-9:

$$w_{AC} = \frac{q_A - q_B}{k_A - k_B} = \frac{2500 - 1000}{31.25 - 12.5} = 80 km/h \tag{4-17}$$

This equals the free flow speed. Graphically, we understand this because both states can be found at the free flow branch of the fundamental diagram (figure 4-4a) and the shock wave speed is the slope of the line segment connecting these states. Because the fundamental diagram is triangular, this slope is equal to the slope at the origin (i.e., the free flow speed).

Upstream of the bottleneck a congested state forms (B). The flow in this area must be the same as the flow which can pass the bottleneck. This is because at the bottleneck no new vehicles can be formed. That means state B is a congested state with a flow of 1000 veh/h. From the fundamental diagram we find the matching density in the congested branch, 387.5 veh/km. The speed at which the shock between states A and B now moves, can be calculated using equation 4-9:

$$w_{AB} = \frac{q_A - q_B}{k_A - k_B} = \frac{2500 - 1000}{31.25 - 387.5} = -4.2\text{km/h} \tag{4-18}$$

The minus sign means that the shock wave moves in the opposite direction of the traffic, upstream. We could also graphically derive the speed of the shock wave by the slope of the line between point A and B in the fundamental diagram.

Once the temporal bottleneck has been removed, the vehicles can drive out of the queue: state D. Because state C is congested, the outflow will be capacity (or the queue discharge rate in case there is a capacity drop). That is for this case a flow of 3x2000 veh/h = 6000 veh/h. Realising the vehicle speed equals the free flow speed of 80 km/h, the density can be found using equation 4-1:

$$k_D = \frac{q_D}{u_D} = \frac{6000}{80} = 75\text{km/h} \tag{4-19}$$

The shock wave between state B and D moves backward. The speed thereof can be found by applying equation 4-9

$$w_{BD} = \frac{q_B - q_D}{k_B - k_D} = \frac{1000 - 6000}{387.5 - 75} = -16\text{km/h} \tag{4-20}$$

The negative shock wave speed means the wave moves upstream. Intuitively, this is right, since the vehicles at the head of the queue can accelerate out of the queue, and thus the head moves backwards.

4-4 Stop and go waves

On motorways, often so called stop-and-go waves occur. These short traffic jams start from local instabilities, and the speed (and flow) in the stop-and-go waves is almost zero. For a more detailed explanation, see section 5-3. Now the speed of the boundaries of the traffic states are known, we can apply this to stop-and-go waves, and understand why this typical pattern arises (figure 4-5).

Stop-and-go waves arise in dense traffic. The traffic demand is then mostly near capacity, or at approximately the level of the queue discharge rate. The upstream boundary of such a shock moves upstream. When it is in congested conditions, both the upstream state (the congested condition) and the downstream state (the standing traffic) are congested, hence the upstream boundary moves backwards with a speed equal to the wave speed of the fundamental diagram. Once it gets out of congestion, the inflow is most likely approximately equal to the capacity of the road. That means that the upstream boundary moves with a speed which is equal to the slope of the line in the fundamental diagram connecting the capacity point with the point of jam density, which is the wave speed.

The downstream boundary separates the jam state (standing traffic) with capacity (by default: vehicles are waiting to get out of the jam, hence a capacity state occurs). The speed between these two states is found by connecting the points in the fundamental diagram. The resulting speed is the wave speed of the fundamental diagram.

After the first stop-and-go wave has moved upstream, the inflow in the second stop-and-go wave equals the outflow of the first stop-and-go wave. These are the upstream respectively the downstream state of the second wave, which thus are the same. In between, within the wave, there is another, jammed state. According to shock wave theory the shock between state A and B (i.e., the upstream state and the state within the stop-and-go wave) and the shock between B and A (i.e. the state within the stop-and-go wave and the downstream state, which equals the upstream state) is the same. This speed is approximately equal for all roads, 15-20 km/h (Schreiter et al., 2010). Hence, the length of the stop-and-go wave remains the same.

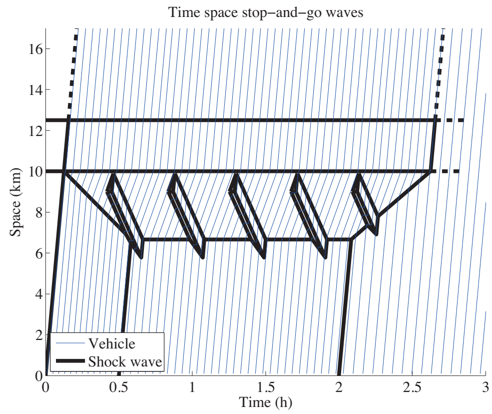


Figure 4-5: Stop and go waves in time and space

This way, stop-and-go waves can travel long distances upstream. Section 5-4 will discuss the stop-and-go waves and their characteristics in more detail.

Selected problems

This chapter is a part of what was previously a combination of fixed bottlenecks (this chapter) and moving bottlenecks (now chapter 16). The following exercises mostly start with a fixed bottleneck, but at some point include a moving bottleneck. Students are supposed to be able to answer these questions up to the moving bottleneck. These problems are: A-1-4, A-2-6, 67, 72, A-3-4, A-4-4, 101, 102, A-5-4, A-6-3, A-7-4, A-8-3, A-8-4, A-9-3, 259, 275, 287.

Chapter 5

Traffic states and Phenomena

After this chapter, the student is able to:

- Comment on different levels of stability for traffic
- Recognise traffic states from traffic measurements, and derive the causes for the observed traffic states
- Comment on the capacity drop, and show this in empirical data
- Comment on the three states of traffic according to Kerner's theory

This chapter discusses some of the important phenomena in traffic.

5-1 Stability

In traffic, we can differentiate between three levels of stability: local, platoon, and traffic. This is indicated in table 5-1, and shown graphically in figure 5-1.

Table 5-1: Identification scheme for different levels of stability

Name	relevant vehicles	criterion
Local	One vehicle pair	Does the follower relax into a steady state speed after a speed change of the leader?
Platoon	A platoon of vehilces	Does the change of speed of a leader increase over the vehicle number in the platoon?
Traffic	More than one platoon	Does the speed disturbance propagate to the next platoon?

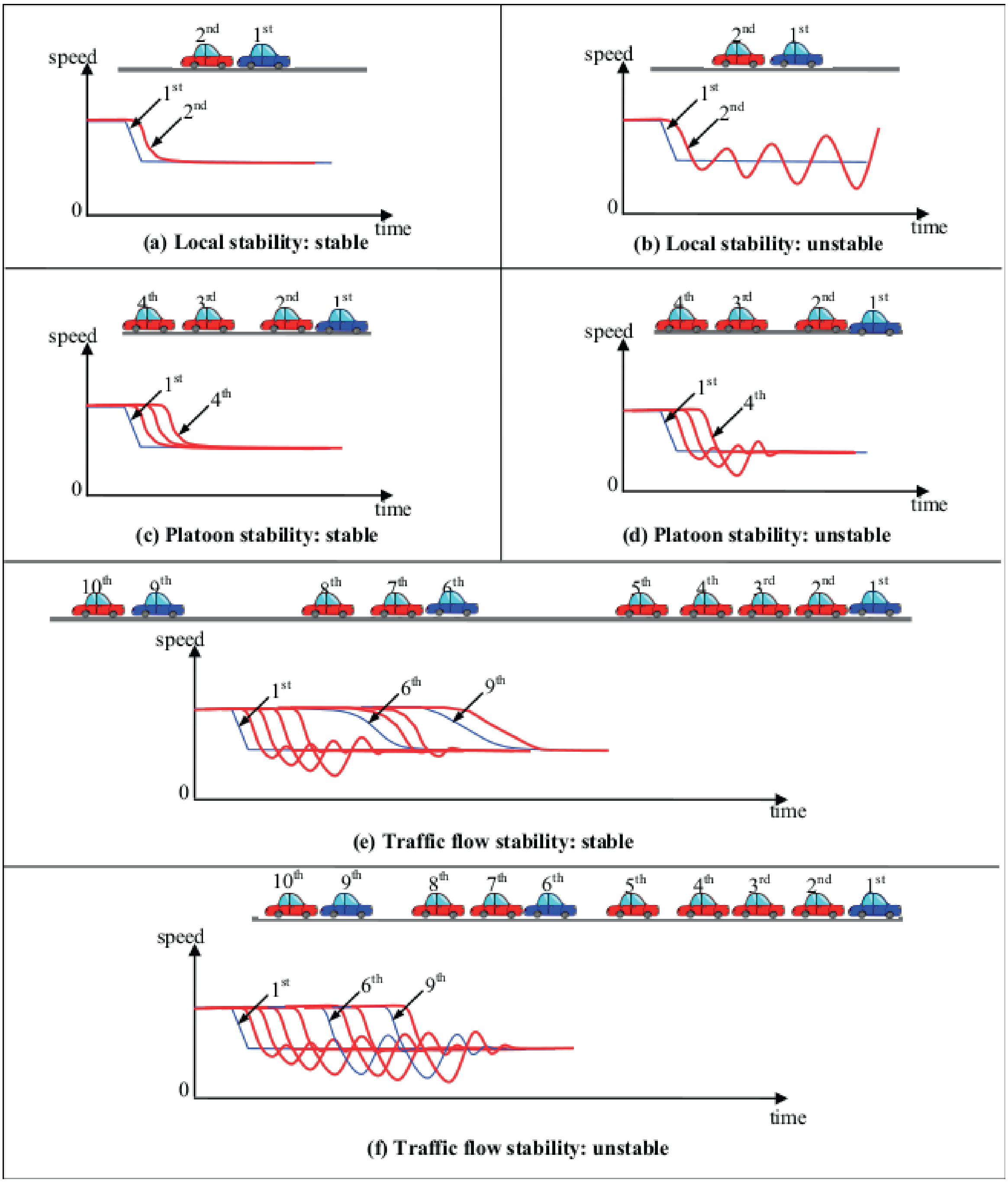


Figure 5-1: Graphical view of the three different types of stability. The graphs show the speeds as function of time for different vehicles. Figure concept from Pueboobpaphan and van Arem (2010)

5-1-1 Local stability

For local stability, one has to consider one vehicle pair, in which the second vehicle is following the leader. If the leader reduces speed, the follower needs to react. In case of local stability, he does so gracefully and relaxes into a new state. Some car following models in combination with particular parameter settings will cause the following vehicle to continuously change speeds. The following vehicle then has oscillatory speed with increasing amplitude. This is not realistic in real life traffic.

5-1-2 Platoon stability

For a platoon, it should be considered how a change of speed propagates through a platoon. This is called platoon stability or string stability. Consider a small change in speed for the platoon leader. The following vehicle needs to react (all vehicles in the platoon are in car-following mode). Platoon stability is whether the speed disruption will grow over the vehicle number in the platoon. In platoon stable traffic, the disruption will reduce, in platoon unstable traffic, this disruption will grow.

Note that platoon stability becomes only relevant to study for locally stable traffic. For locally stable traffic, one can have platoon instability if for a single vehicle the oscillations decrease over time, but they increase over the vehicle number, as shown in figure 5-1.

5-1-3 Traffic flow stability

The third type of stability is traffic flow stability. This indicates whether a traffic stream is stable. For the other types of stabilities, we have seen that stability is judged by the speed profile of another vehicle. For traffic flow stability it matters whether a change in speed of a vehicle will increase over a long time to any vehicle, even exceeding the platoon boundary. Traffic which is platoon stable is always traffic stable. The other way around, platoon unstable traffic is not always traffic unstable. It could happen that the unstable platoons are separated by gaps which are large enough to absorb disruptions. In this case, the traffic flow is stable, but the platoons are not. If the disruption propagates to other platoons, the traffic flow is unstable.

5-1-4 Use of stability analysis

There are mathematical tools to check analytically how disruptions are transmitted if a continuous car-following model is provided. For platoon stability this is not very realistic since all drivers drive differently. Therefore an analytical descriptions should be taken with care if used to describe traffic since one cannot assume all drivers to drive the same. In fact, driver heterogeneity has a large impact on stability. The other way around, the stability analysis could indicate whether a car-following model is realistic.

From data, one knows that traffic is locally stable and mostly platoon unstable. Depending on the traffic flows, it can be traffic unstable or stable. For large number of vehicles, traffic is generally stable (otherwise, there would be an accident for any speed disruption).

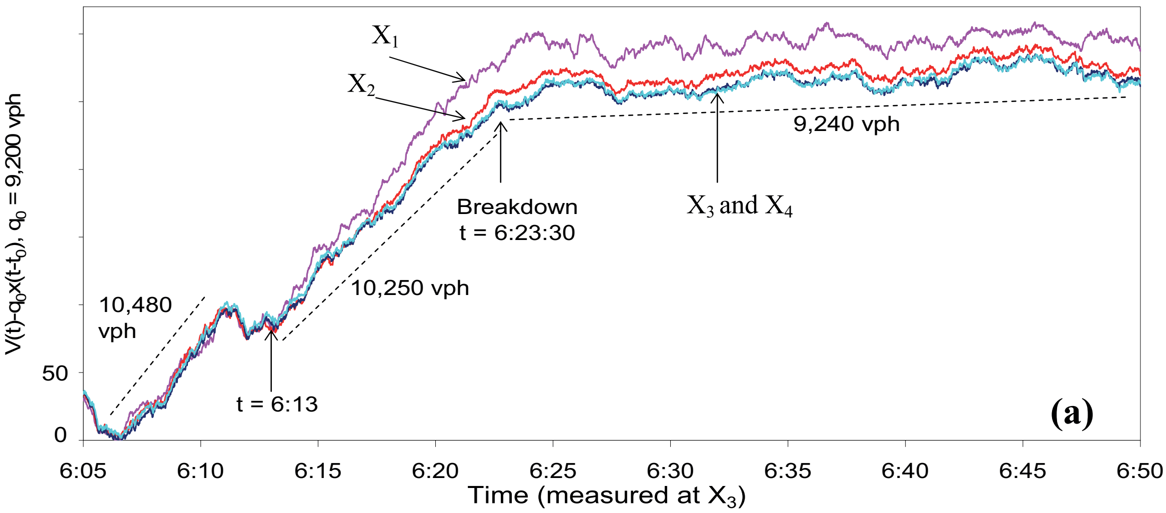


Figure 5-2: Example of an analysis with slanted cumulative curves, from Cassidy and Rudjanakanoknad (2005)

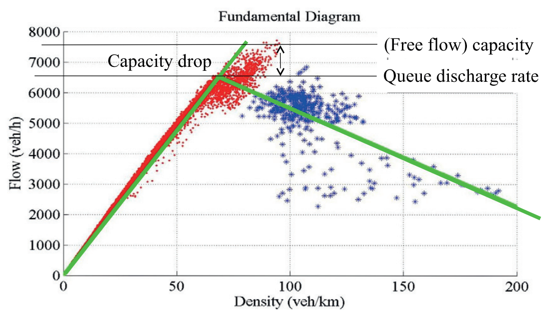


Figure 5-3: Illustration of the capacity drop in the fundamental diagram using real-world data

5-2 Capacity drop

5-2-1 Phenomenon description

The capacity drop is a phenomenon that the maximum flow over a bottleneck is larger before congestion sets in than afterwards. That is, once the bottleneck is active, i.e. there is a congested state upstream of the bottleneck and no influence from bottlenecks further downstream, the traffic flow is lower. This flow is called the queue discharge rate, or sometimes the outflow capacity.

The best way to analyse this is by slanted cumulative curves. Generally, the flow during congestion would be the same. That is, the queue outflow does not depend on the length of the queue. Therefore, a natural way to offset the cumulative curves is by the flow observed during discharging conditions. One would observe an increasing value before congestion sets in.

In the fundamental diagram, one would then observe a congested branch which will not reach the capacity point. Instead, one has a lower congested branch, for instance the inverse lambda

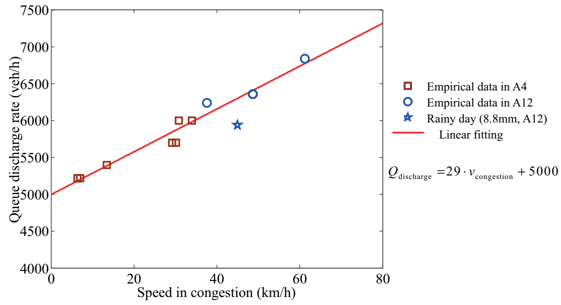


Figure 5-4: The dependency of the queue discharge rate of the speed in the queue (from Yuan et al. (2015a)

fundamental diagram. The capacity drop is the difference between the free flow capacity and the queue discharge rate. This is shown in figure 5-3.

5-2-2 Empirics (from Yuan et al. (2015a))

Capacity with congestion upstream is lower than the possible maximum flow. This capacity drop phenomenon has been empirically observed for decades. Those observations point out that the range of capacity drop, difference between the bottleneck capacity and the queue discharging rate, can vary in a wide range. Hall and Agyemang-Duah (1991) report a drop of around 6% on empirical data analysis. Cassidy and Bertini (1999a) place the drop ranging from 8% to 10%. Srivastava and Geroliminis (2013) observe that the capacity falls by approximately 15% at an on-ramp bottleneck. Chung et al. (2007) present a few empirical observations of capacity drop from 3% to 18% at three active bottlenecks. Excluding the influences of light rain, they show at the same location the capacity drop can range from 8% to 18%. Cassidy and Rudjanakanoknad (2005) observe capacity drop ranging from 8.3% to 14.7%. Oh and Yeo (2012) collect empirical observations of capacity drop in nearly all previous research before 2008. The drop ranges from 3% up to 18%. The large drop of capacity reduces the performance of road network.

In most of observations, capacity drop at one bottleneck only exhibits a small day to day deviation (Chung et al., 2007; Cassidy and Bertini, 1999a). However, it is possible to observe a large difference in the capacity drop empirically at the same location. Srivastava and Geroliminis (2013) observe two different capacity drop, around 15% and 8%, at the same on-ramp bottleneck. Yuan et al. (2) observe different discharging flows at the same freeway section with a lane-drop bottleneck upstream and estimate the outflow of congestion in three-lane section ranging from 5400 vph to 6040 vph. All of these studies show that the capacity drop can be controlled and some strategies have been described to reach the goal (Chung et al., 2007; Carlson et al., 2010; Cassidy and Rudjanakanoknad, 2005). Those control strategies strongly rely on the relation between the congestion and the capacity drop.

Recent research Yuan et al. (2015a) clearly shows this relationship with speed, see figure 5-4.

The main cause of the capacity drop is not identified yet. Some argue it is lane changing, others argue it is the limited acceleration (not instantaneous), whereas others argue it is the *difference* in acceleration (not all the same). This remains an active field of research, both for causes of the capacity drop and for ways to control it.

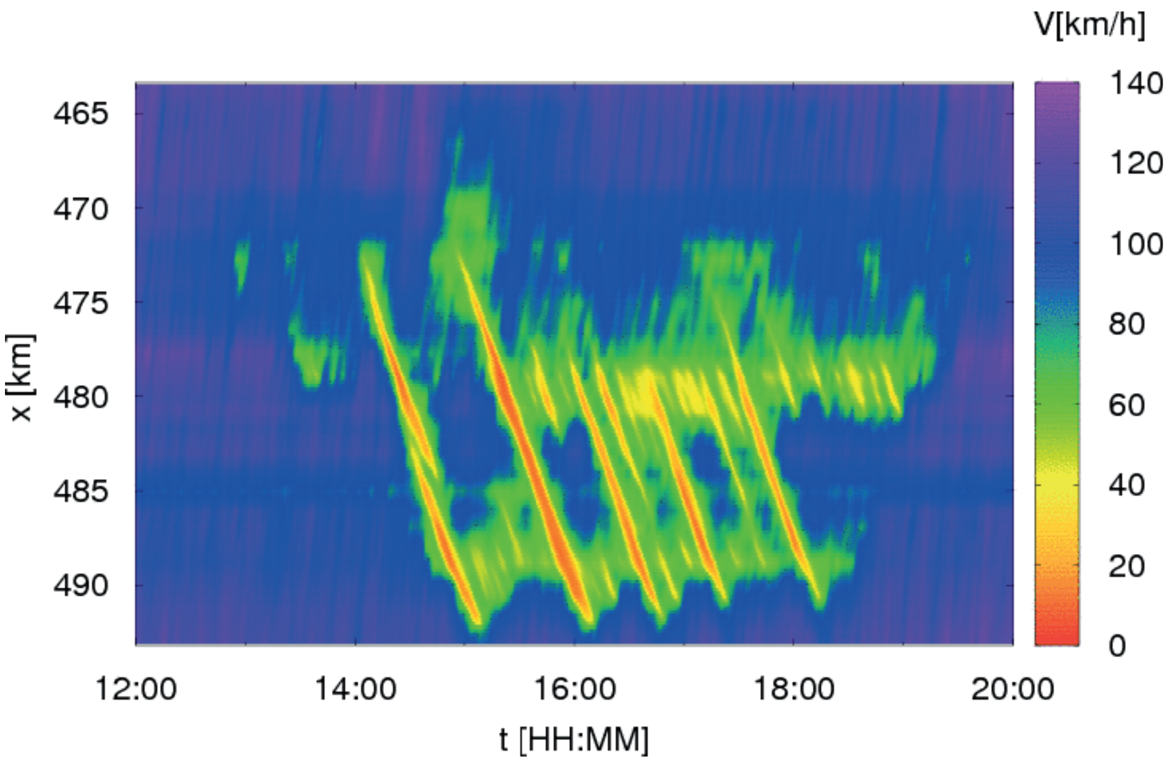


Figure 5-5: Space-time plot of the traffic operations of the German A5, 14 October 2001 – figure from traffic-states.com

5-3 Stop-and-go wave

5-3-1 Phenomenon description

Stop-and-go waves are a specific type of traffic jams. Generally, all traffic states at the congested branch are considered congested states. Sometimes traffic experiences short so-called stop-and-go waves. In these short traffic jams, vehicles come to (almost) a complete standstill. The duration of the queues is a few minutes. At the downstream end of the jam, there is no physical bottleneck. For many drivers, it is surprising they have been in the queue "for no obvious reason".

The outflow of these stop-and-go waves is at the point of the queue discharge rate. Since the traffic in the jam is at almost standstill, the density in the stop-and-go wave is jam density. The speed at which the head moves, can be determined with shock wave theory. Using the fact that the wave speed is the difference in flow divided by the difference in density, we find that the head propagates backward with the wave speed (the slope of the congested branch of the fundamental diagram).

These short jams usually start in congestion. Only if traffic is unstable, short disturbances can grow to jams. In free traffic, the spacing between platoons is usually large enough to ensure these jams do not grow. That means that upstream of the stop-and-go wave, traffic conditions are probably close to capacity. A similar traffic state upstream and downstream of the stop-and-go wave means that the upstream and downstream boundary move at the same

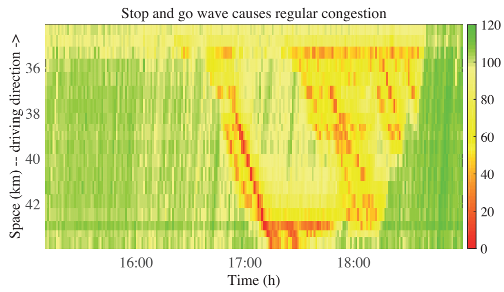


Figure 5-6: Example of a stop-and-go wave triggering a standing queue at the A4 motorway

speed. Therefore, the jam keeps its length.

Often, these type of jams occur regularly, with intervals in the order of 10 minutes. Then, the outflow of the first stop-and-go wave (queue discharge rate) is the inflow of the second one. Then, the upstream and downstream boundary must move at the same speed. Hence, the jams move parallel in a space time figure – see also figure 5-5.

Since traffic in the jam is (almost) at standstill, no flow is possible. That means that the upstream boundary must continue propagating backwards through a bottleneck – the inflow stays at the same level and the vehicles which join the queue have no choice but join the queue keeping their desired jam spacing to their predecessor. Once the jam passes the bottleneck, the desired outflow would be the queue discharge rate of the section where the congestion is. However, the capacity of the bottleneck might be lower.

In such a case, the stop-and-go wave triggers a standing congestion, as is seen in figure 5-6. A stop-and-go wave propagates past a botteneck near location 43km. The outflow of the stop and go wave would be larger, but now the flow is limited. Between the stop-and-go wave and the bottleneck, a congested state occurs, of which the flow equals the flow through the bottleneck (a localised bottleneck, so at both sides the same flow). This congested state is at the fundamental diagram. For most applications where the congested branch of the fundamental diagram is considered a straight line, this congested state will lie on the line connecting the queue discharge point with the jam density. The speed of the downstream boundary of the stop-and-go wave is now determined by the slope of the line connecting the "stopped" jam state with the new congestion state. Since these are all on the same line, the speed is the same, and the stop-and-go wave continues propagating.

5-4 Kerner's Three Phase Traffic Flow Theory

In the first decade of this century, the number of traffic states were debated. Kerner (2004b) claimed that there would be three states (free flow, congested, and synchronised). He claimed that all other theories, using two branches of the fundamental diagram – a free flow branch and a congested branch – assumed two states. This led to a hefty scientific debate, of which the most clear objections are presented by Treiber et al. (2000) and Helbing et al. (1999). A main criticism of them is that the states are not clearly distinguished. Whereas most scientist now agree that three phase traffic flow theory is not a fully correct description of traffic flow, it includes some features which are observed in traffic. In order to discuss these features and introduce the names in that framework, the theory is included in the course. This reader provides a short description for each of the states: an interested reader is referred to Kerner's book for full information (Kerner, 2004b), or the wikipedia entry on three phase traffic flow theory for more concise information.

5-4-1 States

Kerner argues there are three different states: free flow, synchronized flow and wide moving jams.

Free flow

In free flow traffic, vehicles are not (much) influenced by each other and can move freely. In multi-lane traffic, this means that vehicles can freely overtake. Note that as consequence, the traffic in the left lane is faster than the traffic in the right lane. The description of Kerner's free flow state therefore has similarities with the description of the two-pipe regime in Daganzo's theory of slugs and rabbits (see section 7-2).

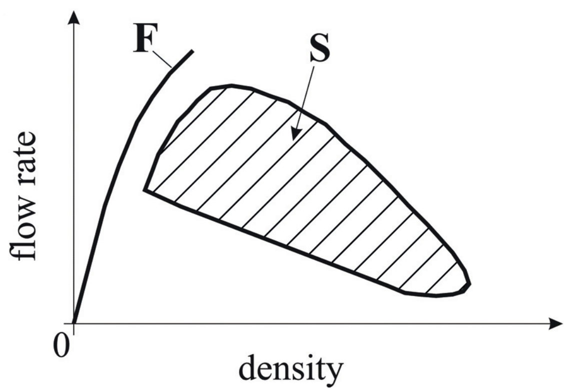
Synchronized flow

Synchronized flow is found in multi-lane traffic and is characterized by the fact that the speeds in both lanes become equal – hence the name of the state: the speeds are "synchronized". This could be seen similar to the one-pipe regime of Daganzo.

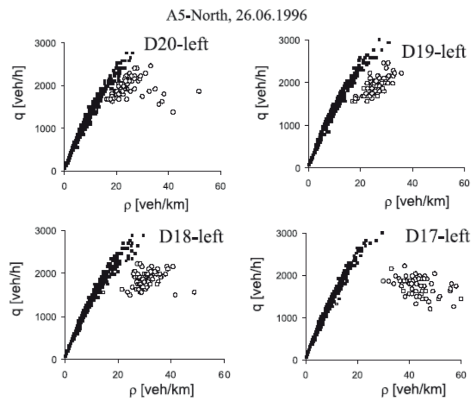
In Kerner's three phase traffic flow theory this is one of the congested states, the other one being wide moving jams. The characterising difference between the two is that in synchronised flow, the vehicles are moving at a (possibly high) speed, whereas in wide moving jams, they are at (almost) standstill. Kerner attributes several characteristics to the synchronized flow:

1. Speeds in all lanes are equal
2. Speeds can be high
3. The flow can be high – possibly higher than the maximum free flow speed. Note that in this case, the "capacity" of the road is found in the synchronized flow.
4. (Important) there is an area, rather than a line, of traffic states in the fundamental diagram.

The area indicated by an "S" in figure 5-7a is the area where traffic states with synchronised flow can be found. Note that in the figure, the free flow branch ("F") does not extend beyond the maximum flow. However, note that for a certain density, various flows are possible. The explanation by Kerner is that drivers have different equilibrium speeds, leading to different flows.



(a) Concepts, figure from Wikipedia



(b) Observations, figure from Kerner (2004b)

Figure 5-7: Synchronized flow

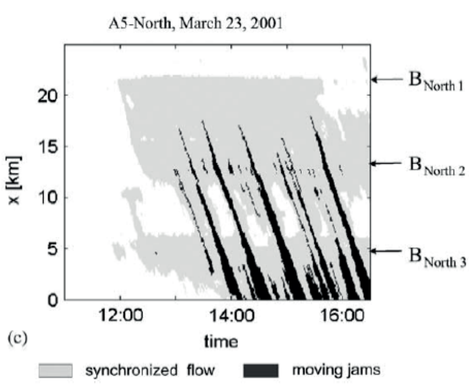


Figure 5-8: Observations of wide moving jams propagating through synchronised flow; figure from Kerner (2004b)

Wide moving jams

A wide moving jam is most similar to what is more often called a "stop and go wave". Stop-and-go relates to the movement an individual vehicle makes: it stops (briefly, in the order of 1-2 minutes), and then sets off again. Kerner relates the name to the pattern (it moves in the opposite direction of the traffic stream), and "wide" is related to the width (or more accurately, it would be called length: the distance from the tail to the head) of the queue compared to the acceleration/deceleration zones surrounding the jam. The speed in the jam is (almost) 0. In the flow-density plane, the traffic state is hence found at flow (almost) 0 and jam density.

By consequence hence the tail moves upstream with every vehicle attaching to the jam. The head also moves upstream, since the vehicles drive off from the front. Hence the pattern of the jam moves upstream. This phenomenon is purely based on the low flow inside the jam, hence it can travel upstream for long distances (dozens of kilometers). It can pass through areas of synchronised flow, and pass on and off ramps, see figure 5-8

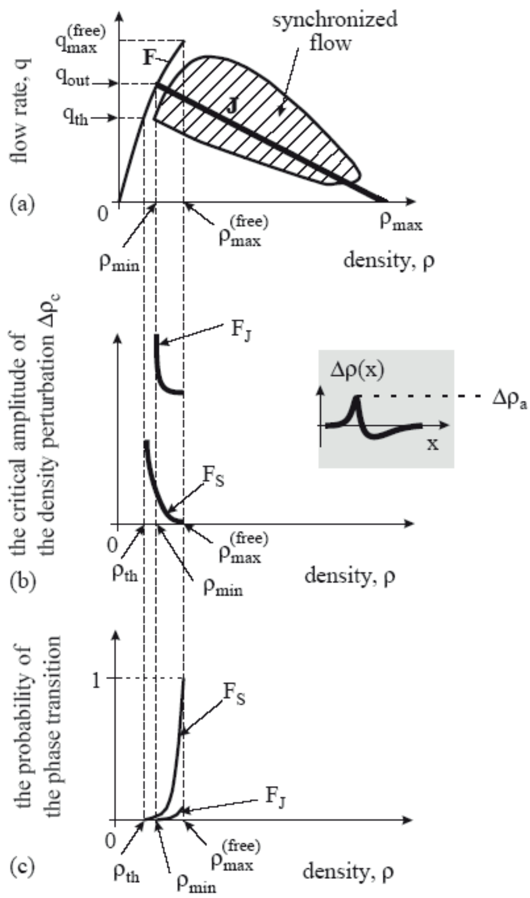


Figure 5-9: Transitions between the phases, figure from Kerner (2004b)

5-4-2 Transitions

The other aspect in which Kerner's theory deviates from the traditional view, is the transitions between the states. Kerner argues there are probabilities to go from the free flow to the synchronized flow and from synchronized flow to wide moving jams. These probabilities depend on the density, and are related to the size of a disturbance.

Let's consider the transition from free flow to synchronized flow as an example. As the density approaches higher densities, the pertubation which is needed to move into another state, is reducing (since the propagation and amplification increases). One can draw the minimum size of a disturbance which changes the traffic state as function of the density, see figure 5-9. Besides, one also knows with which probability these disturbances occur. Hence, one can describe with what probability a phase transition occurs, see figure 5-9.

Selected problems

For this chapter, consider problems: A-1-2, 12-13, 25, 30, 31, A-2-2, 64, 99, 107, 108, 109 110, A-5-1, A-5-2, 153, 161, 183, 180, 191, 192, 189, 218, 243, 244, A-10-2, 278, 279

Chapter 6

Car-following

After this chapter, the student is able to:

- Predict the vehicle's trajectory given a CF model (know Newell's model by heart, others given in eqns)
- Interpret a model/FD at the other aggregation level (microscopic and macroscopic)

A car following model describes the longitudinal action of the vehicle as function of it's leader(s). That can be, it describes its position, speed or acceleration. There are many different forms, and all have their advantages and disadvantages. They all aim to describe driving behavior, and human behavior is inconsistent and hence difficult to capture in models.

This chapter does not give an overview of all models, nor a historical overview. That is given by Brackstone and McDonald (1999). Instead, in this chapter we discuss the simplest, Newell's, and discuss some characteristics which one might include.

6-1 Newell's car following model

The most easy car-following model is the model presented by Newell (2002a). It prescribes the position of the following x_{i+1} car as function of the position of the leader x_i. The model simply states that the position (and – by consequence – also speed, acceleration or jerk) of the follower is a distance s_j upstream of the position (or respectively speed, acceleration or jerk) of the leader of a time τ earlier

$$x_{i+1}(t) = x_i(t - \tau) - s_j \tag{6-1}$$

This means that the follower's trajectory is a copy of the leader's trajectory, translated over a vector $\{\tau, s_j\}$ in the xt-plane (see figure 6-1). This vector has also a direction, which can be computed by dividing its vertical component over the horizontal component.

$$w = \frac{s_j}{\tau} \tag{6-2}$$

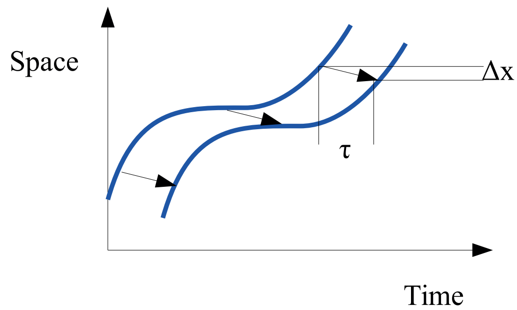


Figure 6-1: Newell's car-following model is translating a leader's trajectory over a vector $\{\tau, s_j\}$

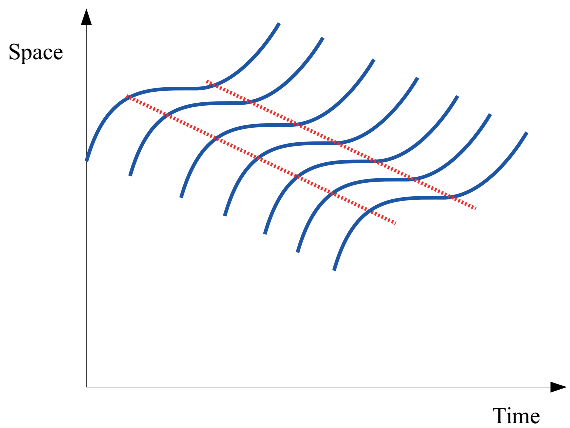


Figure 6-2: The average of the car-following behavior of drivers each with different parameters of the Newell model leads to a shock wave speed

This slope is the speed at which information travels backward in congested conditions, and hence is the slope of the congested branch of the fundamental diagram.

For each driver, different values for τ and s_j can be found. From empirical analysis (Chiabaut et al., 2010) it is shown that averaging the slopes over more than 12 drivers would yield a constant shock wave speed. This is illustrated in figure 6-2. Although the parameters of the car-following model are all different, the wave speed is an average translation of the disturbance, indicated by the red dotted line.

6-2 Characteristics

6-2-1 Dependencies

The car-following model describes the action of the following vehicle. That of course can depend on the movement of the leading vehicle. Some elements which are often included in a car-following model are:

- Acceleration of the leader: if the leader accelerates, the follower can get closer

- Speed: a higher speed would require a longer spacing
- Speed difference: if a follower is approaching his predecessor at a high speed, he needs to brake in time
- Spacing: if the spacing is large, he might (i) accelerate and/or (ii) be less influenced by its leader
- Desired speed: (i) the faster he wants to go, the more his desire to close a gap. But also (ii) even if the predecessor is far away, the follower will not exceed its desired speed

All these elements occur in car-following models. This list combines elements used for different type of models. For instance, some models might prescribe a distance, and hence use speed as input, whereas others might prescribe a speed, and use distance as input.

Surprisingly, many of the available car-following models are "incomplete", i.e. they lack one or more of the above elements and are therefore limited in their use. Using them for a dedicated task is of course allowable. A user should ensure that the model is suited for the task.

6-2-2 Reaction time

Human drivers have a reaction time. Mostly, the models are evaluated at time steps, which are chosen small, often in the order of 0.1 second. A good model allows to set the reaction time of the driver separately. Then, only information which is more than a reaction time earlier can influence a driver's acceleration.

Some models will use the model time step as reaction time. In that case the speed or acceleration of the model is evaluated every time step, which then is typically 1 second. Information of the previous time step is immediately used in the next time step, and in between there are no accelerations considered. The effect of reaction time on the traffic flow is described in Treiber et al. (2006). Generally speaking, a large reaction time can make the traffic flow unstable.

6-2-3 Multi leader car-following models

Human drivers can anticipate on the movement of their leader by looking ahead and considering more leaders. This has empirically been shown by Ossen (2008). Generally, one would adapt the car-following model allowing n times the spacing for the nth leader. Also, one might expect that a follower acts less sensitive on inputs from leaders further away – that might be in terms of lower acceleration (i.e., closer to zero) or later (i.e., a higher reaction time).

Two "mistakes" are common, leading to multi-leader car-following models which seem to make sense, but are not realistic.

1. In the car-following behavior, only vehicles in the same platoon should be considered, or vehicles which might influence the following vehicle. A car-following model which always takes three leaders regardless of the spacings might take a leader which is a long distance away, which in reality will not influence the driver.

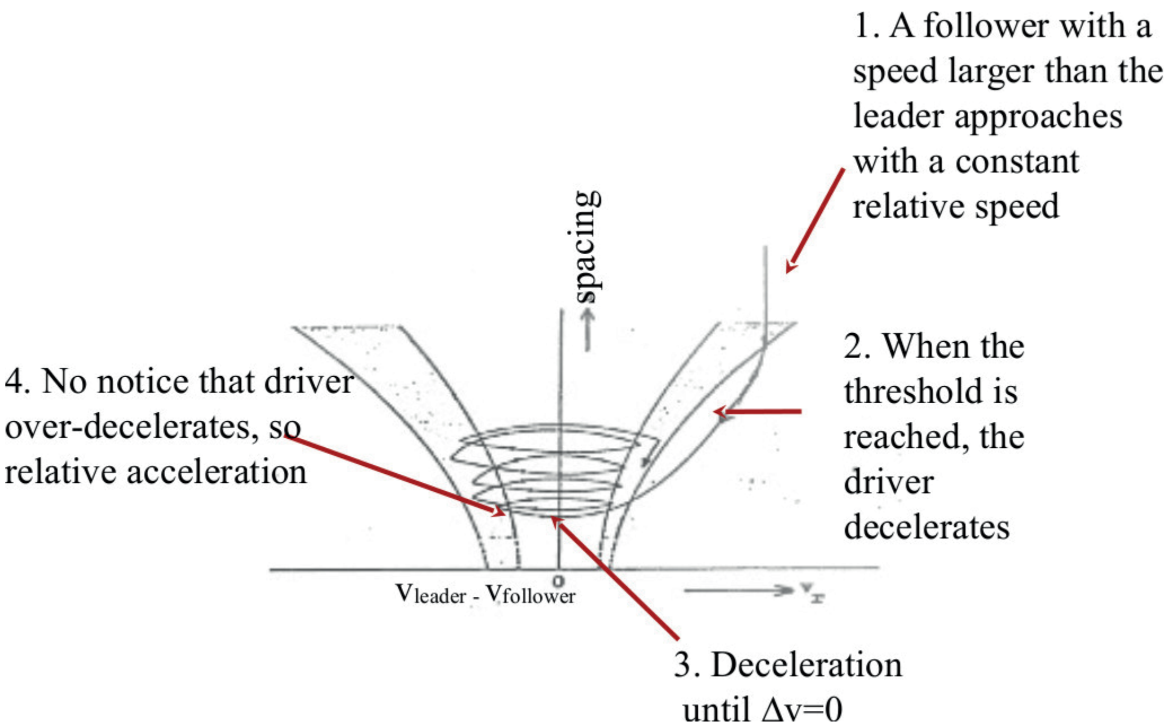


Figure 6-3: Schematic overview of the principles by Wiedemann

2. If the car following model predicts the *average* acceleration caused by each of the leaders, it might be that the follower crashes into its leader. Instead, a minimum operator is usually more suitable.

6-2-4 Insensitivity depending on distance

Most car-following models indicate drivers will adapt their speed based on the action of their leaders. Wiedemann (1974) states that drivers only react if the required action is above a certain threshold. Specifically, he indicates that within some bounds, drivers are unable to observe speed differences. One might also argue that within some boundaries, drivers are unwilling to adapt their speeds because the adaptation is too low. These bounds depend on the spacing between the leader and follower: the larger the spacing, the larger the speed difference needs to be before a driver is able to observe the speed difference. As an example, a 1 km/h speed difference might be unobservable (or unimportant to react upon) on a 200 meter spacing, but if the spacing is 10 meter, this is observable (or important).

It is relevant to draw a diagram relating the relative speed to the spacing, as is done in figure 6-3. The thresholds are indicated: within these bounds, the driver is insensitive and is considered to keep his current speed. Outside the bounds, the driver accelerates to match the traffic situation.

When the leader brakes, the follower closes in. That means that point is in the right half (the follower has a higher speed) and the line is going down (the spacing reduces because the follower is driving faster). At a certain moment, the follower comes closer to his leader and

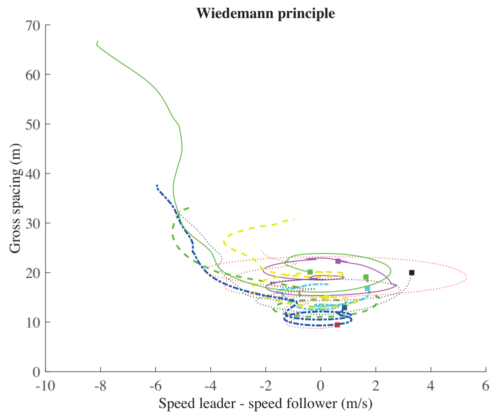


Figure 6-4: Empirical observations of leader-follower pairs in the relative speed - spacing plane

starts reacting to the speed difference, i.e. he will reduce speed. He can either not brake enough and have still a higher speed, or overreact. If the braking is not enough, the situation is the same as the starting situation, and the same situation will occur. Therefore, we will assume for the further reasoning that the follower overreacts.

The overreaction means that the follower will have a lower speed than the leader. That means that the speed difference will go through zero (the y-axis in the figure). At that moment, the speeds are equal, so the spacing remains momentarily the same. Therefore, when crossing the y-axis, the line is horizontal. After that, the leader drives at a higher speed than the follower, and the spacing will start increasing. For low speed differences, the speed difference is under the observation threshold and the follower does not adapt his acceleration. Only once the next bound is reached, he will notice the leader is shying away and start accelerating again.

If he will again overreact, the speed difference will become positive. Again, the speed difference will go through zero (the y-axis) where the line has a horizontal tangent line (a zero speed difference means that at that moment the spacing remains constant). One can continue constructing this figure, and it shows that in this phase plane (i.e., the plane relative speed - spacing) the leader-follower pair makes circles.

The wider the circles are, the wider the observation thresholds. The thresholds are considered to be higher for larger spacings. Empirical studies (e.g., Knoop et al. (2009); Hoogendoorn et al. (2011) indeed reveal these type of circles.

Caution is required in interpreting these figures. A Newell car-following model does not implement these type of observation thresholds. However, the relative speed - spacing figures will show circles due to the reaction time. Namely, once the leader brakes, the follower will not yet, and in the relative speed-spacing plane this will be shown as a part of a circle. A correct analysis tracks the car-following behaviour not at the same moments in time (vertical lines in the space time diagram), but at lines moving back with the shock wave speed. When doing so for the Newell model, one would not find any circles at all. A detailed analysis of this method is presented by Laval (2011)

Finally, some words on the principle in relation to car-following. Car-following models prescribe the position, speed or acceleration of the follower based on the position, speed, or acceleration of their leaders. The principle laid out in this section only mentions when the speed is *not* adapted, but does not specify the acceleration outside the thresholds. As such, it therefore cannot be considered a car-following model. Instead, it can be combined with another car-following model to get a complete description.

6-3 Examples

This section shows the some frequently used models, apart from the Newell car-following model described in 6-1.

6-3-1 Helly

The first model we describe here is the Helly model (Helly, 1959). The Helly model prescribes a desired spacing s^* as function of the speed v:

$$s^* = s_0 + Tv \tag{6-3}$$

Note that this could be considered as a spacing at standstill (jam spacing) plus a dynamic part where T is the net time headway (subtracting the jam spacing from desired spacing).

Now the acceleration is determined by a desire to drive at the same speed as the predecessor and a desire to drive at the desired headway. The model prescribes the following acceleration:

$$a(t) = \alpha \left(\Delta v(t-\tau)\right) + \gamma \left(s(t-\tau) - s^*\right) \tag{6-4}$$

In this equation, Δv is the speed difference, t is a moment in time and τ a reaction time.

6-3-2 Optimal Velocity Model

The optimal velocity model proposed by Bando et al. (1995) is a car-following model specifying the acceleration a as follows:

$$a = a_0(v^* - v) \tag{6-5}$$

In this equation, v is the speed of the vehicle, and a_0 a reference acceleration (tunable parameter, constant for a specific vehicle-driver combination). v^* is determined as follows:

$$v* = 16.8(\tanh(0.086(s-25) + 0.913)) \tag{6-6}$$

In this equation, s is the spacing (in meters) between the vehicle and its leader, giving the speed in m/s.

6-3-3 Intelligent Driver model

Treiber et al. (2000) proposes the Intelligent Driver Model. This prescribes the following acceleration:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = a_0 \left(1 - \left(\frac{v}{v_0}\right)^4 - \left(\frac{s^*(v, \Delta v)}{s}\right)^2\right) \tag{6-7}$$

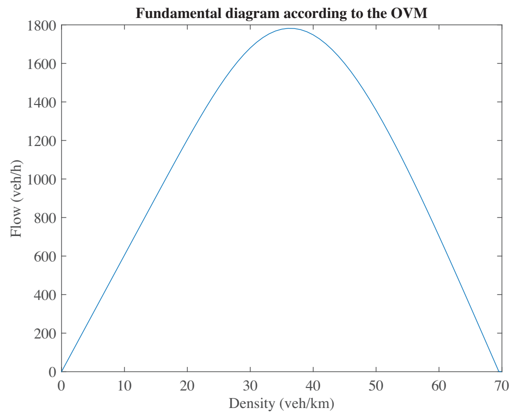


Figure 6-5: Fundamental diagram according to the OVM

with the desired spacing s^* as function of speed v and speed difference Δv:

$$s^*(v, \Delta v) = s_0 + vT + \frac{v\Delta v}{2\sqrt{a\,b}} \tag{6-8}$$

In this equation, Δv is the speed difference between the leader and the follower, a is acceleration and b a comfortable deceleration. a_0 is a reference acceleration (parameter).

6-4 Relation to fundamental diagram

The fundamental diagram gives in its usual form the relation between the density and the flow for homogeneous and stationary conditions. Also car-following models can describe homogeneous and stationary conditions. In that case, all vehicles should drive at the speed and should not change speed (stationary). Therefore, the acceleration of all vehicles must be zero.

Many car-following model prescribe the acceleration based on speed and spacing. The equilibrium conditions mean that the acceleration is zero (and one might argue the relative speeds as well). The relation between spacing and speed then gives an implicit equation for the spacing-speed diagram. Using that the density is the inverse of the (average) spacing, and the flow is (average) speed times density, one can reformulate the spacing-speed diagram from the car-following model into a fundamental diagram.

Take for example the OVM, equation 6-5 and equation 6-6. Equilibrium conditions prescribe that drivers do not accelerate, hence $a_0(v^* - v) = 0$. That means that either $a_0 = 0$, or

$$(v^* - v) = 0 \iff v^* = v \tag{6-9}$$

Since a_0 cannot be zero – in that case the vehicle would never accelerate. Therefore, in equilibrium conditions, equation 6-9 should hold. Using equation 6-6, we find:

$$v = 16.8(\tanh(0.086(s - 25) + 0.913)) \tag{6-10}$$

This gives a relation between speed and spacing. For a fundamental diagram in flow-density, that gives (using $q = kv$:

$$q = kv = k\,(16.8(\tanh(0.086(s-25)+0.913))) \tag{6-11}$$

The spacing can be changed into a density. Since all vehicles have the same spacing (equilibrium conditions), we may use

$$s = \langle s \rangle - 1/k \tag{6-12}$$

Substituting this in equation 6-11, we find an expression for the fundamental diagram.

$$q = k\,(16.8(\tanh(0.086(1/k-25)+0.913))) \tag{6-13}$$

Note that in this expression with the values as in equation 6-6, density should be in vehicles per meter and the resulting flow is given in veh/s. Converting these units, this results in the fundamental diagram as shown in figure 6-5.

Selected problems

For this chapter, consider problems: 11, A-2-4, A-3-2, A-4-2, 133, 154, 179, 181, 220, 222, A-10-4, 266, 277, 280, 281, 282, 283

Chapter 7

Microscopic lane change models

After this chapter, the student is able to:

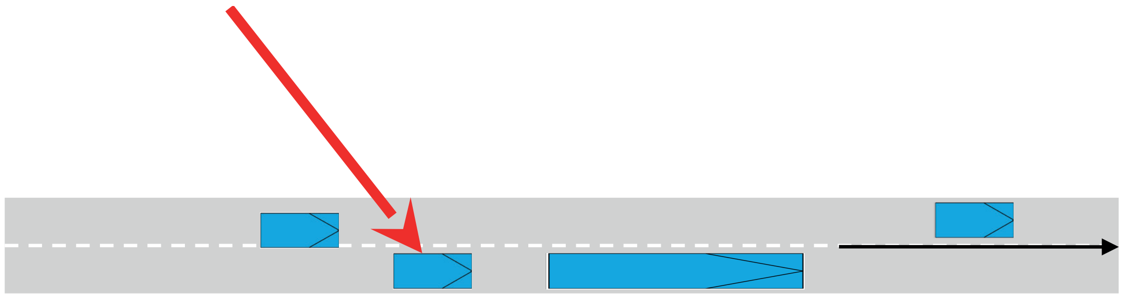
- differentiate between courtesy, mandatory and desired lane changes
- comment on the principles of the Mobil lane change model
- comment on lane selection theories, “slugs and rabbits” in particular
- explain 4 other lane change strategies
- comment and analyse lane distributions
- describe the principle of relaxation

The parts up to now are mainly focussed on the longitudinal driving behaviour and the modelling thereof. This chapter discusses lateral driving behaviour. For multi-lane roads, the lane changing plays an important role. In fact, lane changing is claimed to be the main cause for traffic breakdowns (Ahn and Cassidy, 2007). They argue that a single lane change manoeuvre might lead to a disturbance which grows and due to overreaction, leads to a stop and go wave. Moreover, the lane flow distribution (LFD) is not equal at bottlenecks, implying that in some lanes there is capacity remaining even though the combined flow gets overcritical.

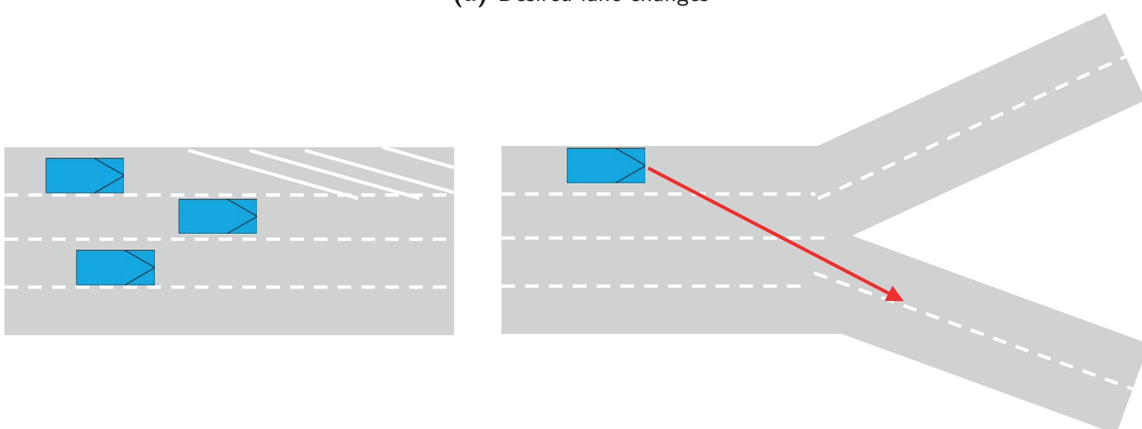
Firstly, this chapter discusses a lane selection model from a psychological perspective (section 7-2). Then, in section 7-3 a microscopic lane change model is introduced. Section 7-4 discusses more advanced lane change models, combining lateral and longitudinal movements.

7-1 Type of lane changes

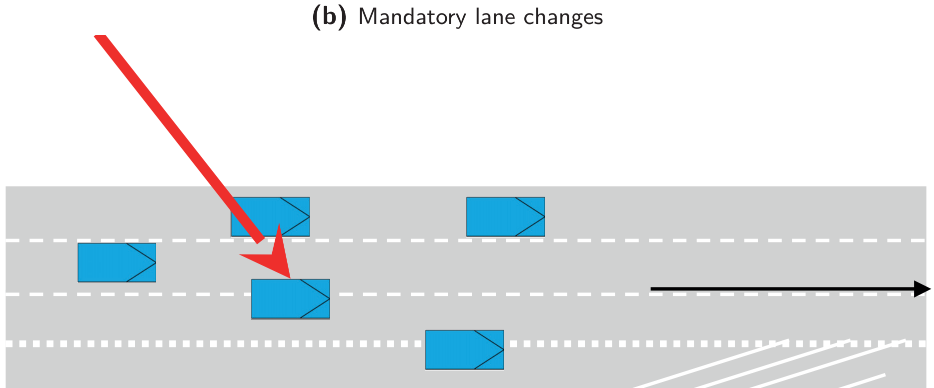
Three types of lane changes are distinguished. These are show in figure 7-1 and explained below



(a) Desired lane changes



(b) Mandatory lane changes



(c) Coutesy lane changes

Figure 7-1: Different type of lane changes

1. Desired lane changes are lane changes which a driver does to be better off, usually because the speed in the adjacent lane is larger.
2. Mandatory lane changes are lane changes which a driver does because he has to to continue to its destination. This might be because the lane ends and the driver has to merge, or because the road splits into roadways with different directions, see figure 7-1b
3. Courtesy lane changes are lane changes which a driver makes to help other drivers. There is no direct benefit for the driver that makes the lane change.

7-2 Slugs and rabbits

This section describes a theory of lane selection, posed by Daganzo (2002a,b).

7-2-1 Theory

This theory poses there are slugs and rabbits, two different types of drivers each with a specific type of lane selection. They are defined as follows:

Category	free speed	lane choice
Slug	Low	always right
Rabbit	High	Fastest

This means that the left lane(s) can operate at a different speed, and even in a different regime than the right lane.

7-2-2 Traffic operations

Consider a situation with low demand. Then, slugs drive at their desired speed in the right lane. This speed is lower than the desired speed of the rabbits. Rabbits hence drive their desired speed and stay in the left lane(s). Both (type of) lanes have a different speed, and operate independently. That is why this state is called a two pipe regime.

If the density in the left lane(s) increases, the speed can decrease to values below the free flow speed of the right lane. Rabbits then choose to change lanes towards the right lane, since they will change to the fastest lane. This will increase the density on the right lane, and lower the speed. The density on the left lane(s) will decrease and the speed will increase. Thus, the difference in speeds between the right and the left lane(s) decreases. This process of changing lanes continues until the speed differences between the left and the right lane have decreased to zero. Then, the complete flow operates at one state, and this flow is called a one pipe regime.

7-2-3 Loading

The section above describes traffic operations at a continuous motorway stretch. Interesting phenomena occur at on-ramps. At an on-ramp, traffic merges onto the main road, but always into the right lane. Consequently, the rabbits are not in their desired lane. They need a distance to perform the additional lane change into the left, and faster, lane. If the density in the left lane already is high, the traffic in the left lane might become overcritical, and speeds might reduce. Note that this is happening *downstream* of the point of merging, namely at a point where the rabbits change towards the fast lane.

7-2-4 Consequences

Although the model of lane selection as proposed by Daganzo is simple, it does explain the following features observed in traffic.

- Boomerang effect. This effect (Helbing, 2003) says that a traffic disturbance starts at an discontinuity (for instance an onramp), than travels a while with the traffic, before traffic breaks down to very low speeds and the disturbance will travel upstream back to the discontinuity where it started.
- Uneven lane distribution. It is found that near capacity conditions, the left lane has a much higher traffic flow than the right lane (see e.g., (Knoop et al., 2010).
- Capacity drop, elaborated in section 5-2.

7-3 Utility model

The theory of slugs and rabbits is a psychological theory which explains some phenomena. It can not yet directly be implemented in a model. The lane change model MOBIL, abbreviation for Minimizing Overall Braking Induced by Lane Changes, can. It describes for each time step whether or not a vehicle will change lanes. The model is explained in this section. The section first describes the idea of the model. Then, the model is formulated in terms of equations.

7-3-1 Model idea

The basic idea of the model is that for each of the lanes a *utility* is calculated. This utility is based on the following items (note: all of the utilities below have a different value for a different lane choice):

- Foreseen acceleration of the driver: the more it can accelerate the better it is
- Foreseen acceleration of the other drivers: one would prefer not to hinder others
- (For European driving:) how far right: the rules prescribe to keep right unless overtaking.

These utilities are calculated for each of the possible decisions: change left, change right or stay in current lane. Then, they are weighted per decision and summed. This gives the utility for a certain lane.

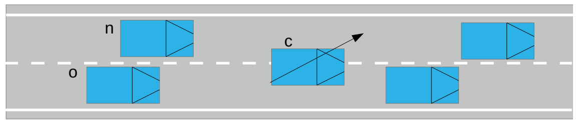


Figure 7-2: The lane changing vehicle and the surrounding vehicles. In this figure, the vehicle is considering a lane change to a left lane. If a lane change to the right lane is considered, vehicles in the other lane should be considered, which will be indicated with the same symbols

7-3-2 Model equations

The model is based on utility, calculated per lane, which we indicate by U_λ. Drivers take their own utility into account, as well as the utility of others. Figure 7-2 shows the other vehicles involved in a lane change. The lane changing vehicle is indicated by a c, the new follower by a n and the old follower by a o.

The total utility for a driver is considered a weighted sum of the utility for himself and the other vehicles:

$$U_{\text{tot}} = U_c + \mathcal{P} \sum_{i \in \text{other drivers}} U_i = U_c + p\left(U_o + U_n\right) \tag{7-1}$$

The utility for the vehicle is expressed by its instantaneous acceleration a, as computed using the IDM car-following model (see section 6-3-3). The utility U for vehicle i is then expressed as

$$U_i = a_i = a_0 \left(1 - \left(\frac{v}{v_0}\right)^{\delta} - \left(\frac{s^*(v\Delta v)}{s}\right)^2\right) \tag{7-2}$$

For the interpretation of the variables, we refer to section 6-3-3.

A lane change is performed if the utility of the driver for the other lane is at least a_{th} higher than the utility of the driver in the current lane. The variable a_{th} acts as a threshold variable in this case: as long as the utility gain is lower than that threshold, a driver will not change lanes. From a behavioral point of view, a_{th} can be seen as the cost of lane changing, which can differ per driver.

European driving rules

In Europe, driving rules dictate to stay right unless overtaking. In the United Kingdom, the rule is to stay left unless overtaking. In the description here, we will refer to left and right for right hand driving as in continental Europe; for changes in the UK system, change left and right in the following description.

Two changes are made to the model:

1. The utility in the right lane is maximised at the utility of the left lane. By that change, one accounts for the fact that overtaking at the right is not allowed.

2. A bias towards the right is introduced, a_{bias}. This variable has a value larger than 0, and shows how much utility people attach to keeping right. A lane change to the right is performed if

$$U^{\text{right}} - U^{\text{left}} \geq a_{\text{th}} - a_{\text{bias}} \tag{7-3}$$

A lane change to the left is performed if

$$U^{\text{left}} - U^{\text{right}} \geq a_{\text{th}} + a_{\text{bias}} \tag{7-4}$$

7-4 Integrated modelling

Usually, models for lane changing separate a desire to change lanes from the manoeuvre itself, which is for instance modelled by gap acceptance Gipps (1986). The model *MOBIL*, as described in 7-3, incorporates the longitudinal accelerations into the choice for the lane. Alternatively, the integrated model by Toledo et al. (2007, 2009) starts with the lane change decision as lead and determines the acceleration.

Intuitively, one might think that drivers prepare for a lane change in their acceleration, leave gaps for others merging into the traffic stream and only gradually adapt their headways after someone merged in front, a phenomenon which is called relaxation. These concepts are incorporated in the *LMRS* model, the Lane Change model with Relaxation and Synchronisation (Schakel et al., 2012).

Recent experimental (Keyvan-Ekbatani et al., 2016) findings show that drivers have four distinct lane change strategies:

1. Speed leading: drivers choose a speed and in order to keep that speed, they change lanes
2. Speed leading with overtaking: drivers choose a speed, and change lanes to overtake. In contrast to the speed leading strategy, drivers will speed up in order to reduce the time of overtaking.
3. Lane leading: drivers choose a certain lane and adapt their speed to the traffic in that lane
4. Traffic leading: drivers claim they “adapt to the traffic”, i.e. they do not have a certain lane or speed in mind.

Surprisingly, although there are 4 distinct strategies, drivers expect that they drive “as anyone else”, and they are not aware that other drivers can follow a different strategy.

Selected problems

For this chapter, consider problems: 2, A-3-5, A-4-3, A-5-2, 137, 138, 182, 190, A-9-4, 249, A-11-7

Chapter 8

Use of traffic models

After this chapter, the student is able to:

- choose the appropriate level of simulation
- comment on the value of calibration and validation
- (in principle) calibrate and validate a model
- comment on the number of parameters in models
- comment on stochasticity in models, and the use of random seeds

Traffic simulation models are a very useful tool to assess the impact of a road design or a new traffic management measure. In assessing measures, we differentiate between ex ante and ex post analyses. This means respectively that the analysis is carried out before the measure is implemented in practice, or after. Especially in ex ante analyses, traffic models are very useful.

The traffic flow theory and proposed simulation tools as discussed in this book can form a useful basis to perform these analyses. However, they need to be used with care. This chapter describes some steps a researcher or engineer needs to take before the models can be used. These steps are calibration (section 8-3) and validation (section 8-3) of the models. Section 8-2-2 discusses the stochasticy of models, and how to handle this. The last section discusses examples where it can go wrong. Before the steps are taken the goal of the model has to be defined, which will be explained in the following section.

8-1 Goal of the model

The goal of the model has to be defined in advance and the aspects the performance of the model will be judged upon also have to be defined in advance. This is called the measure of performance or measure of effectiveness. Examples of this measures of effectiveness are:

- average speed at (x,t)
- travel time
- delay

Even if the measure of performance has been set, it has to be determined how this will be measured. This is called the goodness of fit. This can be for instance the root mean square error of the travel time.

8-2 Type of models

Traffic models can be differentiated at different dimensions. First, we differentiate the level at which a model operates. Furthermore, the model might have one or more driver/vehicle types. Finally, another dimension is the stochasticity included.

8-2-1 Level

There are different levels at which one can simulate traffic. The levels considered in this course are

- Microscopically
- Macroscopically
- (Network level)

In case of microscopic simulation, one describes all vehicles. In case of macroscopic simulation, one describes the traffic states on the road. Hence, this is one level higher than the microscopic simulation. In this book, we only consider traffic simulations with dynamics. Often, in planning, static assignment models are used; these are not considered here.

The choice of the level of simulation depends on the requirements, for instance with regard to the:

- unknown items (e.g., effect of influencing driving behavior, connected vehicles)
- network size
- allowed time for the simulation

The most important reason to choose a microscopic simulation model is if one wants to analyse the effects of individual behaviour on the traffic operations.

8-2-2 Stochasticity

Many traffic simulation models are stochastic. That means that not all simulation *runs* are the same. The runs can be seen as representation of different days for which the same input conditions hold. Typical changes from run to run are:

- Moment of entry into simulation, i.e. headway.
- Car/driver specific:
  - driver type
  - routing information
  - equipped driver (connected)

To get a representative idea of the traffic situation, multiple random seeds (also called replications or iterations) are required. Generally, it is advised to use proper statistics to determine the right number of random seeds. A good way to assess this is to first test what the model needs the represent, the so called measure of effictiveness. In a few runs, determine the standard of this measure of effectiveness for various runs. Also determine the required accuracy of the measure of effectiveness for the study at hand. A very accessible approach is found in Driels and Shin (2004).

Let's consider an example. Suppose one wants to assess the average travel time in a simulation and one would like to know with 95% certainty that the calculated travel time does not differ more than 15% from the mean from the real. Several test runs show as initial values a mean travel time of 28 minutes ($\bar{x}$), and a standard deviation (variation between the different random seeds, not variation between the drivers) of 6 minutes (σ), which is the maximum allowed error ($\epsilon_{\max}$). From statistics, we know that 95% of the observations fall within a margin of 1.96 standard deviations, so we choose a z-value of 1.96.

From statistics it is known that the bandwidth for the mean depends on the number of replications:

$$\epsilon_{\max} = \frac{zS}{\sqrt{n}} \tag{8-1}$$

In this equation, z is used to obtain the likelihood the observation is in between the bandwidth (1.96 in the example case). S is the standard deviation of the process, here approximated by the standard deviation of the test set σ, 6 minutes. n is the number of replications, to be determined. The maximum error is 5% of the mean of the process, here estimated by 5% of the mean of the test set, being 5% of 28 minutes, is 1.4 minutes.

Inversing equation 8-1 gives the equation for the required number of random seeds:

$$n \geq \left(\frac{zS}{\epsilon_{\max}}\right)^2 \tag{8-2}$$

Filling the values for the example, we obtain:

$$n \geq \left(\frac{1.96 \times 6}{1.4}\right)^2 \approx 70 \tag{8-3}$$

This process can be used. Note that the derivation assumes that the estimation for S is accurate, so one needs already a small set of replications to estimate this. Alternatively, one

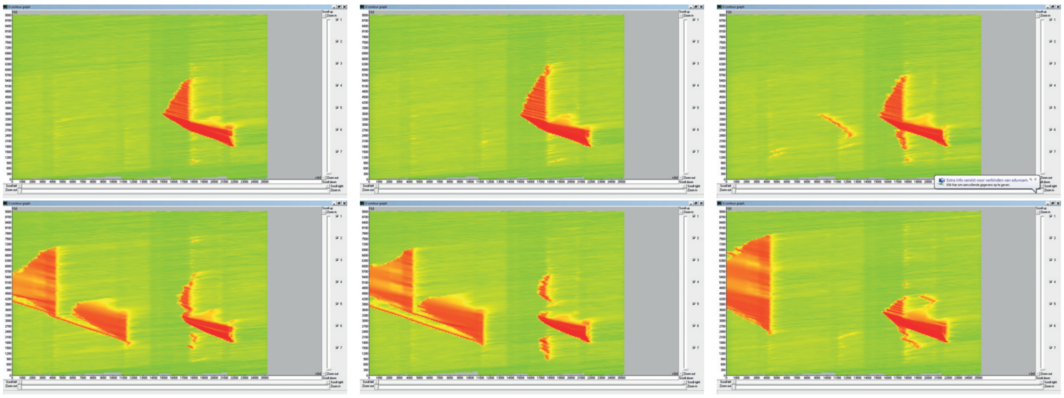


Figure 8-1: Examples of different random seeds in a simulation model. Colors indicate speed in a time-space plane (space is horizontal, time increases from bottom to top

can check at the end whether the used number of replications also holds if S is re-evaluated based on all runs. Moreover, it is a stochastic process and the equations assume a normal distribution. They hence can only be applied for larger sets of replications.

Finally a note from what is been used in practice. Often, due to the long computation times of traffic simulation programs – especially microscopic simulation times – traffic engineers consider a "manual inspection" of the typical patterns, and use 10-15 replications. This will hence not give a statistically significant outcome.

8-2-3 Different user classes

Moreover, a traffic model can include different user classes. For instance, this could be trucks and passenger cars, but one can include more classes: different destinations, different aggressiveness of driving, different type of information.

8-3 Calibration

The goal of the calibration is to make the model representative to the reality. The background will be discussed in the next section. Section 8-3-2 will provide some techniques for calibration.

8-3-1 Use of calibration

Whereas some aspects of traffic are universal (conservation of vehicles), the driving behaviour is not constant. It might differ for instance per:

- Road layout
- Day of the week
- Weather condition

- Lighting condition
- Region/country
- Time of the day (including peak / non peak)

That also means that developers of the model cannot provide a model which can be used in any condition. Therefore, the user has to change the parameters in the model such that it fits the situation he wants to use the model for.

The calibration entails adapting the driving parameters such that the model is in agreement with reality. It is therefore essential that the user has data of traffic available to calculate the goodness of fit of the model.

For microscopic models, if the microscopic behaviour has to be represented correctly, it is needed that microscopic data is available. For an overview of the techniques and uses of microscopic calibration and validation, we refer to the thesis of Ossen (2008). Some microscopic models only aim at representing macroscopic characteristics. For these cases, the model can be calibrated using macroscopic data that uses macroscopic measures of performance. It can not be assumed, though, that the underlying microscopic properties are correct.

The end result of a calibration is a model including a parameter set which is representing the situation for which it has been calibrated. To test and check its predictive power for other situations, a validation (section 8-3) is carried out.

8-3-2 Techniques

The idea is to get the best performance of the model, i.e. optimizing the goodness of fit. In this section, two frequently used values will be discussed, being the root mean square error and the log-likelihood.

Root mean square error

An often used measure to indicate the quality of the simulation is the root mean square error, or RMSE. For instance the RMSE of the speeds at each time and space interval indicates how far the speeds are off "on average". The root mean square error is the difference in speed (or in general: the indicator) squared, and then averaged over all observations, in most cases several observations over space and time. Of this average of squares, one takes the mean to get a value for the difference. This gives an indication how far the speeds are different from the base values.

Note that the RMSE of speeds is not always a good measure. For instance, speeds for the free flow branch, the traffic state can differ considerably, with an (almost) equal speed. Choosing density as predictor for a traffic state is more meaningful as representative of the traffic state. However, the core of the calibration process lies in which is the measure of performance the user wants to reproduce. If the user is satisfied if the speeds are correctly predicted and the traffic states are off, the RMSE of speeds can be a good way. Moreover, if there are stop-and-go waves which originate at random times, the RMSE of speed can be very far off if the waves are predicted in the "other phase" (i.e., start is predicted when traffic is stopped and vice versa). In that case, all speeds are wrong, and a simulation predicting no congestion would be better, although from a traffic flow perspective that is further from the truth. All in all, it lies in the requirements for the user: if the user really wants speeds, minimising the RMSE might be a good option.

Log-likelihood

For stochastic models, the model predicts the outcome of a certain traffic state with a probability. Let us denote the probability on a traffic state $P(S|\Pi)$. This depends on the parameters of the model, Π. The likelihood indicates how likely it is that the data is found given the model. Assuming independency between the observations, one can express the likelihood as:

$$L\left(\Pi\right) = \prod_{\text{All observed data}} P(S_{\text{observed}}|\Pi) \tag{8-4}$$

Theoretically, one likes to maximize the likelihood of the model giving the observations by adjusting the parameters. Hence, the optimal parameter set Π^* is found by:

$$\Pi^* = \arg\max_{parset} L\left(\left(\Pi\right)\right) \tag{8-5}$$

The likelihood is very small. That is because each datapoint contributes a multiplicative factor of 0 to 1 to the likelihood, being the probability that the model with that parameter set predicts that outcome. For many observations, that value will be very small. For computational reasons, it might therefore be advisable to maximize the logarithm of the likelihood instead, also called the loglikelihood.

The loglikelihood $\mathcal{L}$ is defined as

$$\mathcal{L} = \log L \tag{8-6}$$

Because the likelihood is a product of probabilities, the log likelihood changes into a sum of the logs of the probabilities:

$$\mathcal{L} = \log L = \sum_{\text{All observed data}} \log P(S_{\text{observed}}|\Pi) \tag{8-7}$$

8-3-3 Number of parameters

A higher number of parameters gives more degrees of freedom to fit the model to the data. However, this also comes at a risk of overfitting. One might choose the parameters to best fit the data, but the interpretation of the parameters could be wrong because the parameters are actually fitted random effects in the data rather than the underlying structure of the data.

For model developers, the advice is that a parsimonious model should be used – a model as simple as possible which represent the data. For model users, the advice is to only fit the parameters which can be estimated reliably and which have an influence on the data which are collected.

A typical example of where (too) many parameters are to be estimated is in calibrating a microscopic model, including all driver parameters, using macroscopic data. As is indicated in

Ossen (2008), one cannot calibrate a microscopic model using macroscopic data since many combinations of parameters in the microscopic model lead to the same macroscopic traffic patterns. If one would like to calibrate the microscopic model nonetheless, one should refer to microscopic data.

8-4 Validation

The section describes the validation of a model. In the validation, is is tested how well the model is performing for the cases it is intended for.

8-4-1 Need of validation

The result of a calibration is a model and parameters which are optimized. The purpose of validation is checking whether the model indeed does as it is intended to. It could be that all data from the calibration fits perfect, but for the cases in validation there are other patterns. In that case the model is calibrated, but not validated. In general, the calibration consists of finding an optimal parameter set, for which no specific quality level needs to be given. For the validation, the goal of the model must be set, as well as a validity range.

The validity range is the range for conditions under which the model is working at a quality level defined by the user on beforehand. This validity range could be the same conditions as the validation, another day, another location, another country, another number of lanes – this is to be specified by the user. However, note that the model is only validated for those conditions the validation has been carried out for. For example, if the validation is carried out for a similar day on the same road layout, it is not necessarily a correct model for another road layout.

The essence of validation is to make sure the model indeed performs to the required level. Before that is done, no statements of the quality of the model can be made.

8-4-2 Data handling for calibration and validation

For calibration and validation different data sets should be used. This prevents for problems of overfitting: if the parameters are fit to the specific (random) properties of the data, this should not show in the validation.

If the validity range is different from the calibration, one data set could be used for the calibration and another dataset can be used for validation. That could for instance be two different locations. If no specific conditions could be identified, the data should at random be split in a part for calibration and validation. A typical distribution is to allow 2/3 for calibration and 1/3 for validation.

8-4-3 Techniques

As mentioned before, the first thing one needs to do is set the aims for the model: under which conditions the model needs to perform at which level. These are preferably intuitive measures; for instance a log-likelihood value is hard to interpret and hence not very useful.

The process then is straightforward: run the model with the correct input variables, and then consider the outcome in the chosen measure of performance and goodness of fit. To which extent does it meet the required (or desired) criteria.

For instance, we want to validate a macroscopic traffic flow model for another day than the day used for the calibration. Data are available for that day, showing the flows and the speeds. We require that with the right inflow, measured at each km and aggregated over 5 minutes, the RMSE of the speeds (compared to the ground truth data) is below 5 km/h. The inflow at the beginning of the road stretch is put into the model, the model predicts the speeds. These are compared with the data, and the RMSE of the speeds is calculated. If the RMSE is below 5 km/h, the model is validated. For a good overview of calibration and validation, see chapter 16 of Treiber and Kesting (2013).

8-5 Often made mistakes

In traffic engineering this often goes wrong. In this section, some examples are given.

The first, and most serious mistake is that uncalibrated models are being used. That could mean that a model is developed such that some behaviour is in there. For commercial micro simulation packages and drivers on freeways, that means that typical car-following behaviour is implemented, just as a lane change model. However, drivers drive differently all over the world and in different conditions. If one would simply implement a road layout and put it in the specific package assuming that the drivers drive as the developer put into the model, the risk of having a very wrong model is large. This risk is further enhanced by the fact that the vehicle trajectories look realistic. This is because the underlying car movements make sense. However, they are not necessarily true for the case at hand.

Calibrated but unvalidated models are also common in practice. Data are being collected and parameters of the model are optimized. However, without validation, one knows the model is the best prediction possible (although a risk of overfitting exists), but one does not know how good it is. It can be very far off the reality.

In microsimulation a vehicle demand is given. It can happen that the tail of congestion in the network is spilling back to the entry of the network. In that case, the demand cannot be put on the network due to the congestion. Various simulation programs then act differently. Some will not generate new vehicles, whereas others store the vehicles in “vertical queues”. Since it is generally not known what models do, it is best practice to avoid congestion spilling back to the network entry. Note that network entries also can be onramps halfway the network. In this case, one can consider elongating the onramp.

The final mistake commonly made is the validated models outside their validation range. For instance, a model could be validated for Dutch freeways with no slope at a speed limit of 100 km/h and dry weather conditions. That model, with those parameter settings, is then only validated for that range. A road in Italy, or road with slopes, or with a different speed limit or with different weather conditions would then fall outside the validation range.

Selected problems

For this chapter, consider problems: 163, 219, 221

Part II

Traffic Flow Modelling and Control Part I – Control

Chapter 9

Performance measures and control objectives

9-1 Introduction

A good performance of road traffic networks is beyond doubt highly important. But what is good performance? It depends on who you ask. Traffic networks have in general many users and stakeholders, and their performance can be expressed in several different types of performance measures.

Different stakeholders in traffic control problems will have different interests. Road managers (such as a national ministry of transportation, or provincial governments) are often concerned about the network performance, safety, and possibly other aspects, such as emissions and noise. However, individual road users are more concerned about their own travel time and fuel consumption. These objectives may conflict with each other. In addition, other interests may be represented by other stakeholders, such as municipalities that are responsible for local roads, industry in industrial areas, a harbor, amusement parks, or stadiums that attract high traffic volumes, public transport companies, passenger associations, motorist associations, individual drivers, cyclists, and pedestrians.

Due to this wide range of stakeholders with usually conflicting interests, policy should be formulated. Policy formulates how the different objectives are prioritized: which parts of the networks should remain congestion free, what type of traffic should go on which road (e.g., local or and long-distance traffic), or what balance is sought between conflicting objectives, such as maximum speed and emissions. Such policies may also be user-specific, e.g., which users should be on which roads, or whether public transport gets priority; but may also specify what service level should be provided on different roads; or what traffic management measures are used.

However, policy is usually not formulated in terms of mathematical equations, and therefore they need to be translated into a mathematical form before they can be used. The typical applications of the resulting *performance measures* or *objective functions* are the evaluation of

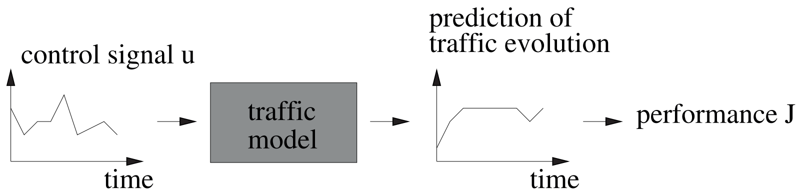


Figure 9-1: In predictive optimization-based control methods the effect of control signal u is predicted using a traffic flow prediction model, and the performance is calculated based on the prediction and the objective function J. The controller aims to find a control signal u that opzimizes the performance expressed by J.

the performance of real or simulated traffic networks, or the design or evaluation of new traffic control approaches. Note that in the context of evaluation, usually the term *performance measure* is used, while for control design the term *control objective*, or *objective function* is used. We will use these terms interchangeably.

Regarding terminology, also note that we use the word *measure* in two different meanings: as a *performance measure* which expresses how well a traffic system performs (such as the average speed in a network), and as *control measure* which influences the dynamics of a traffic system (such as a traffic signal, or speed limit). In addition, we will also have *measurements* that are the physical quantities detected by sensors (such as traffic speed measured by video cameras).

In traffic control system design, objective functions are used in two different ways. One way is to design a traffic control system and use the objective function to evaluate the design by simulation or in reality. The second way is to explicitly include the objective function in the controller. These controllers typically use model-based prediction as shown in Figure 9-1. In such systems the effect of the control signal u is predicted by a traffic flow model and its performance is determined using the objective function J. For example, u can be the green and red times of an urban traffic signal, a model is used to predict how the queues will evolve, and a function J expresses the resulting delay. Since at the end the performance J will directly depend on the choice of u, the optimization problem can be expressed as:

$$\min_{u} J(u) \tag{9-1}$$

in which the controller will find the control signal u^* that minimizes (optimizes) J.

In this chapter we consider the most common performance measures and objective functions. These objectives may be formulated based on macroscopic or microscopic measurements. The performance measures should match the measurements that are available from the detectors, which is in particular important for real-world systems, where the available sensors are a given (differenc detectors can measure different quantities, such as speed or flow at some point, or vehicle trajectories). In this chapter we discuss the different formulations, and the relation between the them.

9-2 Learning objectives

After reading this chapter the reader should be able to:

- give the mathematical definitions of different objectives; and/or explain what the different objectives mean;
- discuss the pros and cons of the different measures for traffic performance;
- apply the relations between the inflow, outflow, density, and total time spent;
- derive the delay for an urban approach;
- define effective green and effective red, explain how they simplify delay calculations;
- analyze the performance improvement of a simple traffic network by some control measure.

9-3 Objective functions

In this section we discuss the most common objective functions, their definition, units, their applicability and compare them to alternatives. For the performance measures it is assumed that they refer to a certain time period, and one link with several segments, as shown by the grey area in Figure 9-2. The extension to networks, or other space-time areas is usually straightforward. For most of the definitions we will use the macroscopic traffic flow variables, such as the flow $q_m(k)$ (veh/h), the speed $v_m(k)$ (km/h), or the density $\rho_m(k)$ (veh/km), where k is a time index referring to time kT, where T (h) is the time step size, and the road segments are indexed by m, with $m \in M$, there M is the set of all segment indices of the considered network, and L_m is the length of segment m. The time period of interest is $k = 1, \ldots, K$. The time-location area of interest is called area A (grey area).

9-3-1 Average speed

The definition of the average speed was discussed in Chapter 1 and defined in (1-9), as the ratio of the total travelled distance (TTD) over the total travel time (TTT). This definition assumes that both, the total traveled distance and the total travel time are measurable. This is the case for example, for floating car data, or for simulations where every vehicle trajectory is measurable. However in practice, often inductive detector loops measure the speed and/or flow, and provide only aggregated values. To express the average speed in macroscopic measurements, we start with Edie's definition. The TTD is expressed using the number of vehicles $\rho_m(k)L_m$ and the travelled distance (per vehicle) $Tv_m(k)$ in segment m, at timestep k (with duration T). The TTT is expressed using the number of vehicles in the

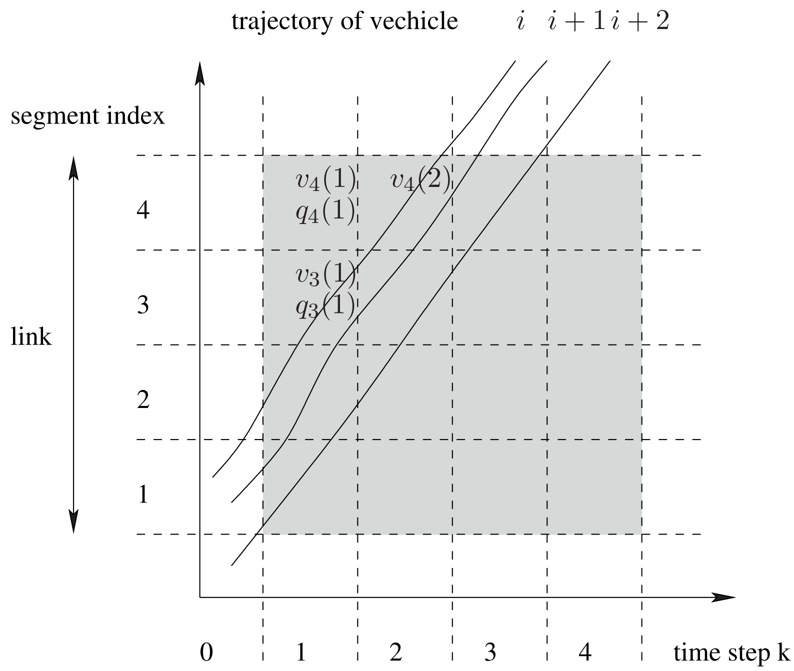


Figure 9-2: Depending on the available detectors, different traffic flow variables can be measured, such as microscopic vehicle trajectories, or macroscopic (aggregated) speed, flow or density. The performance measure is evaluated for a pre-defined time period and segments, indicated by the grey area (to which we refere as area A).

segmnet, times the time duration T that they spend in the segment. This gives:

$$v = \frac{TTD}{TTT} = \frac{\sum_{k=1}^{K}\sum_{m\in M} TL_m\rho_m(k)v_m(k)}{\sum_{k=1}^{K}\sum_{m\in M} TL_m\rho_m(k)} \tag{9-2}$$

$$= \frac{\sum_{k=1}^{K}\sum_{m\in M} N_m(k)v_m(k)}{\sum_{k=1}^{K}\sum_{m\in M} N_m(k)}, \tag{9-3}$$

where $N_m(k) = L_m\rho_m(k)$ is the number of vehicles in segment m at time step k. Note that (9-3) is a weighted average of the speeds $v_m(k)$, weigthed by the number of vehicles in each time-location cell (k, m).

9-3-2 Average flow and total traveled distance

The definition of the average flow was discussed in Chapter 1 in (1-8). Again, Edie's definition is the starting point:

$$q = \frac{TTD}{A} = \frac{\sum_{k=1}^{K}\sum_{m\in M} TL_m \rho_m(k) v_m(k)}{\sum_{k=1}^{K}\sum_{m\in M} TL_m} \tag{9-4}$$

$$= \frac{\sum_{k=1}^{K}\sum_{m\in M} L_m q_m(k)}{\sum_{k=1}^{K}\sum_{m\in M} L_m}. \tag{9-5}$$

where $q_m(k) = v_m(k)\rho_m(k)$. Note that the resulting flow is a weighted average of the measured flows $q_m(k)$.

Since the denominator is constant in (9-5), the maximization of the flow (for control) over fixed time period is equivalent to maximizing the TTD.

While maximizing the average flow (or TTD) may sound attractive, it is not always a good idea. The average flow by definition doesn't express *when* the traffic exits the network, in fact, since the TTD is measured inside the network (and not at the exits only), then it doesn't even tell whether the vehicles have reached their destinations. The TTD only tells how much distance the vehicles have traveled in total. Suppose we compare two cases, with a fixed morning peak demand (and zero demand after it). Case 1: there is no congestion and all traffic exits the network (reaches its destination) within one hour. Case 2: there is a congestion of an hour, and all traffic reaches its destination in the second hour. The TTD will be the same for both cases (averaged over two hours), because all traffic travels from origin to destination, but it is evident that Case 1 is preferred.

Another case, where maximizing TTD is even wrong to do, is when routing is a control measure. Maximizing TTD may mean that traffic is sent to longer routes because then the same vehicles will travel longer distances. This is obviously not what one would want.

For these reasons, care should be take when TTD is chosen as a control objective.

9-3-3 Average travel time

Average travel time is the average time that vehicles reach their destination. Since the travel time is undefined for vehicles that did not complete the trip at the end of the simulations (or measurement period), or for vehicles that were already in the network at the start of the simulation (begin of measurement period), this measure does not include the full behavior of the traffic system. Starting and ending a simulation with an empty network may partially solve the problem, but it may introduce another problem, because the dynamics of the (nearly) empty network is not necessarily the dynamics of interest. For these reasons often the Total Time Spent is preferred, which will be discussed in Section 9-3-6.

Note that the term “travel time” is sometimes confusingly used as the time that vehicles are in motion (opposed to waiting time, when they are stopped), but this is not how we use it here.

9-3-4 Average delay

The general definition of delay is the difference between the nominal travel time (when traffic can drive at the regular speed), and the realized travel time.

The definition for delay[1] J^{delay} (veh-h) that the Dutch Ministry of transport uses is,

$$J^{\text{delay}} = \sum_{k=1}^{K} \sum_{m \in M} \left[L_m \left(\frac{1}{v_m(k)} - \frac{1}{v_m^{\text{free}}} \right) \right] q_m(k) T \,, \tag{9-6}$$

where v_m^{free} is the free-flow speed of link m. The equation expresses the difference of travel time over segment m for the measured speed and the free-flow speed, multiplied by the number of vehicles that have flown over that section during the time step.

Urban queuing and delays: some basic notions

Also in urban context delay is important, and in this section we use a simple queuing model to derive delay.

Let’s consider the individual trajectories of vehicles arriving at a traffic light, as shown in Figure 9-3. The vehicles arriving during the green phase after the queue has dissolved, will pass the stop line without delay. At the end of the green phase the yellow phase will be used by the last vehicles to cross the intersection. The first vehicle that stops is shown as vehicle 1 in Figure 9-3. For that vehicle both trajectories are indicated, the real one, with stopping and re-acceleration after the light has become green, and the trajectory that it would have had, if it had not stopped. Looking at the two trajectories, the *delay* of the vehicle is defined as the time that the vehicle passes a point downstream (at a distance far enough that the vehicle has reached cruising speed again) and the time at which it would have passed if it had not stopped. The *waiting time* or *time in queue* is the time that the vehicle stands at rest in the queue. The delay is obviously larger than the waiting time.

In order to simplify the modeling of the flow that crosses a stop line, usually some assumptions are made. Let’s denote the time when green starts as $t^{\text{g,start}}$, and the time that yellow starts as $t^{\text{y,start}}$. Then it is assumed:

- The flow is zero until time $t^{\text{g,start}} + \lambda_1$ after the start of green. λ_1 (s) is called the *start lag* or *green start lag*. The start lag is partly due to the reaction time of the first driver, but most of the time is the consequence of the fact that vehicles have to accelerate which makes the speed of the first few vehicles lower than in the middle of the green phase.
- The flow is zero after time $t^{\text{y,start}} + \lambda_2$ after the start of yellow. λ_2 (s) is called the *end lag* or *green end lag*, and expresses the part of the yellow phase where vehicles still enter the intersection.

[1] called “voertuigverliesuren” in Dutch

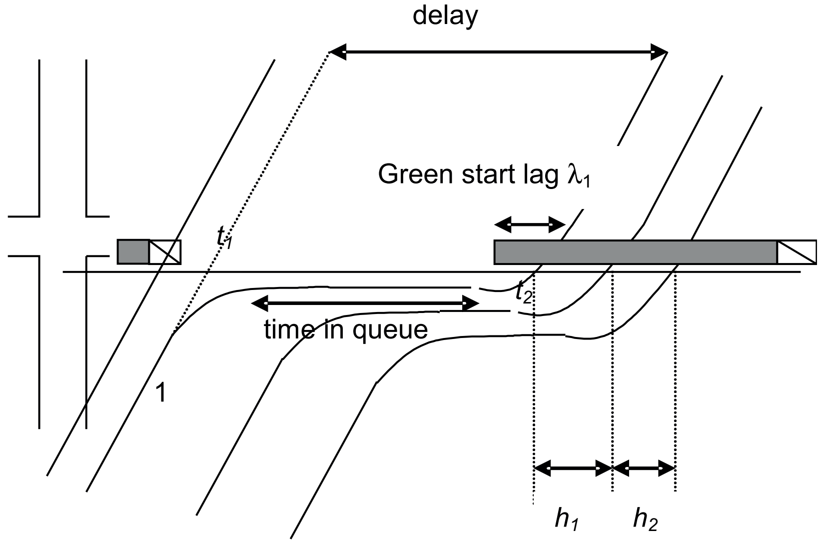


Figure 9-3: Vehicle trajectories at a controlled intersection

- The time between $t^{\mathrm{g,start}}+\lambda_1$ and $t^{\mathrm{y,start}}+\lambda_2$, the flow is assumed to equal the saturation flow s (veh/s), if there is enough traffic. This time period is called the *effective green* time $t^{\mathrm{g,eff}}$ (s), and its duration can be expressed as

$$t^{\mathrm{g,eff}} = t^{\mathrm{g}} - \lambda_1 + \lambda_2\,. \tag{9-7}$$

  The rest of the cycle is called *effective red* time $t^{\mathrm{r,eff}}$ (s) (so there is no effective yellow). Important: the saturation flow and the start and end lags are chosen such that the number of vehicles N that cross the stop line equals the real number of vehicles, and such that

$$N = t^{\mathrm{g,eff}}s\,. \tag{9-8}$$

With these assumptions the vehicles are either crossing the stop line with flow s or are waiting in the queue, and therefore the delay becomes equal to the time in queue. This simplifies the calculations.

Urban queuing and delays: delays at a traffic signal (fixed-time)

This section introduces a simple analytical model for queuing and explains the way to calculate the delay from the queue length as a function of the time.

If vehicles arrive according to a *uniform flow*, that is in a regular pattern with constant gaps between arrivals, the queue will be built up in the way as shown in Figure 9-4. Note that physically the queue length is discrete (the stair shape in the figure), but we will approximate it with straight lines. Let's consider here *vertical* queueing, i.e., queues with the head at the stop line, and without considering their physical length. We will also use the effective green and effective red times.

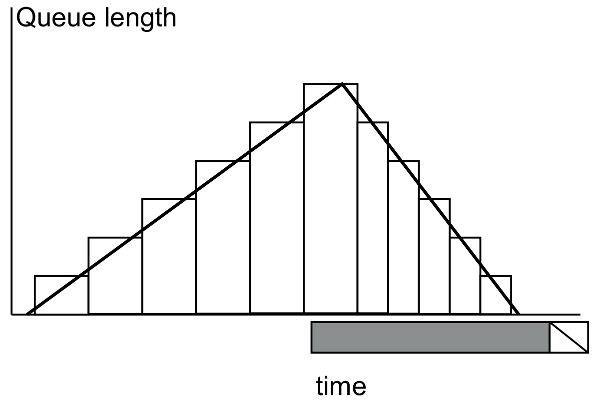


Figure 9-4: Queue lengths as a function of the time

Assume a single road where the inflow is q (veh/s), and that the red phase starts at $t = 0$ with no waiting vehicles (an empty queue), the maximum length $Q^{\max}$ of the queue (at the end of red) is given by:

$$Q^{\max} = t^{\mathrm{r,eff}} q \,. \tag{9-9}$$

During effective green the queue decreases with the rate $s - q$ since the outflow is s and the inflow is still q. So, the queue has cleared at time

$$t = t^{\mathrm{r,eff}} + t^{\mathrm{clear}} \tag{9-10}$$

where t^{clear} is the time needed to clear the queue after the light has become green:

$$t^{\mathrm{clear}} = t^{\mathrm{r,eff}} \frac{q}{(s-q)} \,. \tag{9-11}$$

Let's denote the queue length in front of the stop line if a vehicle arriving at time t^{arrival} by $Q(t^{\mathrm{arrival}})$ (veh). The vehicle that arrives at t^{arrival} will have a delay t^{wait} equal to the remaining red time plus the time needed to clear the queue in front of it:

$$\begin{aligned} t^{\mathrm{wait}}(t^{\mathrm{arrival}}) &= (t^{\mathrm{r,eff}} - t^{\mathrm{arrival}}) + Q(t^{\mathrm{arrival}})/s \\ &= t^{\mathrm{r,eff}} - t^{\mathrm{arrival}} + q t^{\mathrm{arrival}}/s \\ &= t^{\mathrm{r,eff}} - (1-y) t^{\mathrm{arrival}} \end{aligned} \tag{9-12}$$

when

$$0 \le t^{\mathrm{arrival}} < t^{\mathrm{r,eff}} \text{ (when the light is red)} \tag{9-13}$$

where the *load ratio* $y = q/s$ defines the minimum green fraction that is necessary to accommodate the traffic demand.

When the light is green there is no remaining red time anymore, and the queue length equals the maximum queue length minus the reduction of the queue since the start of the green:

$$\begin{aligned} t^{\mathrm{wait}}(t^{\mathrm{arrival}}) &= Q(t^{\mathrm{arrival}})/s \\ &= (t^{\mathrm{r,eff}} q - (t^{\mathrm{arrival}} - t^{\mathrm{r,eff}})(s-q))/s \\ &= y t^{\mathrm{r,eff}} - t^{\mathrm{arrival}} + t^{\mathrm{r,eff}} + y t^{\mathrm{arrival}} - y t^{\mathrm{r,eff}} \\ &= t^{\mathrm{r,eff}} - (1-y) t^{\mathrm{arrival}} \end{aligned} \tag{9-14}$$

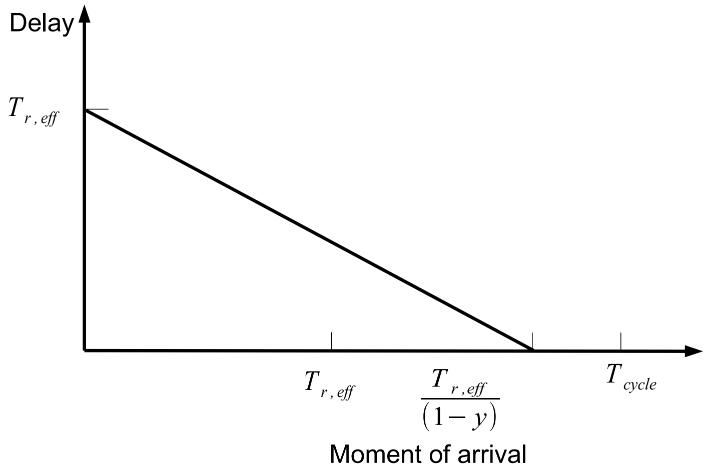


Figure 9-5: Delay as a function of the arrival time in the cycle. The horizontal axis is the arrival time and the vertical axis the delay.

when

$$t^{\mathrm{r,eff}} \leq t^{\mathrm{arrival}} < \min(t^{\mathrm{cycle}}, t^{\mathrm{r,eff}} + t^{\mathrm{clear}}) \tag{9-15}$$

Note that (9-12) and (9-14) are the same.

We assume that the queue will be fully cleared during the cycle, and thus in the green phase there might be also some time where the queue length is zero:

$$t^{\mathrm{wait}}(t^{\mathrm{arrival}}) = 0 \text{ if } t^{\mathrm{clear}} \leq t^{\mathrm{arrival}} < t^{\mathrm{cycle}}. \tag{9-16}$$

Therefore the average delay per vehicle is given for this case by the average height of the delay line in Figure 9-5 over the cycle time:

$$t^{\mathrm{d,mean}} = \frac{1}{2}(t^{\mathrm{r,eff}} + 0)\frac{t^{\mathrm{r,eff}}}{1-y}\frac{1}{t^{\mathrm{cycle}}} = \frac{(t^{\mathrm{r,eff}})^2}{2t^{\mathrm{cycle}}(1-y)} \text{ [s]} \tag{9-17}$$

The total waiting time can also be found as the area under the queue length in Figure 9-4 (the triangle):

$$\begin{aligned}
t^{\mathrm{d,uniform}} &= \frac{1}{2}Q^{\mathrm{max}}[t^{\mathrm{r,eff}} + t^{\mathrm{r,eff}}(q/(s-q))] \\
&= \frac{1}{2}Q^{\mathrm{max}}t^{\mathrm{r,eff}}[1 + q/(s-q)] \\
&= \left(\frac{1}{2}qt^{\mathrm{r,eff}}\right) t^{\mathrm{r,eff}}[s/(s-q)] \\
&= \left(\frac{1}{2}qt^{\mathrm{r,eff}}\right) t^{\mathrm{r,eff}}/[(s-q)/s] \\
&= \frac{1}{2}q(t^{\mathrm{r,eff}})^2/(1-q/s) \\
&= \frac{q(t^{\mathrm{r,eff}})^2}{2(1-y)} \text{ [veh-s]}
\end{aligned} \tag{9-18}$$

Expressions (9-17) and (9-18) differ only by a factor qt^{cycle}, which is the number of vehicles arriving in a cycle, since (9-18) expresses the total delay and (9-18) the delay per vehicle.

9-3-5 Number of stops at traffic signals

The number of stops is often evaluated in an urban traffic control context, because stops mean deceleration and acceleration movements, which are strongly related to noise, emissions, and fuel consumption.

The fraction of vehicles that has to stop can be calculated based on the fraction of the cycle time where a flow still has green, while the queue has already gone. Eq. (9-11) gives the time t^{clear} needed to clear the queue. The green time that is left, the *unsaturated green time* is given by:

$$t^{\text{g,unsaturated}} = t^{\text{g,eff}} - t^{\text{clear}} = t^{\text{g,eff}} - t^{\text{r,eff}} \frac{q}{s-q} \tag{9-19}$$

On the average the fraction P^{stop} of the all vehicles that has to stop, is given by

$$P^{\text{stop}} = 1 - \frac{t^{\text{g,unsaturated}}}{t^{\text{cycle}}} \tag{9-20}$$

The number of stops N^{stops} in one cycle is then

$$N^{\text{stops}} = P^{\text{stop}} q t^{\text{cycle}} . \tag{9-21}$$

9-3-6 Total time spent (TTS)

The time that vehicles spend in a traffic network is one of the most important and most frequently used control objective in traffic management problems. We will see that it expresses the best what we mean by the informal "optimize travel times", even better than the average travel time or the average network flow. This measure is often referred to as the *Total Time Spent (TTS).*

Intuitively, the TTS can be understood as follows. Assume that the inflow of a network is a given by the demand, which means that the moment when vehicles start spending time in the network is fixed and independent of any control action. A good traffic control strategy will help to let as many as possibly vehicles leave the network as soon as possible. That means that they will spend less time in the network, which is exactly what TTS expresses.

One of the forms of the TTS, based on microscopic vehicle trajectories, is already defined in Chapter 1 in (1-8) and (1-6). Here we give another definition, based on macroscopic measurements.

Suppose we have a traffic networks, with links indexed by m, the length of link m is denoted by L_m (km) and the traffic density on the link by $\rho_m(k)$ (veh/km), where k is the time index indicating a time period $[kT, (k+1)T)$, with T the time step length. Then the TTS over a time period $k = 1, \ldots, K$ can be expressed as:

$$J^{\text{TTS}} = \sum_m \sum_{k=1}^{K} T L_m \rho_m(k) . \tag{9-22}$$

Derivation of the TTS

Leaving the network earlier or at a higher rate corresponds to lower TTS. In this section we derive the relationship between the outflow of the traffic network (which depends on the control measures) and TTS. We also show that a relatively small improvement in the outflow can significantly improve the TTS. The derivations follow the lines of Papageorgiou and Kotsialos (2000).

Suppose we have a traffic network with several entrances and exits where the total demand at those entrances sum up to $q^{\mathrm{d}}(k)$ (veh/h) and the total outflow at the exits is $q^{\mathrm{out}}(k)$ (veh/h). The number of vehicles $N(k)$ (veh) can be calculated by taking the number of vehicles in the previous time step and the number of vehicles that entered or exited the network:

$$N(k) = N(k-1) + T(q^{\mathrm{d}}(k-1) - q^{\mathrm{out}}(k-1)) \,, \tag{9-23}$$

where the initial number of vehicles is $N(0)$. This can be rewritten for $k = 1, 2, \ldots$ as

$$N(k) = N(0) + T \sum_{j=0}^{k-1} (q^{\mathrm{d}}(j) - q^{\mathrm{out}}(j)) \,, \tag{9-24}$$

The time spent by $N(k)$ vehicles in one time step is $TN(k)$ (veh-hours), and the total time that the vehicles spend in the network over a period $k = 1, \ldots, K$ (K time steps) is the summation

$$J^{\mathrm{TTS}} = T \sum_{k=1}^{K} N(k) \,. \tag{9-25}$$

Substituting (9-24) in (9-25) gives

$$J^{\mathrm{TTS}} = T \sum_{k=1}^{K} \left(N(0) + T \sum_{j=0}^{k-1} (q^{\mathrm{d}}(j) - q^{\mathrm{out}}(j)) \right) \,, \tag{9-26}$$

which can be rewritten as

$$J^{\mathrm{TTS}} = TKN(0) + T^2 \sum_{k=0}^{K-1} (K-k)(q^{\mathrm{d}}(k) - q^{\mathrm{out}}(k)) \,. \tag{9-27}$$

This expression can be used when a traffic controller is simulated, and its performance in terms of TTS improvement is evaluated. Usually the inflow of the network is defined by the traffic scenario (the demand q^{d}) for which the controller is evaluated, and the TTS for the uncontrolled and controlled case can be determined by measuring the outflow q^{out} and evaluating (9-27). However, occasionally the inflow of the network is different between the uncontrolled and controlled cases, which may bias the results if the *measured* inflow is used. For example, when in the uncontrolled case a jam prevents the entrance of some vehicles, the TTS will be lower (which seems favorable, but isn't). Therefore, the correct way is to include all vehicles in the calculations, also the ones that wanted to enter the network but couldn't. This can be done by taking a fixed (possibly virtual) demand $q^{\mathrm{d}}(k)$ that is the same for all cases, and use that in all calculations, irrespective of the real inflow of the network. The outflow should still be measured, of course.

Minimizing the J^{TTS}

Based on (9-27) if we want to minimize the TTS, and observing that $TKN(0)$ and T^2 are constants, and $q^{\mathrm{d}}(k)$ is given and not influenced by the control actions, then we have to minimize:

$$\sum_{k=0}^{K-1} (K-k)(-q^{\mathrm{out}}(k)) \,. \tag{9-28}$$

or, in other words, maximize

$$\sum_{k=0}^{K-1} (K-k)q^{\mathrm{out}}(k) \,. \tag{9-29}$$

The interpretation of (9-29) is that the TTS is smaller as the outflow is higher, and the TTS is smaller if vehicles flow out earlier (which is expressed by the weight $(K-k)$: smaller k leads to a higher weight).

Relation between the outflow improvement and TTS improvement

Here we show that a small improvement of q^{out} can lead to a significant improvement of J^{TTS}. For simplicity we make the following assumptions:

- the initial state of the network is empty, $N(0) = 0$,
- the traffic demand and the outflows are constant, $q^{\mathrm{d}}(k) = q^{\mathrm{d}}$, and $q^{\mathrm{out}}(k) = q^{\mathrm{unctr}}$ for the case without control, and $q^{\mathrm{out}}(k) = q^{\mathrm{ctr}}$ for the controlled case.

The relative improvement of J^{TTS} by control can be expressed by

$$\Delta J^{\mathrm{TTS}} = \frac{(J^{\mathrm{TTS}})^{\mathrm{unctr}} - (J^{\mathrm{TTS}})^{\mathrm{ctr}}}{(J^{\mathrm{TTS}})^{\mathrm{unctr}}} \,. \tag{9-30}$$

Substituting (9-27) and the assumptions above in (9-30) gives (after removing the common factors)

$$\Delta J^{\mathrm{TTS}} = \frac{q^{\mathrm{ctr}} - q^{\mathrm{unctr}}}{q^{\mathrm{d}} - q^{\mathrm{unctr}}} \,. \tag{9-31}$$

Example: If the demand exceeds the capacity by 20% and the control improves the outflow with 5% then

$$q^{\mathrm{d}} = 1.2q^{\mathrm{ctr}} \tag{9-32}$$

$$q^{\mathrm{ctr}} = 1.05q^{\mathrm{unctr}} \tag{9-33}$$

and thus

$$\Delta J^{\mathrm{TTS}} \approx 0.2 \,. \tag{9-34}$$

So in this example an improvement of the outflow by 5% results in an improvement of the TTS by approximately 20%. It is important to realize, because some control measures realize only a relatively low capacity improvement, but their effect on the TTS may still be significant.

9-3-7 Reliability

A reliable (predictable) traffic system has a value on its own. Even if there is congestion, if it is the same for every day, then it may be still less bad than when it is strongly fluctuating from day to day. For a significant part of the traffic, the arrival time is important. For example, people traveling for business, often have appointments where they have to be on time. If the travel time is accurately predictable, then the departure time can be chosen such that the person arrives exactly on time, but when the travel time fluctuates from day to day, then an extra margin has to be included in the departure time. This means that in the most cases the person will be too early, which may be a loss of productivity.

In some other cases both, being early and being late is penalized. For example, trucks delivering goods often get a fixed time slot when they should deliver their goods, and if the truck is late or early it will get a penalty, because it is taking a time form someone else. For these reasons reliability is also important.

In general, the variability of the different other performance measures, such as the variance of the average speed of individual vehicles (or the travel time of individual vehicles for a given OD pair), may serve as a measure for reliability of the speeds or travel time.

9-3-8 Equity/fairness

Equity or fairness expresses the idea that road capacity, travel time, etc., should be divided in a fair way among the users. In particular when using traffic control measures, it may be possible to increase the overall network performance at the cost of a part of the users. An example is local ramp metering, where the traffic coming from the on-ramp is held back in order to improve the capacity of the freeway. While in total the system improves (e.g., in terms of TTS), the travel time for some vehicles entering the freeway via the on-ramp, will increase. This may be considered unfair or not equitable. However if all traffic enters the freeway via an on-ramp somewhere, and similar delays apply, it may be equitable. Similarly, traffic signals may favor some directions (green waves, red waves) and disadvantaging others.

In general equity measures express the differences of a certain performance measures (such as delay/km or average speed) for different routes or OD pairs.

9-3-9 Environmental measures

The Environmental measures refer to emissions of NO_x, PM_{10}, or CO_2, to the fuel consumption or energy usage, and the noise production.

We will not go into the details of emission models, but the most important factors that determine the emissions, are:

- speed,
- acceleration and deceleration,
- vehicle type,

- total traveled distance.

To what extent the emissions harm residential areas or nature, also depend on the distance of the road to the relevant area and the wind direction. Furthermore, also other sources may significantly contribute to certain emissions, such as agricultural areas and the sea (PM_{10}), industrial areas (CO_2).

9-3-10 Safety

Traffic safety is often hard to measure, because accidents are rare. In addition when evaluating safety in simulation, the models may not even be able to reproduce accidents. In order to get a better view on the safety of existing traffic systems, or the safety of planned traffic management measures, unsafe, near-accident situations are also considered. These situations can be often characterized by surrogate safety measures. Some typical surrogate safety measures are, or are based on quantities such as speed, density, speed variance, speed differences, time to collision, required braking power, and red light violations (Gettman and Head, 2003). Safety measures are extensively discussed in the Course CIE5810 Traffic Safety.

Selected problems

For this chapter, consider problems:A-13-5

Chapter 10

Introduction to Systems and Control

This chapter presents a theoretical and mathematical framework for ITS from the *perspective of systems and control theory*, with the aim to provide the reader with the tools allowing a deeper understanding of the behavior of a system and its control in the context of ITS.

10-1 Learning objectives

After reading this chapter the student should be able to:

- Describe a traffic control system in terms of subsystems, inputs, outputs, state-space description,
- Explain the difference between continuous time systems and discrete time systems,
- Give the definition of equilibrium points and system stability, and calculate them based on the state-space description.
- Explain observability and controllability of systems.
- Explain the basic terms of the field of systems and control, such as: control signals, feedforward control, feedback control, state-feedback, predictive feedback, model predictive control, objective functions and prediction horizon. Describe how these terms and control approaches can be applied in basic traffic control systems.

10-2 Route guidance example case

Throughout this chapter, we will illustrate the theory by means of an example pertaining to *route guidance* in a simple network. In this example, a VMS provides the passing driver with the best route alternative. Note that 'best' will not necessarily mean optimal from a user's

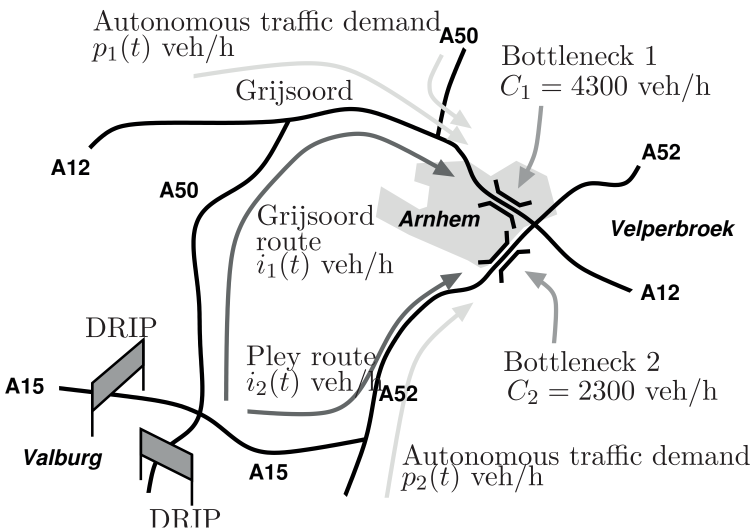


Figure 10-1: The example network around the city of Arnhem (the Netherlands)

perspective, but generally optimal from a system's perspective (cf. system optimum versus user optimum).

Figure 10-1 depicts the example network taken from Toorenburg and Linden (1996), which is used to illustrate the different control approaches. The diamond shaped network offers routing alternatives to the drivers. Part of the traffic at the decision point at Valburg can take one of two possible routes to their destination Velperbroek (the so-called *non-captives*). This decision is dependent on the guidance provided.

The non-captive traffic demand at Valburg destined for Velperbroek at time t is denoted by $i(t)$. This traffic may take the motorway-alternative along Grijsoord (route 1) or the alternative along the Pley-route, which is partially non-motorway. On both routes, a bottleneck exists close to the intersection point near Velperbroek. Bottleneck 2 has a low capacity ($C_2 = 2300$ veh/h) compared to bottleneck 1 ($C_1 = 4300$ veh/h). Consequently, the Pley-route is vulnerable for re-routed traffic flows. Additionally, the unconstrained travel time of route 1 is smaller ($T_1 = 16$ min) than the free travel time of route 2 ($T_2 = 18$ min). Therefore, when no queues are present on route 1, drivers generally prefer this route. Non-captive drivers from Valburg destined for Velperbroek and captive drivers from other routes can cause the bottlenecks to become oversaturated. Whenever this happens, queues build up just before the bottlenecks. The variables $i_1(t)$ and $i_2(t)$ denote traffic choosing the respective routes at time t. Additionally, $p_1(t)$ and $p_2(t)$ denote autonomous traffic demand on route 1 and 2 respectively (as showen in Figure 10-1).

Figure 10-2 shows the traffic demands that are used in the ensuing of the chapter. Notice that these demands are given as a veh/min, and that they resemble a typical peak-hour profile.

10-3 Systems and control theory

The *theory of systems* provides tools for systematic analysis and comparison of systems in general. This is usually done by means of mathematical relations. *Control theory* deals

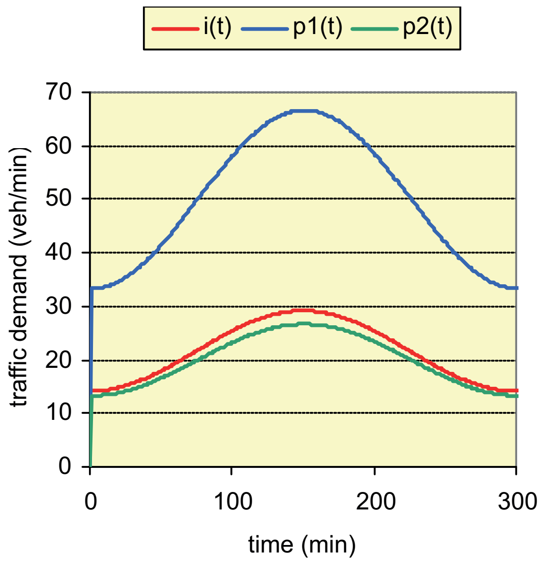


Figure 10-2: Traffic demands for Arnhem network example.

with the basic principles underlying analysis and design of control systems. In this context, to *control* a system means to *influence its behavior to achieve a desired goal*. This goal is generally expressed by some mathematical function (the objective function J) that describes the desired system behavior. For instance, the system can be controlled such that the *collective travel time of all its users is minimized*. Alternatively, the total system throughput may be maximized, the total emissions may be minimized, etc. Alternatively, the objective of the controller might be to steer the system towards a certain desired state. Even in this case, an objective function can generally be given, as the objective is then to minimize the difference *e(t)* between the desired state and the current state.

10-4 Theory of systems

The concept of *systems* is used in many fields of science and technology. Generally, the considered systems are *physical* by nature, although examples of economical, societal, or political systems are known as well. *Systems' theory* enables to study different systems systematically, and to analyze their properties. Although not all systems can be described adequately by a set of mathematical equations (e.g. human behavior), thinking in terms of (sub-) systems provides an *incentive to consider all system's aspects*, and not be satisfied by solutions of sub problems. In this respect, many human and technical activities have side effects, which can be easily overlooked when only considering the main effect. A systems' approach aims to consider *all relevant and only the relevant* aspects.

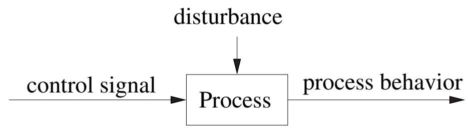


Figure 10-3: A process with the relevant inputs and output.

10-4-1 Inputs, disturbances and outputs

Typically, systems are described as a process with inputs and outputs. The inputs to a process should include all relevant external influences that have an effect on the process.

The inputs that the user (or control system) can influence are called *control inputs.* The inputs that cannot be influenced are called *disturbances.* Disturbances often originate from the environment of the system. However even if the disturbance cannot be influenced, in some cases it can be measured, which can provide useful information about how to change the control signal to still get the desired process behavior. In Figure 10-3 the generic block diagram of a process is shown.

<u>Examples</u>

- Traffic light timing (control input) that infuences the traffic behaviour (process). Turning movement (or route choice) of the drivers (disturbances) cannot be influenced by the traffic controller. (However in future, if for example traffic signals and navigation devices would be integrated, the traffic signals could also influence the route advice to drivers. In such a case the turning movements and routes would become a control input instead of a disturbance.)
- On-ramp flow rate for ramp metering (control input) that infuences the traffic behaviour (process),
- VMS panel (control input) that influences the passing traffic (process),
- Bad weather (disturbance) that influences driving behavior (process),

The inputs and outputs are mathematically described in terms of the *signals* $u(t)$ (control input), $d(t)$ (disturbances), and $y(t)$ (output) which represent the evolution of the values over time t.

The outputs $y(t)$ of a system consist of the measurable components of the system. What the outputs exactly represent depends on the physical (economical, biological) properties of the system and the available sensors. Due to the properties of the system some quantities may not be measurable, or sensors may not provide directly all the desired quantities.

The outputs of a system are typically used for monitoring of the system and the evaluation of its performance. In the context of control the outputs are typically used for determining the necessary control actions.

<u>Example</u>

- _Inductive loop detectors built in the road surface detect vehicle passages (and thus can provide the average flow over some time period) and if they are installed in pairs then they may also provide the average speed over some time interval. The measured speed and flow values provide useful information for monitoring the traffic network and for the evaluation of its performance.
- The traffic density is typically not measured directly as it requires the knowledge of the position of all vehicles on some road stretch, which cannot be measured by loop detectors.

10-4-2 State and state dynamics

The *state* $x(t){\in}X$ of a system at time instant t summarizes the history of the system until instant t. The state of a system is a non-unique quantity that is defined by the following requirement:

> If $x(t_0) \in X$ is the state of a system S at instant t_0, and the input signal $u(t) \in U$, and the disturbance $d(t) \in D$ is given/known for $t \geq t_0$, then both the state $x(t) \in X$, as well as the output $y(t) \in Y$ are uniquely determined for $t \geq t_0$.

This requirement implies that, when the state of the system is known at instant t_0, the systems dynamics for $t{\geq}t_0$ can be determined, even if the input signal $u(t)$ and disturbance *d(t)* for $t{<}t_0$ (control history) is not known. Note that the choice of the state is not necessarily unique, so different state choices may result in a fully determined future state trajectory. For example, one may choose density and speed, or choose speed and flow as the two state variables, since $q = \rho \cdot v$, and two of the three is enough to unambiguously determine how the traffic system in the future will evolve.

The formalization of the state $x(t)$ is (to a large extent) dependent on the problem at hand and the models that used to describe the system's behavior. For example, in Hoogendoorn and Bovy (1996), the state of the system consist of a vector of traffic densities and average velocities of different user-classes, e.g. passenger-cars, trucks, business related traffic, on a number of freeway segments.

As a counterexample, we mention that the travel time on a freeway stretch cannot be considered as the (full) state of the system in general, since the same travel time may correspond to several combinations of speeds and densities over the stretch which may result in different traffic dynamics (and thus different travel times) in the future. Consequently the future evolution of the state is not uniquely defined. When the system is formulated properly, usually the speeds and the densities are the states, and the travel time is the output of such a system.

For most physical or technical systems, the system's dynamics (in relation to the system's input) are specified by the *right-hand-side f* of the so-called *state equation*. These state dynamics are generally described by either a *system of ordinary differential equations* (continuous-time systems):

$$\dot{x}(t) = f(t, x(t), u(t), d(t)) \text{ for } t \in \mathbb{R} \tag{10-1}$$

or by a *system of difference equations* (discrete-time systems):

$$x(t+1) = f(t, x(t), u(t), d(t)) \text{ for } t \in \mathbb{N} \tag{10-2}$$

A more elaborate discussion can be found in Sontag (1990). In the remainder, we will primarily focus on discrete-time problems.

Besides the system equations, often a *measurement* or *observation equation* is given. This equation relates the output y(t) to the state x(t). In mathematical terms:

$$y(t) = g(t, x(t)) \tag{10-3}$$

where g maps the state to the output.

10-4-3 System properties

In this section we discuss notions that are frequently used to characterize systems. The full mathematical specification of these notions is beyond the scope of this chapter, however it is important to understand the relevance and meaning of these notions and to know that mathematical techniques exist for the formalization of these notions.

Linear and time-invariant systems

An important class of systems is the class of *linear systems*. For the discrete time case, these linear systems can be described as follows:

$$x(t+1) = A(t)x(t) + B(t)u(t) \text{ for } t \in \mathbb{N} \tag{10-4}$$

$$y(t) = C(t)x(t) + D(t)u(t)\,. \tag{10-5}$$

Where $A(t)$, $B(t)$, $C(t)$, and $D(t)$ are matrices of appropriate sizes.

Time-invariant systems are systems which dynamics do not depend explicitly on the time t. For a time-discrete system, this means:

$$x(t+1) = f(x(t), u(t)), \quad \text{for} \quad t \in \mathbb{N} \tag{10-6}$$

Many of the theoretical results from mathematical system's theory and optimal control have been determined for linear and/or time-invariant systems. These results include the formulation of *reachability, controllability, observability, and stability* of linear time-invariant systems, which will be shortly discussed below. We refer to Sontag (1990) for more information.

Reachability, controllability and observability

When the system can be controlled from the current state x(t) to another state z(s), we call this state reachable from x(t) (within the period [t,s]). If any state z(s) can be reached from any other state x(t) we call the system controllable. Insight into the controllability of a system is important to understand whether the objectives of applying control are feasible or not.

In DTM, the systems which we are considering are often not fully controllable in a mathematical system theory sense. This means that the system cannot be controlled towards any arbitrary state. In the ramp metering example, for instance, in many cases congestion cannot be avoided completely, but is shifted often from the main road to the on-ramp. In this case, we need to be satisfied with controlling the system to a state which is less severe in the context of the control objective.

Intuitively it is easy to understand that reachability depends on both the available control inputs, and the system dynamics. For example the traffic state in a city can be controlled better (more states can be reached) if there are more traffic control devices (i.e., more control inputs). However, even if all intersections are controlled by traffic lights, some states may not be reachable. For example, it is usually not possible to remove all congestion in a short time.

There are many useful theorems in mathematical systems theory that can be used to determine the controllability of a system and the *set of states to which the system can be controlled*. Without going into detail, we can for instance show that the set $\mathrm{R}^k(\mathrm{x(t)})$ of states z(t+k) that can be reached from x(t) for a linear time-invariant system (Eq. (10-4)) is given by the space that is spanned by the columns of the following matrix:

$$R^k(x(t)) = \ell(B, AB, A^2B, ..., A^{k-1}B) \tag{10-7}$$

Note that it is not required to learn by heart the equation above, but it is important to know that these kinds of procedures exist.

Besides controllability, the *observability* of a system is an important characteristic. With observability we mean that a state x(t) can be reconstructed from the observations y(t), y(t-1), etc., using the state dynamics and the measurement equation. To test the observability, similar mathematical procedures are available as for testing the controllability.

An example of (non-)observability is the example of someone standing in front of an elevator. The display (measurement) shows only the floor at which the elevator is but it does not show the full state (consisting of the position and the speed of the elevator). The measurement in this case does not give the full state, and even for the position it gives only a rough estimate as it shows only integers. Nevertheless, watching the display for some time, it is possible to 'guess' the full state of the elevator: the measurement *signal* gives also information about whether the elevator is moving or not, and about its position between the floors. In this case the speed of the elevator is said to be observable even if it is not measured directly.

An important element in observability is knowledge about the system dynamics: the person in front of the elevator uses his knowledge about the dynamics of the elevator combined with the measurements to estimate the state. It may also happen that the measurements don not provide enough information about the state, for example if the elevator display would only show the direction of the movement (up or down): it would be impossible to estimate the position in most cases. In such a case the position is unobservable. (Exercise: in which cases would it be possible to estimate the state?)

Similarly in traffic systems the detectors usually do not provide the full traffic state (e.g., density is not provided) and state estimation techniques are needed.

Stability

Stability is one of the most important properties in systems theory. Without going into the mathematical details, stability means that the process state converges to a certain state if the inputs to the system are constant. For an unstable system the state does not converge to a fixed value even if the inputs are constant.

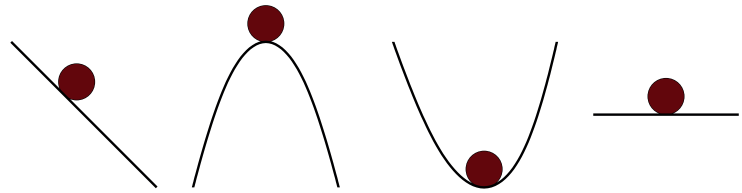


Figure 10-4: The notion of stability makes only sense in the case of an equilibrium.

To analyze stability the notion of equilibrium is essential. A state is in equilibrium if it remains constant for a given (fixed) input. The notion of stability can be further explained by the situations in Figure 1-4. In the leftmost situation the ball is not in equilibrium because the state of the ball is not constant. Therefore this state is not stable. The second left situation represents an equilibrium since the state of the ball is constant. However this equilibrium is said to be unstable, since even the smallest disturbance will cause the ball roll away, and the state of the ball will not converge back to the equilibrium point. The situation to the right from the middle represents a stable equilibrium since the ball will roll back to the equilibrium state after its position has been disturbed. The rightmost figure is said to be meta-stable as multiple equilibrium points exist and a disturbance will always bring the ball in one of the equilibrium states.

The calculation of the equilibrium points is straightforward. Given a state-space system

$$x(t+1) = f(x(t), u(t)), \quad \text{for} \quad t \in \mathbb{N} \tag{10-8}$$

the equilibrium points are given for a certain combination of u^* and x^* (where $u^* = u(t)$, $x^* = x(t)$ for $t > t^*$, for some $t^* \in \mathbb{N}$. Since the state is constant the expression becomes

$$x^* = f(x^*, u^*), \tag{10-9}$$

which can be solved for x^* if u^* is given, or for $u*$ if x^* is given. In some cases there may be multiple solutions or no solution, which means that the system has multiple equilibrium points or no equilibrium points respectively.

10-4-4 Route guidance case

Let us now consider the route guidance case presented in Section 10-2 in more detail. In this section, we will present the model that will be used to describe the system dynamics. We will also define the state of the system and derive the state equation for the discrete-time system.

In the route guidance case, the input of the system is the route guidance provided to the non-captive drivers. The disturbances of the system are the total demand *i(t)* that is present at Valburg and the autonomous traffic demands p_1 and p_2 since these demands are inputs to the system, but cannot be influenced by the control measures.

The system is in this case the combined user-response to the guidance (dependent on amongst other things the compliance to the guidance given) and the resulting system dynamics (e.g. queues occurring at the bottlenecks on the Grijsoord route and the Pleyroute).

Let us assume that on both route 1 and route 2 travel times are constant except for the delays encountered in the bottlenecks. Let us define $r_j(t)$ by the number of vehicles in queue j at time instant t. Changes in the queue lengths and thus in the waiting delays follow from the balance equations (conservation of vehicles).

At time t (in minutes), the number of vehicles $\mathrm{r}_j(\mathrm{t})$ at bottleneck j increases with the number of vehicles arriving at the bottleneck (equaling $p_j(t) + i_j(t - T_j)$). The length of the queue decreases according to the capacity C_j of the bottleneck. Respecting the non-negativity constraint of the queue-length, this results in (see Hoogendoorn and Bovy (1997)):

$$r_j(t+1) = r_j(t) + \max\{p_j(t) + i_j(t - T_j) - C_j, 0\} \tag{10-10}$$

for routes j=1,2.

Eqs. (10-11) and model route-choice when *route-guidance* is operational:

$$\alpha(t) = 1 - \gamma \cdot u(t) \tag{10-11}$$

and

$$i_1(t) = \alpha(t) i(t) \text{ and } i_2(t) = (1 - \alpha(t)) \cdot i(t). \tag{10-12}$$

The variable $0 \leq u(t) \leq 1$ describes the fraction of the sample period t in which the route guidance indicates the Pleyroute as 'best alternative'. If each of the drivers complies to the guidance provided, the compliance rate γ equals one.

State and state dynamics for the test-case example

Due to the *delays* T_1 and T_2, Eq. (10-10) is not a valid state-space description (form of Eq. (10-6)). This can be remedied by defining an appropriated state-vector $x(t)$. Since the state of the system in t-th sample period should capture the entire history of the system, the state can be defined as follows:

$$x^{(1)}(t) = \begin{pmatrix} r_1(t) \\ i_1(t-1) \\ \vdots \\ i_1(t-T_1) \end{pmatrix}, \quad x^{(2)}(t) = \begin{pmatrix} r_2(t) \\ i_2(t-1) \\ \vdots \\ i_2(t-T_2) \end{pmatrix}, \quad \text{and} \quad x(t) = \begin{pmatrix} x^{(1)}(t) \\ x^{(2)}(t) \end{pmatrix} \tag{10-13}$$

Note that the length of the state vectors $\mathrm{x}^?$ and $\mathrm{x}^?$ are respectively $\mathrm{n}_1 = \mathrm{T}_1+1$ and $\mathrm{n}_2 = \mathrm{T}_2+1$. In the remainder, we will use the fraction of vehicles u(t) (or equivalently the fraction of a period) using the (alternative) Pleyroute as the control input. The traffic demand $\mathrm{i}_1(\mathrm{t})$ for the Grijsoord route then equals:

$$i_2(t) = \gamma \cdot u(t) \cdot i(t) \text{ and } i_1(t) = i(t) - i_2(t) \tag{10-14}$$

It is left to the reader to show that the state dynamics of the substate $\mathrm{x}^{(j)}(\mathrm{t})$ become:

$$x^{(1)}(t+1) = \begin{pmatrix} x_1^{(1)}(t) + \max\{p_1(t) + x_{n_1}^{(1)}(t) - C_1, 0\} \\ (1 - \gamma \cdot u(t)) \cdot i(t) \\ x_2^{(1)}(t) \\ \vdots \\ x_{n_1-1}^{(1)}(t) \end{pmatrix} = f^{(1)}(x^{(1)}(t), u(t)) \tag{10-15}$$

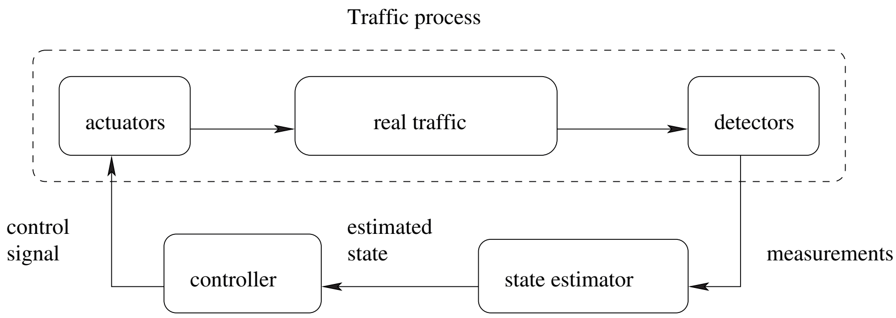


Figure 10-5: Block diagram of a general traffic control system.

and:

$$x^{(2)}(t+1) = \begin{pmatrix} x_1^{(2)}(t) + \max\{p_2(t) + x_{n_2}^{(2)}(t) - C_2, 0\} \\ \gamma \cdot u(t) \cdot i(t) \\ x_2^{(2)}(t) \\ \vdots \\ x_{n_2-1}^{(2)}(t) \end{pmatrix} = f^{(2)}(x^{(2)}(t), u(t)) \tag{10-16}$$

Since:

$$x(t+1) = \begin{pmatrix} x^{(1)}(t+1) \\ x^{(2)}(t+1) \end{pmatrix} = \begin{pmatrix} f^{(1)}(x^{(1)}(t), u(t)) \\ f^{(2)}(x^{(2)}(t), u(t)) \end{pmatrix} = f(x(t), u(t)) \tag{10-17}$$

a valid state-space description has thus been determined.

10-5 Control theory and applications

Control theory is concerned with the study and control of input/output phenomena. The emphasis is on the system's dynamic behavior, i.e. how do characteristic features of these phenomena (such as input, state and output) changes over time, and what are the relationships (as functions of time).

A real-life traffic control system consists of various blocks, namely the real traffic system, the monitoring systems (*traffic sensors*), the state estimator, the controller (*computation of control signals*), and actuators. Figure 10-5 depicts a general traffic control system. The different elements depicted in the figure are briefly discussed in the ensuing sections.

Note that from a *systems and control* point of view the controlled process includes the actuators and sensors too, and the signals between the sensors, state estimator, controller, and actuators are purely mathematical quantities, while the signals between the actuators, traffic system and sensors are physical quantities (lights, inductive loops, etc.).

10-5-1 Actuators

Actuators are the actual physical objects which execute the control signals determined by the controller. Basically, the actuators provide a controller interface to influence the system.

Examples of actuators are the VMS-signs, a DRIP or radio message providing travel time information or a ramp metering installation.

10-5-2 Sensors

Monitoring is the collection of real-life data from the system using sensors. These sensors basically provide an *interface* with the real-life system to the controller. Considering traffic systems, data collection is performed using various sensors, which measure the response of the system. Examples of such sensors in the DTM research field are pneumatic tube detectors, infrared detectors, induction loops, radar detection, video cameras, instrumented (probe-) vehicles, etc.

In the Netherlands, the main sources of information are the induction loops from the MTM (*Motorway Traffic Management*) system. Other systems are the MONICA (MONItoring Casco) system and the MIRA (Motorway Information in the Region of Amsterdam). These data convey among other things traffic counts and speed measurements during a specific collection interval, which is usually one minute. Another information source are video cameras in tunnels (incident management purposes).

10-5-3 State estimation

The raw data from the sensors are in general corrupted by noise. That is, the sensors may not detect the presence of a vehicle (miscount), may detect a vehicle when no vehicle is present (false count). In addition, speed measurements may also be inaccurate, for instance due to the finite precision of the data processing hardware. State estimation techniques aim to estimate the state of the system from the traffic measurements.

To improve the quality of the data, data cleaning techniques are available. For example, a measurement may be qualified using the fundamental diagram: if the one minute average speed measurement does not correspond to speed expected from determination using the fundamental diagram and the measured flow-rate, an observation can be discarded. However, more elaborate techniques exist as well. For instance, Kalman filtering techniques optimally combine measurements and modelling results to provide estimates for the current state (see chapter on traffic monitoring). However in DTM applications currently operational in the Netherlands, state of the art filtering techniques are not commonly applied. In the literature, a number of applications of filtering techniques can be found, e.g. Smulders (1989a), and DACCORD (1996).

10-5-4 Controllers

A large number of control methodologies exist. These include feed-forward, (state) feedback control, predictive control, and optimal control. This section discusses these methods to compute the *control signal*. A control signal is a sequence of instructions at instants or intervals t, which (using the actuators) aim to control the system under consideration.

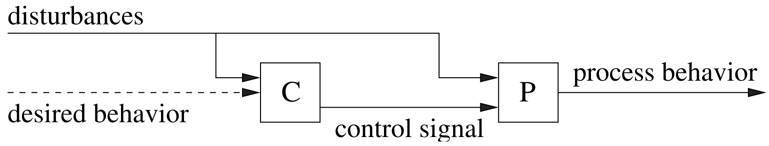


Figure 10-6: The block diagram of a feedforward control structure. The controller only takes the measurement into account, not the resulting process outputs. C indicates the controller and P the process.

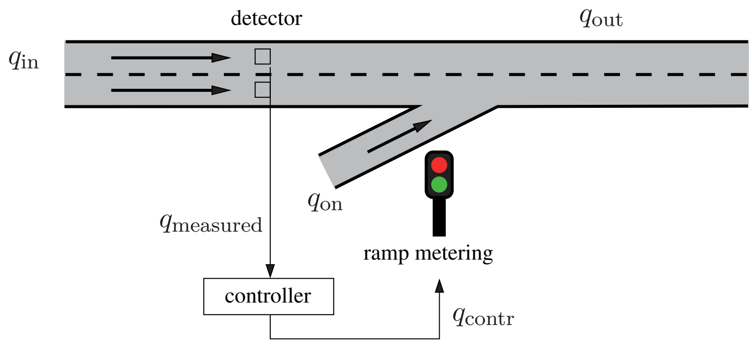


Figure 10-7: The lay-out of the demand-capacity algorithm. In free-flow the control structure is feedforward.

Feed-forward control

Feed-forward control (also called *open loop control*) methodologies do not incorporate information regarding the (estimated) current state x(t) or output y(t) of the system, but only regarding the measurable disturbances d(t) as shown in Figure 10-6.

The advantage of feedforward controllers is that if the process is stable and the controller is also stable (which is a matter of proper design) then the combination of the two systems in a feedforward structure will always result in a stable overall system.

Example

Examples of time-dependent open loop control (where time is considered as the disturbance) are overtaking prohibition for trucks during peak hours, tidal flow lanes, and disclosure of shoulder lanes during peak hours. Also, open loop control has frequently been applied in non-traffic dependent signalized intersection control (i.e., fixed or time-dependent control).

A dynamic example in current traffic management, applications of feedforward control is the Rijkswaterstaat algorithm (demand-capacity algorithm) which computes the remaining capacity by subtracting the flow-rate measurement (the disturbance) on the freeway upstream of the on-ramp from the capacity of the bottleneck caused by the on-ramp (Figure 10-7). This remaining capacity determines the number of vehicles per minute allowed to enter the freeway (control signal). In free flow the effect of the ramp metering (on the traffic state) cannot be measured at the upstream detector location so the control is of the feedforward type.

Route guidance case For the case-study, a feed-forward control would be to apply no control (u(t)=0), or to always guide 30% of the non-captives along the Pleyroute (u(t)=0.3), or measuring the disturbance *i(t)* only, but not the resulting queues. Figure 1-8 shows an

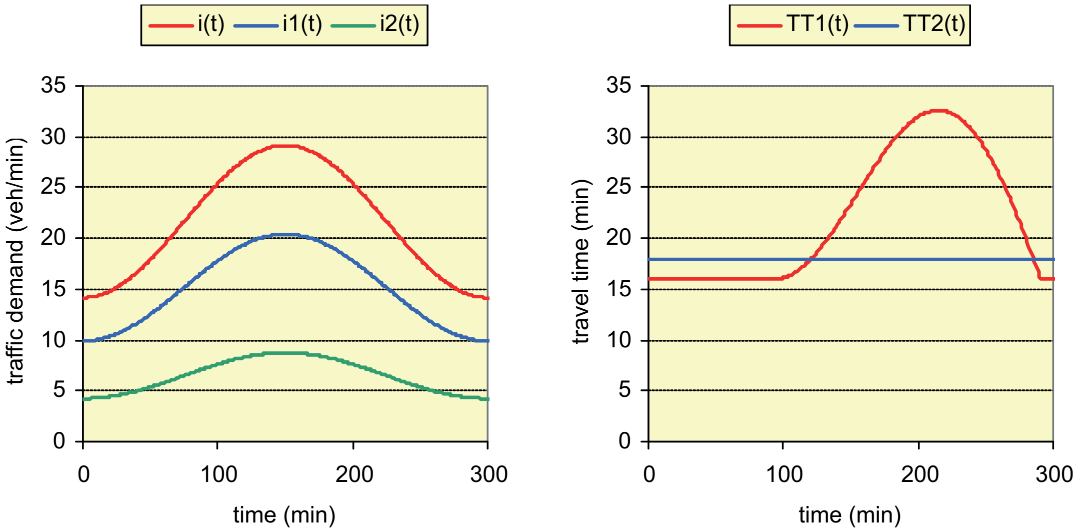


Figure 10-8: Example of open-loop control where u(t) = 0.4 (40% of the drivers are re-routed via the Pleyroute).

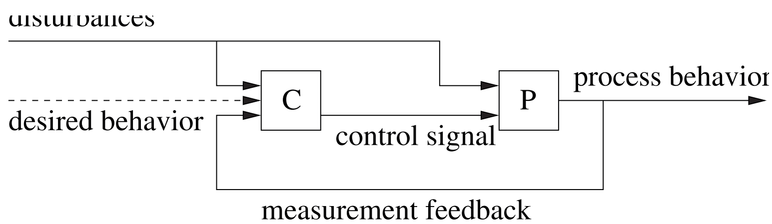


Figure 10-9: The block diagram of a feedback control structure.

example of static open-loop control. Note that open-loop control can be dynamic in the sense that it depends on time or the disturbances; it does however never depend on the current state x(t) of the system.

Feedback control

Another standard control methodology is *feedback control*. Feedback control uses the *measurements y(t) or the estimations of the current state x(t)* to determine the control signal *u(t)* as shown in Figure 10-9. The feedback connection creates a loop in the block diagram. For this reason feedforward control is also called open-loop control, and feedback control is called closed loop control.

In general the measurements $y(t)$ are fed back to the controller in order to determine the control signal $u(t)$. A special case where the state is directly measurable, or can be estimateed from the available measurements, and is used to determine the control signal, which is called *state feedback*.

For example, a linear state feedback controller employs the error $e(t)= x(t)\text{-}x_0$ and a gain K to determine the control $u(t)\text{-}u_0\text{=-}Ke(t)$ or $u(t)$ =-$Ke(t)+u_0$, thereby aiming to steer the

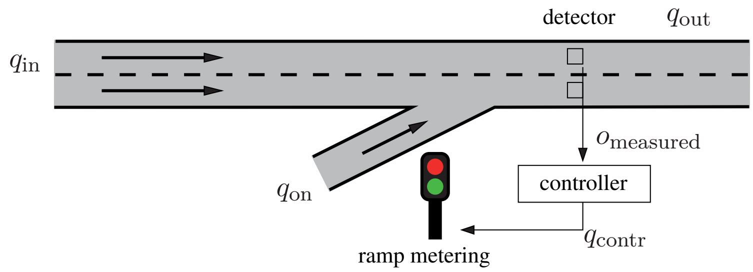


Figure 10-10: In the ALINEA algorithm the detector is placed downstream of the merge area. The effects of the control actions are measured by the detectors, so there is a loop in the information flow, which means feedback.

system to the desired state x_0 (where u_0 is the equilibrium input corresponding to x_0).

The advantage of feedback controllers is that the results of the control actions and the consequences of unmeasurable disturbances (which can only be seen in the process output) will be used in the subsequent control actions. This usually leads to better performance.

The disadvantage is that the feedback *loop* may lead to instabilities even if both the controller and the process are stable by themselves. The example of the route guidance case at the end of this section is such a case. Proper design of the controller can prevent instabilities.

Examples

In current traffic management, applications of state feedback control are numerous. For example:

- *Automatic incident detection*: in the Netherlands, the speed regulations on the VMS signs above the freeway are determined from the current speed measurements of the accompanying detection loops, and the current state of the first VMS downstream. That is, if the downstream VMS indicates a 50km/h speed limit, the current VMS can only display 50km/h or 70km/h (in order to prevent accidents due to vast differences in speed limits)
- The ramp metering algorithm ALINEA (as shown in Figure 10-10) uses a detector downstream of the merge area of the freeway and the on-ramp. Varying ramp metering rates can be observed by the detectors in the form of occupancy changes, in both free flow and congested flow. The control structure is therefore of the feedback type.

Route guidance case

For the route guidance case presented in this chapter, we can define the desired state as the situation in which the instantaneous travel times on both routes are equal, i.e.:

$$e(t) = TT_1(t) - TT_2(t) = T_1 + \frac{r_1(t)}{C_1} - \left(T_2 + \frac{r_2(t)}{C_2}\right) = T_1 - T_2 + \frac{x_1^{(1)}(t)}{C_1} - \frac{x_1^{(2)}(t)}{C_2} \quad (10\text{-}18)$$

The feedback law would then be:

$$u(t) = K \cdot e(t) \;\text{ subject to }\; 0 \leq \mathrm{u(t)} \leq 1 \quad (10\text{-}19)$$

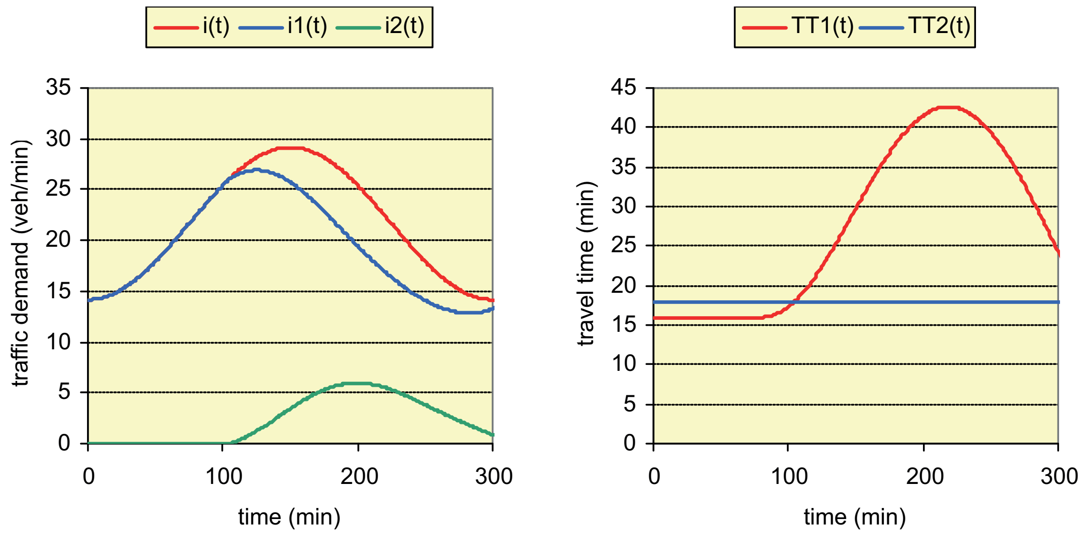


Figure 10-11: Example traffic flow operations of feedback control with $K = 0.01$.

Figure 10-11 shows the resulting traffic conditions when a feedback controller with K=0.01 is used. The figure shows both the traffic demand i(t) and the traffic demands for the respective routes, assuming that all drivers comply with the guidance given ($\gamma = 1$). The figure clearly shows that in this case some of the traffic is rerouted via route 2.

When the feedback gain is increased, undesired effects may occur. Figure 10-12 shows the results of applying feedback control is case K=0.2. It turns out that the control becomes oscillatory (unstable). This has to do with the fact that the system response to the control u(t) is in fact delayed, i.e. it takes time before the control affects the queues at the bottlenecks. For instance, if at some time t route 2 is optimal (in instantaneous travel time terms), traffic will be diverted towards route 2. Since it will take time before this traffic reaches the bottleneck (in fact, it will take T_2 minutes), the instantaneous travel time on route 2 will only be affected after T_2 minutes. In the meanwhile, traffic will still be rerouted towards that route.

This example shows the important result that high feedback gain (large K) in combination with a delay causes instability in the system.

Predictive feedback control

The key to predictive control is the use of forecasting models, predicting the future state evolution. Often, these models are given in the state-space forms (1-1) or (1-2)

Figure 10-13 depicts the scheme of a general model-based predictive feedback controller. The figure shows that both, the current state estimate and the demand prediction (or anythign else that is necessary) are input for the prediction model. Using these inputs, the model predicts the traffic operations for a specific future time horizon N_{p}. That is at instant t a state prediction $x_{\mathrm{pred}}(s)$, for $t \leq s \leq t + H$ is determined using the current state estimate at time t, demand predictions with respect to the traffic conditions within the planning period, and the prediction model. The predictive feedback controller determines the actual control

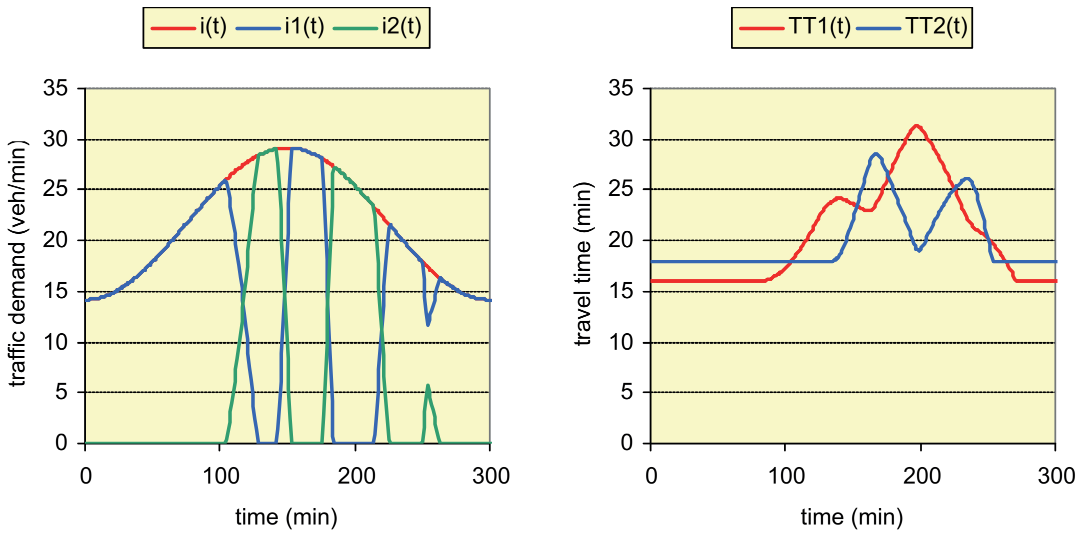


Figure 10-12: Example traffic flow operations of feedback control with $K = 0.2$.

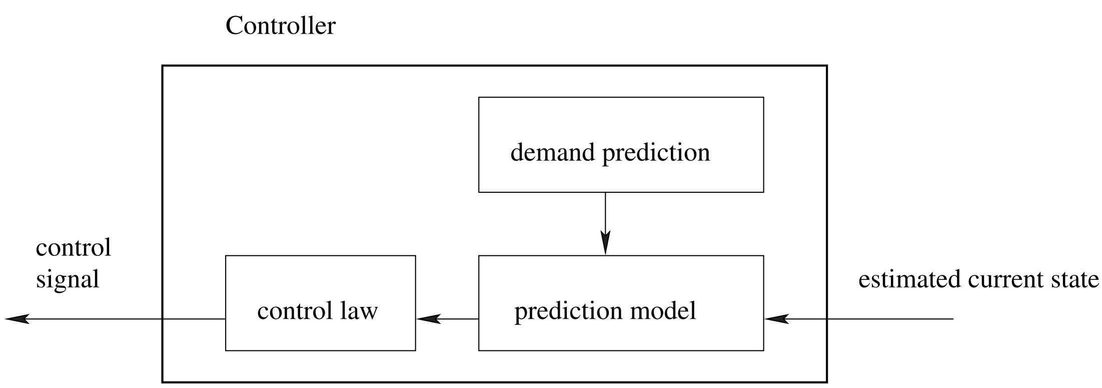


Figure 10-13: A predictive feedback controller

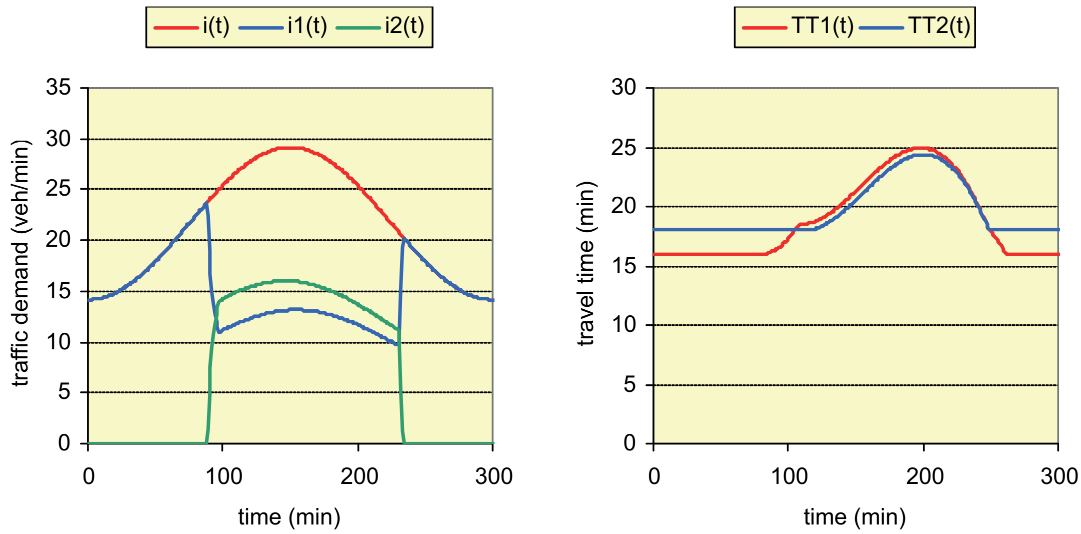


Figure 10-14: Example predictive control with K = 1.

signal based on these predictions. For example, the control signal can be based on simple feedback algorithms: $u(t) = Ke(t+T)$, where T is a certain constant time, and $e(t+T)$ is the prediction T timesteps ahead of the error.

Examples

An example of predictive control is the provision of travel time predictions to road users. In this case, a model (e.g. a neural network) is used to predict the travel time drivers will experience, given the current state of the system and the envisaged system dynamics.

Route guidance example

For the route guidance example, a simple form of predictive control entails using the *predicted* difference in the queue lengths. This means that traffic arriving at the decision point near Valburg is guided in line with the predicted difference in predicted travel times, i.e.:

$$e(t) = T_1 + \frac{r_1(t+T_1)}{C_1} - \left(T_2 + \frac{r_2(t+T_2)}{C_2}\right) = T_1 - T_2 + \frac{x_1^{(1)}(t+T_1)}{C_1} - \frac{x_1^{(2)}(t+T_2)}{C_2} \quad (10\text{-}20)$$

The feedback law would then again be:

$$u(t) = K \cdot e(t) \ \text{ subject to } \ 0 \le \mathrm{u(t)} \le 1 \quad (10\text{-}21)$$

An example of application of predictive control is shown in Figure 10-14. The figure shows that the oscillatory behavior that resulted from feedback control has now disappeared completely. This is a well-known characteristic of predictive control, which is caused by the fact that the controller anticipates on future traffic conditions. Also note that the travel times on both routes are now nearly the same.

Model Predictive Control (MPC)

Model predictive control is a form of feedback control. Also in case of MPC prediction models are used to predict the future state evolution. However, in MPC the control signal is

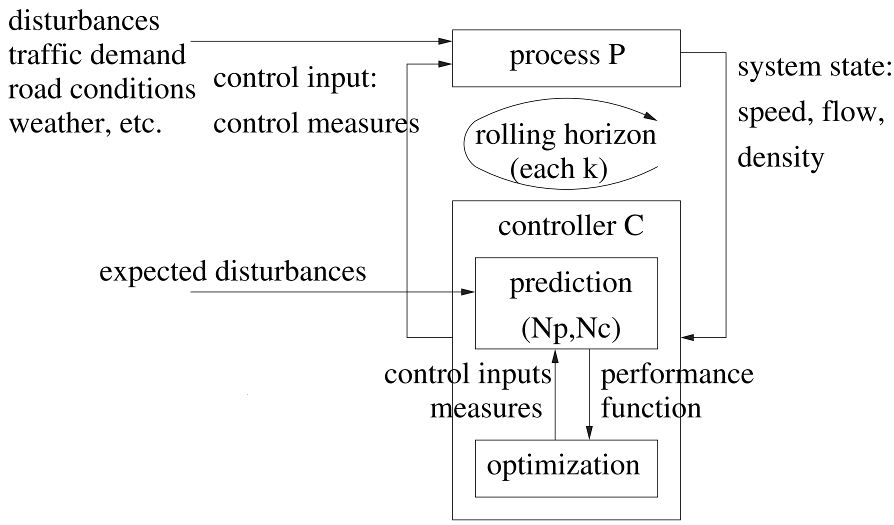


Figure 10-15: The conceptual scheme of an MPC controller

not determined by a control law, but by optimization.

Figure 10-15 and 10-16 depict the schematics of a general model-based predictive controller. The figure shows that both, the current system state and additional data (disturbances, including the predicted traffic demand) are input for the prediction model. Using these inputs *and an assumption on the future control signal* (e.g. the speed limits that will be applied during the prediction horizon), the model predicts the traffic operations for a specific future time horizon N_p. That is at time step k a prediction of the state $\hat{x}(\tilde{k})$ is made, for $k \leq \tilde{k} \leq k + N_\mathrm{p}$ using the current state estimate at time step k and the prediction model, and the control signal plan $u(k), \ldots, u(k + N_\mathrm{p})$. From this prediction the predicted performance of the system can be derived, such as the total travel time. Now, since the prediction depends on the control signal plan, the future performance can be optimized by selecting the best possible control plan. When this plan is found, it is applied to the real process, until a new measurement (estimate) of $x(k)$ is available. Then the whole prediction-optimization process is repeated. This repeated prediction is called *rolling horizon* since the prediction horizon shifts a bit to the futere every time that there is a new measurement available.

This approach is flexible in the choice of the control objective. Basicall any function can be chosen, that accurately expresses the performance measure that the user (road manager) find important. Several objectives were discussed in Chapter 9. By choosing the objective function appropriately, various objectives can be incorporated by means of weighting. These objectives can reflect the goals of different actors involved in the control process. For instance, when ramp metering is considered, the national road authority will often have the objective to reduce congestion on the main road. However, the municipality will not want the queues on the on-ramps to block the surface street intersections.

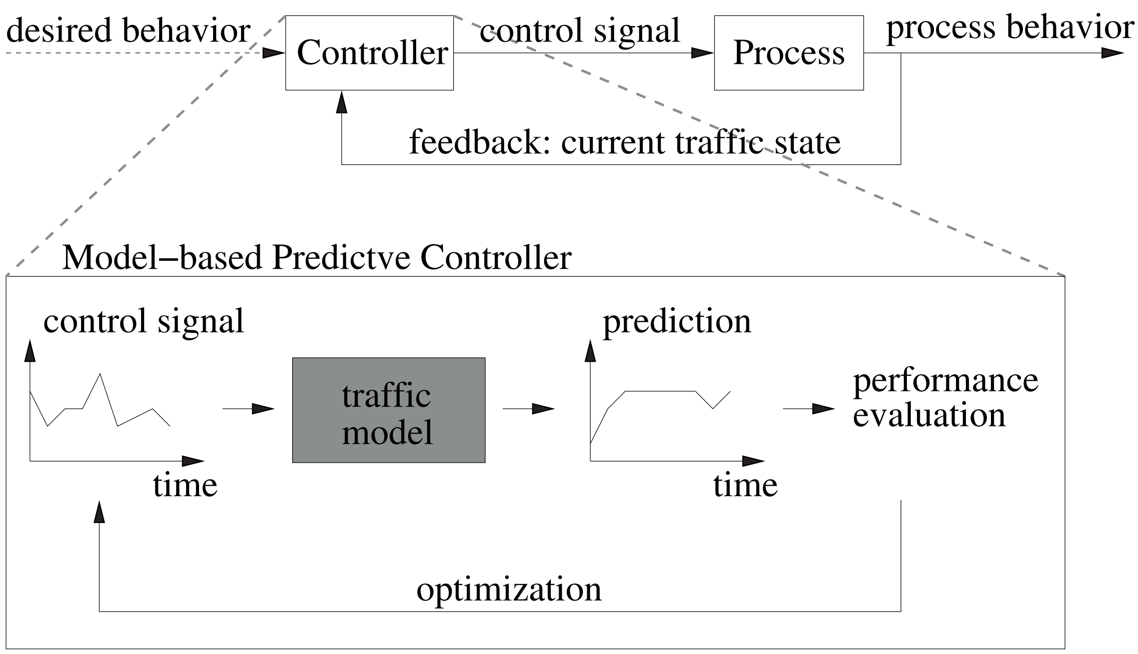


Figure 10-16: The optimization process in the MPC controller.

Chapter 11

Ramp metering

Ramp metering is one of the most investigated and applied freeway traffic control measures. In this chapter we discuss the technical realization of ramp metering – to some depth, but more important, we discuss the relevant traffic phenomena on freeways that play a role in ramp metering, and the main approaches that exist in theory and practice to ramp metering.

11-1 Technical realization of ramp metering

A ramp metering set-up is implemented as a traffic signal that is placed at the on-ramp of a freeway as shown in Figure 11-1. The required metering rate is implemented by appropriately choosing the phase lengths of the traffic signal. Several ramp metering implementations can be distinguished Chaudhary and Messer (2000), e.g., single-lane with one vehicle per green ramp metering, single-lane with multiple vehicles per green ramp metering (bulk metering), and dual-lane ramp metering. In The Netherlands the typical implementation allows one car per green per lane (with a cycle length of few seconds). In other countries there exist implementations that allow two or more cars per green (with a longer cycle length of, e.g., 60 s).

For the one-car-per-green type of ramp metering the green time is determined by the time that the actual vehicle needs to pass the stop line when the light gets green. The passage of the stop line is measured by a loop directly downstream of the stop line, and when a passage is detected, the light immediately turns yellow. The yellow light is only used due to legal regulations, and the yellow time is short. To achieve a desired flow of q_{contr} the cycle time has to be

$$TC = n_{\mathrm{lanes}} \frac{1}{q_{\mathrm{contr}}} \tag{11-1}$$

for a ramp metering set-up with n_{lanes} lanes. Since

$$TC = TG + TY + TR \tag{11-2}$$

Figure 11-1: Ramp metering at the freeway A13 in Delft, The Netherlands. One car may pass per green phase. To prevent red-light running the control is enforced.

and the green and yellow times are given, the only possibility to achieve the desired cycle time is by selecting an appropriate red time.

In the Netherlands the typical minimum red and amber (yellow) times are respectively 2 s and 0.5 s, which define the maximum flow achievable by ramp metering, being around $3600\,\mathrm{s.h}^{-1}$ / $2.5\,\mathrm{s.veh}^{-1} = 1440$ veh.h^{-1}.

11-2 Relevant traffic phenomena

There are several traffic phenomena that are related to on-ramps and therefore also to ramp metering. Most of these phenomena describe the various circumstances in which jams at on-ramps are created. Depending on the jam creation mechanism the ramp metering solution that prevents or resolves the jam may be different. In this section we discuss the phenomena and the general ramp metering approach that could be used for the given phenomenon.

11-2-1 Breakdowns

One of the goals of ramp metering is to prevent or resolve a traffic breakdown (creation of a traffic jam). Here we discuss why a breakdown is undesired and what traffic scenarios may cause a breakdown at an on-ramp.

The main reason why a breakdown is considered to be undesired is related to the *capacity drop.* Capacity drop is the phenomenon that the outflow of a traffic jam is significantly lower

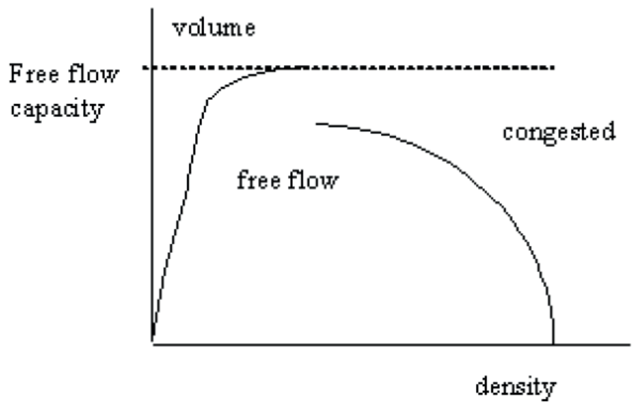


Figure 11-2: A schematic fundamental diagram that represent the capacity drop

than the maximum achievable flow at the same location in free flow. A fundamental diagram that shows a capacity drop is given in Figure 11-2.

Capacity drop at fixed bottlenecks has been reported in field studies Cassidy and Bertini (1999b); Hall and Agyemang-Duah (1991), and the decrease in flow ranges from 0–15 %. Since the capacity drop is not observed in all cases, traffic data from the bottleneck that is to be controlled has to be studied carefully.

Ramp metering may be used in order to maintain the capacity that corresponds to free flow. In the first place ramp metering should aim at *preventing* a breakdown and the corresponding capacity drop. In the case that prevention has failed, the second aim of ramp metering should be to *restore* free flow.

A breakdown can be caused in several traffic situations. The most important situations are:

- **High traffic demand at on-ramp and freeway.**
  Obviously one of the main causes of jams at on-ramps is that the demand on the freeway and the on-ramp exceed together the capacity of the freeway section at or downstream of the on-ramp. If this is due to a relatively slow increase in demand, such as at the beginning of the peak hour, then ramp metering can be switched on before the demand exceeds the capacity.

  In such a case the ramp demand is limited in order to keep the freeway in efficient operation. A possible consequence is that the ramp queue will grow, and may block urban intersections, which is also undesired. In practice this is often prevented by setting the ramp metering cycle time to its minimum when the ramp queue length has exceeded a certain threshold value. A detector can be placed at the location which the queue should not exceed.

- **High-flow, high-density forward waves on the freeway** Even if the total freeway and on-ramp demand does not exceed the capacity, a forward propagating wave with high speed and high density (a group of vehicles traveling close to each other in free flow) may trigger a ramp jam. An example is shown in Figure 11-3.

  A solution might be to limit the on-ramp flow for the time that the forward wave is around the on-ramp.

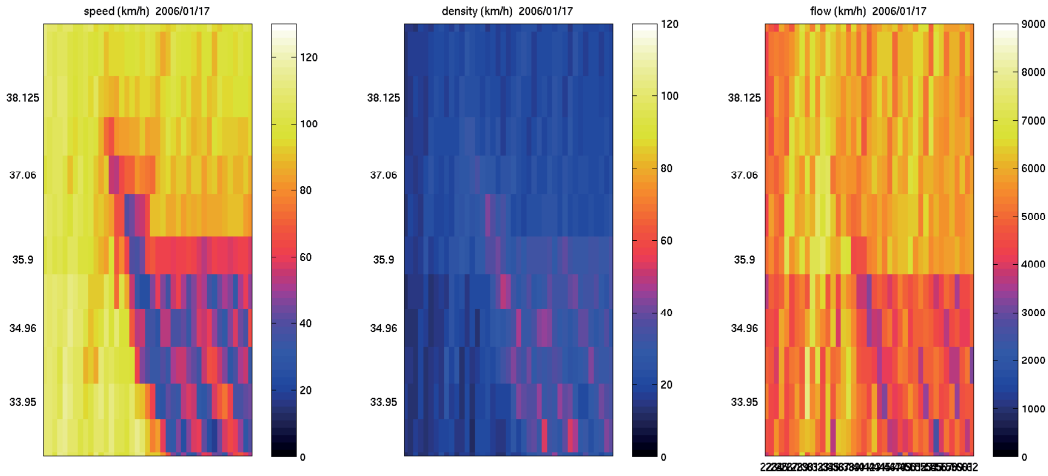


Figure 11-3: A forward propagating wave triggers a ramp jam at 35.9 km. The forward wave first creates a high-density area downstream of the on-ramp, which starts to propagate backward when the breakdown has fully occurred.

- **Platooning: peak in on-ramp demand**
  In many cases there is a signalized intersection at the upstream end of the on-ramp. This causes that the vehicles enter the on-ramp in a platoon in the green phase. If this platoon enters the freeway at once, it may create a breakdown due to the sudden peak in the on-ramp demand. A ramp metering approach that spreads the vehicles, but does not limit the flow, would be sufficient in this case:

$$q_{\text{contr}} = q_{\text{demand}}. \tag{11-3}$$

- **Back-propagating moving jams**
  Back propagating moving jams can also trigger jams at on-ramps. Examples are given in Figure 11-4. The concept of the solution would be similar to the case of the forward propagating waves: to limit the ramp flow for the short period that the backward propagating wave is passing by. However in the case of the backward propagating jam an earlier limitation of the ramp flow may even reduce the length of the moving jam, as a lower ramp flow reduces the inflow to the jam as well.

- **Slow traffic**
  As shown in Figure 11-5 slow traffic may also trigger jams at on-ramps. An example of slow traffic is a snow plow or the transport of exceptionally large or heavy goods (such as parts of bridges). The exact mechanism why or under which conditions jams are triggered by slow traffic is not investigated in scientific literature yet. However slow traffic does not frequently occur on freeways, and therefore the relevance of developing targeted ramp metering algorithm that prevent breakdowns in these cases is relatively low.

11-2-2 Off-ramp blocking

Another reason for ramp metering is that the tail of a jam that is created at an on-ramp may block the traffic that wants to leave the freeway at an upstream off-ramp as the example

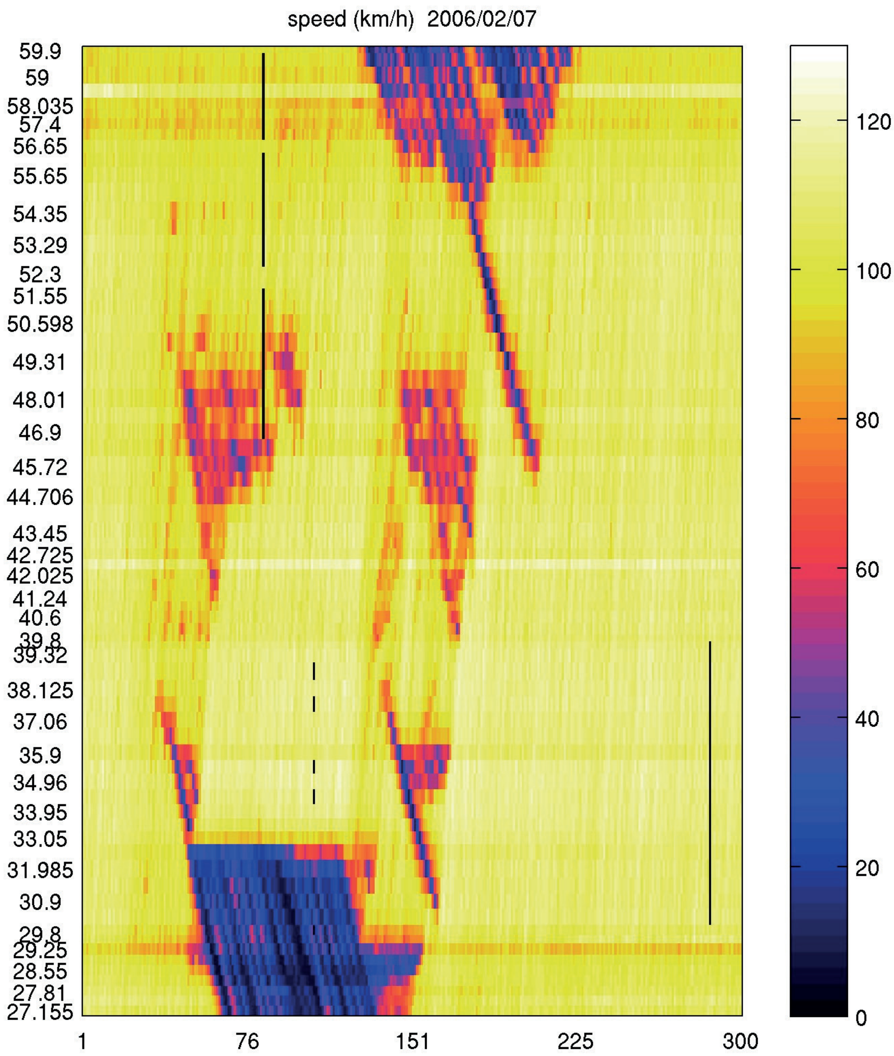


Figure 11-4: Two examples of back-propagating jams that trigger on-ramp jams: (1) at 60 min and 33.5 km, and (2) at 140 min and 35.9 km.

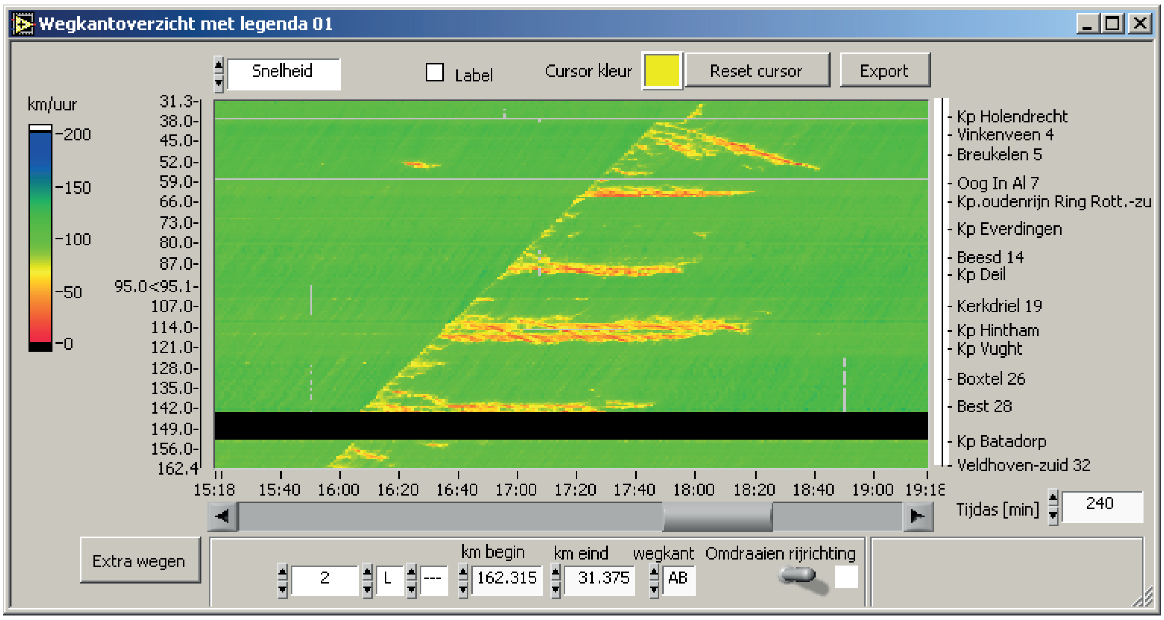


Figure 11-5: Slow traffic travels over a long freeway. At several on-ramps jams are triggered that remain existent for a long time.

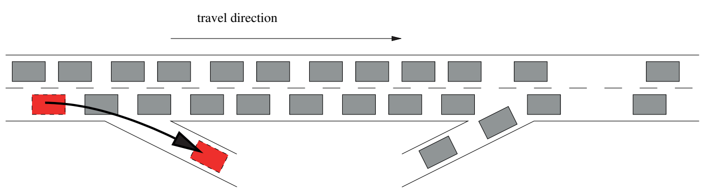


Figure 11-6: Congestion caused by high on-ramp demand could also result in the blocking of an upstream off-ramp.

in Figure 11-6 shows. Off- and on-ramps are often located close to each other. This may substantially reduce the volume that can leave the freeway, which means therefore a reduction of the freeway performance.

Let's compare the outflow of the off-ramp in the situation when it is not blocked, with the situation when it is blocked, as shown in Figures 11-7 and 11-8. Assume that γ is the fraction of vehicles that comes from the upstream freeway and wants to leave via the off-ramp. Let d the on-ramp demand, and q_{cap} the capacity of the on-ramp section. For simplicity, assume that there is no capacity drop. When the off-ramp is not blocked, the outflow is

$$q_{\text{off-ramp}} = \gamma q_{\text{in}} \,. \tag{11-4}$$

When the off-ramp is blocked, then the flow on the freeway just upstream of the on-ramp will be $q_{\text{cap}} - d$, for which holds that

$$q_{\text{cap}} - d = (1 - \gamma)\tilde{q}_{\text{in}} \tag{11-5}$$

where $\tilde{q}_{\text{in}}$ is the flow just upstream of the off-ramp, which is blocked so $\tilde{q}_{\text{in}} \leq q_{\text{in}}$. Then the

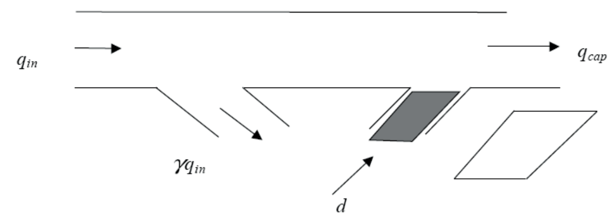


Figure 11-7: The off-ramp is not blocked.

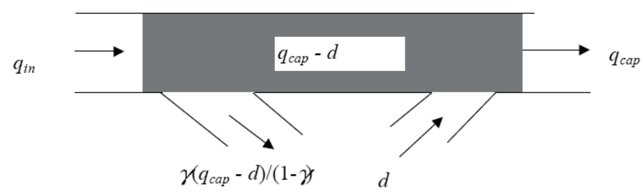


Figure 11-8: The off-ramp is blocked by an on-ramp-related jam.

outflow of the off-ramp is

$$q_{\text{off-ramp}} = \gamma \tilde{q}_{\text{in}} = \frac{\gamma}{1-\gamma}(q_{\text{cap}} - d)\,. \tag{11-6}$$

11-2-3 Route choice effects

Ramp metering in general delays the traffic that enters the freeway via the on-ramp. Consequently some drivers may choose another route if there exist routes with shorter travel times. In areas where on-ramps are closely located (e.g., in urban areas) this may change the drivers' route choice towards other on-ramps. In such cases it is often better to meter all on-ramps, not only one.

Theoretically, ramp metering can also be used to discourage drivers to take an undesired route (from the network manager's point of view). However, this is not done in practice.

11-3 Ramp metering objectives

Summarizing, ramp metering can be used for the following purposes:

- **Prevention of breakdowns.** When traffic is dense, ramp metering can prevent a traffic breakdown on the freeway by adjusting the metering rate such that the density on the freeway remains below the critical value. Preventing a traffic breakdown has not only the advantage that it results in a higher flow, but also that it prevents the creation of a jam that could block the off-ramp upstream of the on-ramp (as shown in Figure 11-6). These effects are studied in detail by Papageorgiou and Kotsialos Papageorgiou and Kotsialos (2002). Daganzo Daganzo (1996) has quantified the role of ramp metering in avoiding the activation of freeway gridlocks.

- **Influencing route choice.** Ramp metering can be implemented to influence the traffic demand and traffic routing. The impact of ramp metering on the traffic state and on the travel times is taken into account by the drivers in their routing behavior Zhang (2007). Banks Banks (2005) has described a theory to apply ramp metering to influence traffic routing to avoid freeway bottlenecks. Based on a similar idea Middelham Middelham (1999) has performed a synthetic study on the route choice effects of ramp metering.

- **Localization of traffic jams.** According to Kerner Kerner (2004a) ramp metering can prevent the back-propagation of traffic jams and shock waves occurring at on-ramps. This could be beneficial on the network level since it could localize the traffic jam, and it could also be beneficial to the traffic throughput.

11-4 Systems and control background notions

Before we go to the specific approaches, we discuss some basic notions from the systems and control field which will be used in the discussion of the ramp metering approaches in the next section.

In systems and control a process is often described by its state, the inputs that can act on the process and the measurable outputs of the system. In our case the state is the speed, flow and density (or at least two of the three) of the traffic on the freeway and the queue on the on-ramp, the inputs are the traffic demands on the freeway and the on-ramp and upstream propagating jams. An input that we can influence is the ramp metering rate. This is called the control input. The inputs that we cannot influence are considered as 'disturbances'. The outputs of the system are for example the detector measurement, such as the flow or the speed. Note that density is an example of a state that is difficult to measure.

The goal of control is to find the control inputs that lead to a desired system state or behavior (which is observed via the outputs). For this two basic structures are used: *feed-forward* and *feedback* control. We discuss these structures below, and their consequences for the stability of the overall system (controller+process).

It holds for both control structures that in general controllers contain some parameters that need tuning. We will discuss the tuning of the parameters of the ramp metering algorithms in Section 11-6.

11-4-1 Feed-forward control

The block diagram of a feed-forward control structure is shown in Figure 11-9 Åström and Wittenmark (1997). The behavior of process **P** can be influenced by the control inputs. As a result the outputs (measurements or observations) show a given behavior. The controller **C** determines the control inputs in order to reach a given desired behavior of the outputs, taking into account the disturbances that act on the process. In the feed-forward structure the controller **C** translates the desired behavior and the measured disturbances into control actions for the process.

The term feed-forward refers to the fact that the direction of the information flow in the system contains no loops, i.e., it propagates only "forward".

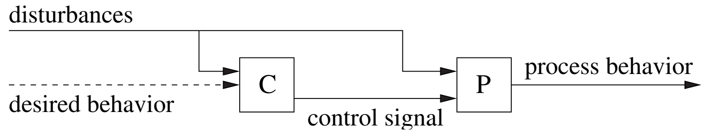


Figure 11-9: The feed-forward control structure.

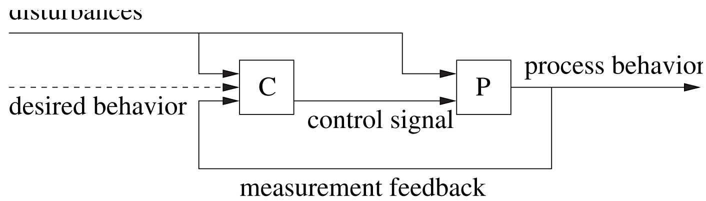


Figure 11-10: The feedback control structure.

The main advantages of a feed-forward controller are that the complete system is stable if the controller and the process are stable, and that its design is in general simple.

11-4-2 Feedback control

In Figure 11-10 the feedback control structure is shown Åström and Wittenmark (1997). In contrast to the feed-forward control structure, here the behavior of the outputs is coupled back to the controller (hence the name feedback). This structure is also often referred to as "closed-loop" control.

The main advantages of a feedback controller over a feed-forward controller are that (1) it may have a quicker response (resulting in better performance), (2) it may correct undesired offsets in the output, (3) it may suppress unmeasurable disturbances that are observable through the output only, and (4) it may stabilize an unstable system. An example of stabilization of an unstable system is a person (=controller) riding a bicycle (= unstable system).

On the other hand a feedback controller may also destabilize a system if the controller is not designed properly. An example is an overreacting cruise control, that accelerates too much, which leads to a too high speed, and then decelerates too much, which leads to too low speeds, and so on. . .

11-5 Control in dynamic traffic management

Dynamic traffic management systems typically operate according to the feedback control concept known from control systems theory, as shown in Figure 11-11. The traffic sensors provide information about the current traffic state, such as speed, flow, density, or occupancy. The controller determines appropriate control signals that are sent to the actuators[1]. The reaction of the traffic system is measured by the sensors again, which closes the control loop. If the new measurements show a deviation from the desired behavior (caused, e.g., by unforeseen disturbances), the new control signals are adapted accordingly. Note that there

[1]Depending on the system the changes in the control signal may be implemented instantly or may need to be phased in.

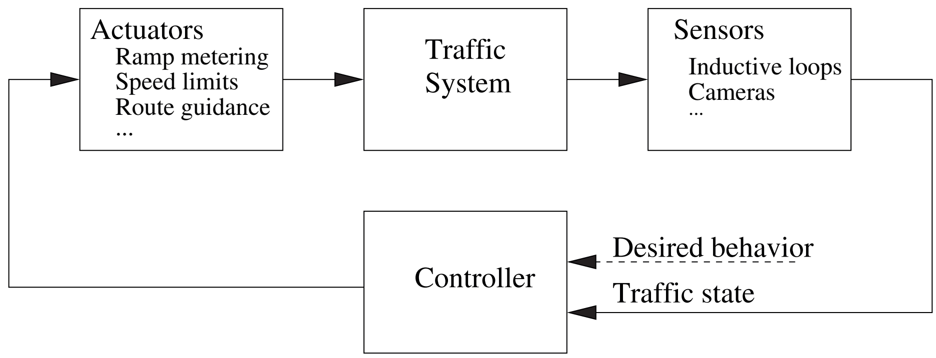


Figure 11-11: Schematic representation of the dynamic traffic management control loop. Based on the measurements provided by the sensors the controller determines the control signals sent to the actuators. Since the control loop is closed, the deviations from the desired traffic system behavior are observed and appropriate control actions are taken.

also exist traffic control systems that have a feed-forward structure, e.g., the demand-capacity ramp metering approach that will be discussed in Section 11-6.

11-6 Ramp metering strategies

The control strategies that have been developed for restrictive ramp metering can be classified as static or dynamic, fixed-time or traffic-responsive.

11-6-1 Fixed-time ramp metering

Fixed-time ramp metering strategies use (possibly time-of-day dependent) fixed metering rates that are determined off-line based on historical demands. This approach was first suggested by Wattleworth Wattleworth (1965), and was extended to a dynamic traffic model by Papageorgiou Papageorgiou (1980). The disadvantage of fixed-time strategies is that they do not take into account the day-to-day variations in the traffic demand or the variations in the demand during a period with a constant metering rate, which may result in under-utilization of the freeway or inability to prevent congestion.

11-6-2 Traffic-responsive strategies

Traffic responsive control strategies solve the issues with fixed-time ramp metering by adjusting on-line the metering rate as a function of the prevailing traffic conditions. These strategies also aim at the same objectives as the fixed-time strategies, but use direct traffic measurements instead of historical data to prevent or to reduce congestion. Below we discuss the most frequently used traffic-responsive ramp metering strategies.

Demand-capacity (RWS)

One of the best known strategies is the *demand-capacity* strategy Papageorgiou and Kotsialos (2000) in which it is assumed that the on-ramp section has a fixed capacity q_{cap} and that the

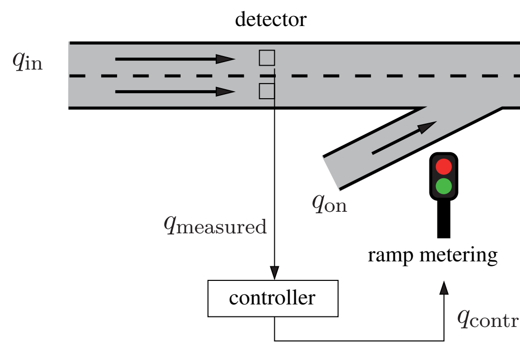


Figure 11-12: The demand-capacity algorithm has a detector upstream from the on-ramp.

difference between the capacity and the upstream measured inflow q_{in} can be admitted via the on-ramp:

$$q_{\text{ramp}}(k) = \begin{cases} q_{\text{cap}} - q_{\text{in}}(k-1) & \text{if } o_{\text{meas}}(k-1) \leq o_{\text{cr}} \\ q_{\text{r,min}} & \text{otherwise} \end{cases} \tag{11-7}$$

with $q_{\text{ramp}}(k)$ the admitted ramp flow at time step k, q_{cap} the downstream freeway capacity, $q_{\text{in}}(k)$ the freeway flow measured upstream of the on-ramp at time step k, $q_{\text{r,min}}$ the minimal on-ramp flow during congestion, $o_{\text{meas}}(k)$ the occupancy downstream the on-ramp at time step k, and o_{cr} is the critical occupancy (at which the flow is maximal).

This algorithm works only in free-flow. (Exercise: explain why it doesn't improve flow in congestion.) Since the traffic state on the freeway (congested or free-flow) cannot be determined based on the measurement of the traffic flow alone, the downstream occupancy is measured in order to determine whether congestion is present ($o_{\text{meas}}(k-1) > o_{\text{cr}}$) or not. Note that consequently the algorithm is not suitable to resolve existing jams, only to (hopefully) prevent possible ones.

This is a feed-forward algorithm, since the effect of the ramp metering is not measured and not fed-back to the controller.

The capacity q_{cap} is a tuning variable that has to be selected by the designer of the controller. In practice capacity is not a crisp value, but rather a range of flow values that can be maintained for some time. The higher this value the shorter the time that it can be maintained. For this reason q_{cap} is tuned somewhat lower than the peak of the fundamental diagram, as shown in Figure 11-13.

Another issue is the selection of the detector location. Since there is a delay of one time step between the measurement and the control action, the detector has to be located such that the traffic that is measured is at the on-ramp location when the corresponding control action is taken. However the required distance between the detector and the on-ramp depends on the traffic speed, and therefore it is not possible to locate the detector properly for all traffic situations.

ALINEA

Another approach is the ALINEA Papageorigiou et al. (1991) ramp metering strategy. This strategy uses a (closed-loop) feedback structure, and therefore it allows for a controller formu-

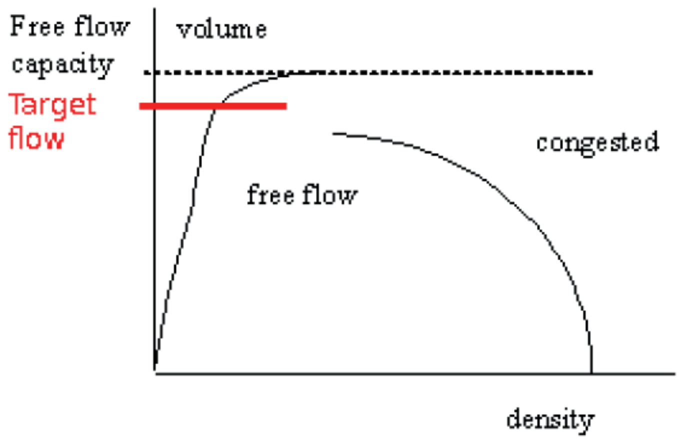


Figure 11-13: The tuning of the demand-capacity algorithm must be somewhat lower than the top of the fundamental diagram.

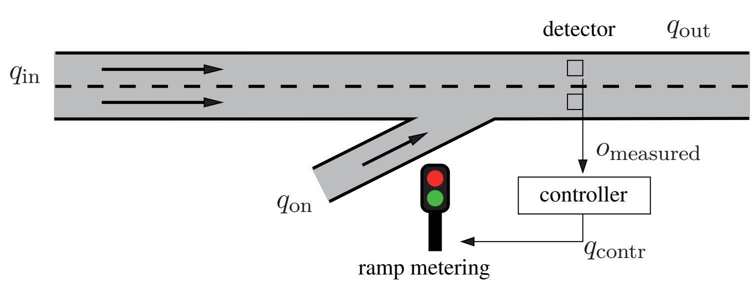


Figure 11-14: The ALINEA algorithm measures the occupancy just downstream of the on-ramp section.

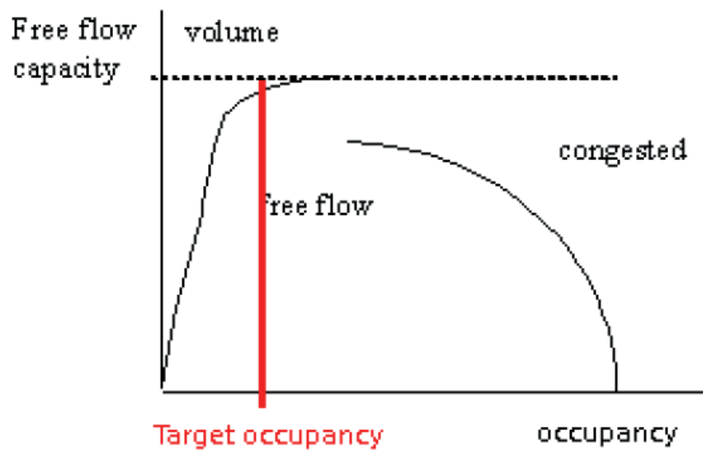


Figure 11-15: The tuning of the desired occupancy in the ALINEA algorithm must be somewhat lower than the occupancy corresponding to top of the fundamental diagram.

lation that can reject disturbances and have zero steady-state error. Its control law is defined as follows:

$$q_{\text{ramp}}(k) = q_{\text{ramp}}(k-1) + K(\hat{o} - o_{\text{meas}}(k)), \tag{11-8}$$

where $q_{\text{ramp}}(k)$ is the metered on-ramp flow at time step k, K is a positive constant, $\hat{o}$ is a set-point (desired value) for the occupancy, and $o_{\text{meas}}(k)$ is the measured occupancy on the freeway downstream of the on-ramp at time step k. ALINEA tries to maintain the occupancy on the freeway equal to a set-point $\hat{o}$, which is chosen in the region of stable operation. As long as the measured occupancy is higher (lower) than the desired value, the on-ramp flow will be decreased (increased) according to 11-8.

The tuning of this algorithm involves the selection of two parameters, the desired occupancy $\hat{o}$ and the controller (feedback) constant K. Given the probabilistic nature of traffic breakdowns (similar tot the reasoning for the demand capacity algorithm), the set-point $\hat{o}$ is often chosen somewhat lower than the critical occupancy in order to guarantee free-flow traffic operation. Although this algorithm is able to resolve a jam[2] it is in general better to prevent a breakdown than to resolve one. See also Figure 11-15.

For the constant K the value can be selected based on a heuristic reasoning. If K is too small then the ramp flow will hardly be adjusted if the measured occupancy is not equal to the desired occupancy. If K too large then the on-ramp flow adjustment will be too strong and will lead to oscillations: when the occupancy is too low, more traffic will be admitted. This will cause an occupancy that is too high, and consequently less traffic will be admitted, and so on...

The detector location should be just downstream of the on-ramp but still in the head of the jam. The algorithm is sensitive to detector location as the typical (which is the reference) occupancy in the jam head depends on the location.

[2] Exercise: explain how.

11-7 Real-world experiences

Several field tests and simulation studies have shown the effectiveness of ramp metering. In Paris on the Boulevard Périphérique and in Amsterdam several ramp metering strategies have been tested Papageorgiou et al. (1997a). The demand-capacity, occupancy, and ALINEA strategies were applied in the field tests at a single ramp in Paris. It was found that ALINEA was clearly superior to the other two in all the performance measures (total time spent, total traveled distance, mean speed, mean congestion duration). At the Boulevard Périphérique in Paris the multi-variable (coordinated) feedback strategy METALINE was also applied and was compared with the local feedback strategy ALINEA. Both strategies resulted in approximately the same performance improvement Papageorgiou et al. (1990). One of the largest field tests was conducted in the Twin Cities metropolitan area of Minnesota. In this area 430 operational ramp meters were shut down to evaluate their effectiveness. The results of this test show that ramp metering not only serves the purposes of improving traffic flow and traffic safety, but also improves travel time reliability Cambridge Systematics, Inc. (2001); Levinson and Zhang (2006).

Chapter 12

SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory

12-1 Learning objectives

This chapter describes the dynamic speed limit control algorithm that was also discussed in the lecture. After reading this chapter the reader should be able to:

- Construct the control scheme in the time-space diagram if the states 1–6 are given in the density-flow diagram.
- Explain in traffic engineering terms the different phases of the control scheme.
- Draw trajectories of vehicles that are travelling through the controlled area.
- Analyse whether a jam wave is resolvable for a given set of states 1–6.
- Explain and draw how the shape of the control scheme changes if one of the following tuning parameters is changed: density of state 4, the effective speed limit (speed of state 3 and 4), the margins on the detected jam wave location, the density and flow of state 5.

12-2 Introduction

On freeways basically two types of traffic jams can occur: jams with the head fixed at a bottleneck location and jams that have an upstream moving head and tail (moving opposite to the travel direction). Here we focus on the second type, which are often called jam waves (Hegyi et al., 2005) or wide moving jams (Kerner and Rehborn, 1996). (Note that by jam wave we mean the whole jam, from head to tail, that propagates opposite to the travel direction. In contrast, a shock wave is just one sharp transition between traffic states. The tail of the jam is an example of a shock wave. The terms shock wave and jam wave are sometimes used interchangeably in literature, which is not entirely correct.) Jam waves are typically short jams (say, 1-2 km) that propagate upstream, due to the incoming vehicles at their tail and the leaving vehicles at their head. These jam waves can remain existent for a long time and distance (Kerner and Rehborn, 1996). As a consequence, every vehicle that enters the freeway upstream of the jammed area will have to pass through the jammed area, which increases travel times, creates potentially unsafe situations, and increases noise and air pollution by braking and accelerating vehicles. Jam waves typically have a significantly lower outflow than the capacity of the freeway, which motivates the idea that traffic flow can be improved by resolving jam waves. The difference between the free flow capacity and the queue discharge rate is typically around 30% (Kerner and Rehborn, 1996).

In the literature two main approaches can be found to dynamic speed limit control aiming at flow improvement. The first emphasizes the homogenization effect (Smulders, 1990; van den Hoogen and Smulders, 1994; Hardman, 1996; Kühne, 1991), whereas the second is focused on preventing traffic breakdown or resolving existing jams by reducing the flow by means of speed limits (Hegyi et al., 2005; Zhang et al., 2005; Popov et al., 2008): The basic idea of homogenization is that speed limits can reduce the speed (and/or density) differences, by which a more stable (and safer) flow can be achieved. The homogenizing approach typically uses speed limits that are above the critical speed (i.e., the speed that corresponds to the maximal flow). So, these speed limits do not limit the traffic flow, but only slightly reduce the average speed (and slightly increase the density). In theory this approach can increase the time to breakdown slightly (Smulders, 1990), but it cannot suppress or resolve jam waves.

The flow reduction approach focuses more on preventing or resolving too high densities, and also allows speed limits that are lower than the critical speed in order to limit the inflow to these areas. The flow reduction by the speed limits is due to two mechanisms: first, at the moment when the speed is reduced the traffic is traveling with the same density but with a lower speed, which leads to a lower flow, and second, the maximum flow that corresponds to speed limits below the critical speed is less than the capacity. The goal of the flow reduction is to resolve jams by limiting the inflow to them. By resolving the jams (bottlenecks) higher flows can be achieved in contrast to the homogenization approach as demonstrated in (Hegyi et al., 2005; Zhang et al., 2005; Popov et al., 2008).

Several control methodologies are used in the literature for speed limit control, such as linear control (Chang et al., 2007; Zhang et al., 2005; Popov et al., 2008), multi-layer control (Li et al., 1995), model predictive control (Hegyi et al., 2005). Other authors use (or simplify their control law to) a control logic where the switching between the speed limit values is based on traffic volume, speed, or density (van den Hoogen and Smulders, 1994; Hardman, 1996; Smulders, 1990; Kühne, 1991).

Most of these approaches are difficult to be applied in practice due to their computational complexity, and the uncertainty about their robustness or to the fact that they incorporate parameters that are difficult to tune. The tuning of a real-life system should be easy, as it implies "experimenting" with real people. The proposed method in this paper is based on a simple principle, has a very low computational demand, and has tuning parameters that have clear physical interpretation. The physical interpretation of the tuning parameters makes tuning straightforward and the proposed method also contains parameters that can improve the robustness of the algorithm.

The proposed algorithm is based on shock wave theory and is called SPECIALIST (SPEed ControllIng ALgorIthm using Shock wave Theory). The theory of the algorithm will be discussed in Section 12-3. The presented theory will be translated to a practically applicable algorithm in Section 12-4. The tuning of the parameters will be introduced and discussed in Section 12-5. The algorithm will be demonstrated in Section 12-6 and Section 12-10 concludes the paper.

12-3 Theory of SPECIALIST

The theory of resolving jam waves by dynamic speed limits is based on the shock wave theory as also applied by Lighthill and Whitham in their famous paper (Lighthill and Whitham, 1955a). Before explaining the approach for jam wave resolution we explain one of the fundamental relationships in shock wave theory.

12-3-1 Shock wave theory

Although shock wave theory goes further than what is presented here, we only present one fundamental aspect, which is necessary to understand the remainder of the paper.

One of the most basic relationships in shock wave theory is the relationship between the time-space graph of the traffic states (as shown on the left in Fig. 12-1) and the density-flow graph (as shown on the right in Fig. 12-1). The time-space graph shows the traffic states on a road stretch (along the vertical axis) and their propagation over time (in the horizontal direction). In the figure a short traffic jam is shown that propagates upstream (area 2) and which is surrounded by traffic in free-flow (areas 1). The density-flow diagram shows the corresponding density and flow values for these states. Shock wave theory states that the front (boundary) between two states in the left figure has the same slope as the slope of the line that connects the two states in the right figure. Note that the slopes in both figures have the unit of km/h. The orange lines (light gray in black and white) indicate the fundamental diagram (as a reference).

The importance of this relationship is that if the different traffic states on a freeway stretch are known, then their future evolution can be predicted by describing the fronts between them. This basic relationship will be used in the theory for resolving jam waves.

12-3-2 Resolving jam waves

The approach to resolve jam waves consists of different phases and starts with a jam wave similar to the example above.

Phase 1. Assume a jam wave is detected on the freeway as shown in Fig. 12-2. (How the jam wave is detected will be explained in Section 12-4.) The jam wave typically has a very low flow, low speed, and a high density. We assume that the traffic state upstream (state 6) and downstream (state 1) of the jam wave is in free flow which is generally the case in real

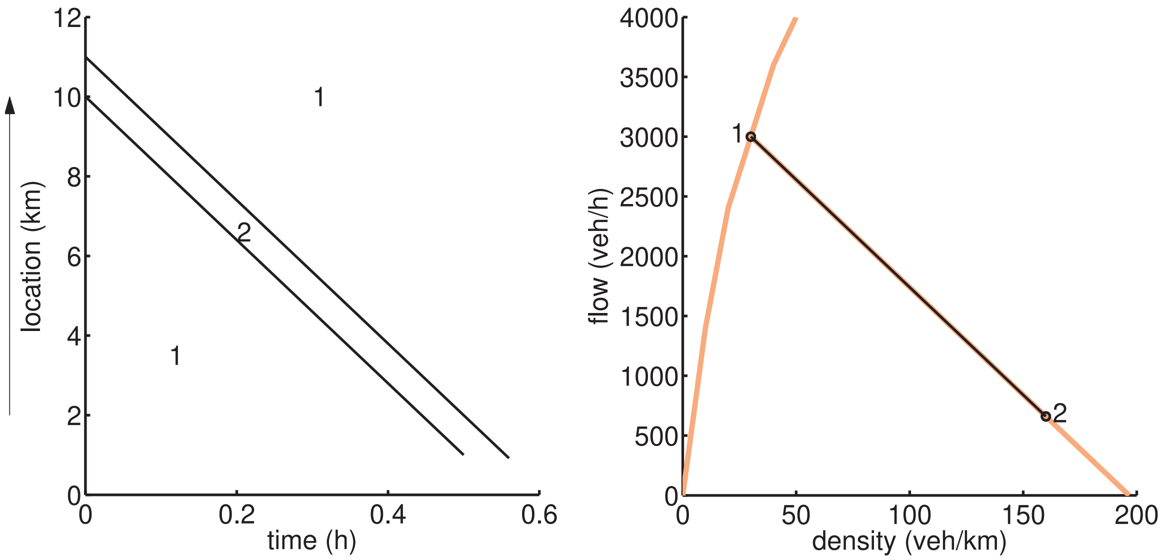


Figure 12-1: According to the shock wave theory the propagation of the front between two traffic states (left figure) has the same slope as the line connecting the two states in the density-flow diagram (right figure). The arrow indicates the travel direction.

traffic. In Fig. 12-2 phase 1 is indicated by the green background (dark gray in black and white) in the diagrams. For the sake of readability of the figures we assume that state 1 and state 6 are equal, but the theory also holds for the case when they are unequal.

Phase 2. As soon as the jam wave is detected the speed limits upstream of the jam wave are switched on. This leads to a state change in the speed controlled area from state 6 to state 3 as shown in Fig. 12-3 on the right, and to the boundary between areas 6 and 3 as shown in the figure on the left. State 3 has the same density as state 6, as the density does not change when the speed limits are lowered on a longer stretch: no vehicles can suddenly appear or disappear. However, the flow of state 3 is lower than that of state 6 due to the combination of the same density with a lower speed.

As shown by the density-flow graph, the front between states 2 and 3 will propagate backwards with a lower speed than the front betwen states 1 and 2, which resolves the jam wave after some time. The required length of the speed-limited stretch depends on the density and flow associated with state 2 and the physical length of the detected jam. We choose the length of the speed-limited stretch such that the creation of state 3 exactly resolves the jam wave.

At the upstream end of the speed-limited area traffic will flow into this area, with the speed equaling the speed limit and with a density that is in accordance with the speed, typically significantly higher than the density of state 3 (which was the density corresponding to free-flow). This state is called state 4, and the front between states 6 and 4 will propagate upstream.

Phase 3. When the jam wave (area 2) is resolved there remains an area with the speed limits active (state 4) with a moderate density (higher than in free-flow, lower than in a jam wave). A basic assumption in this theory is that the traffic from such an area can flow out more efficiently than a queue discharging from full congestion as in the jam wave. So, the

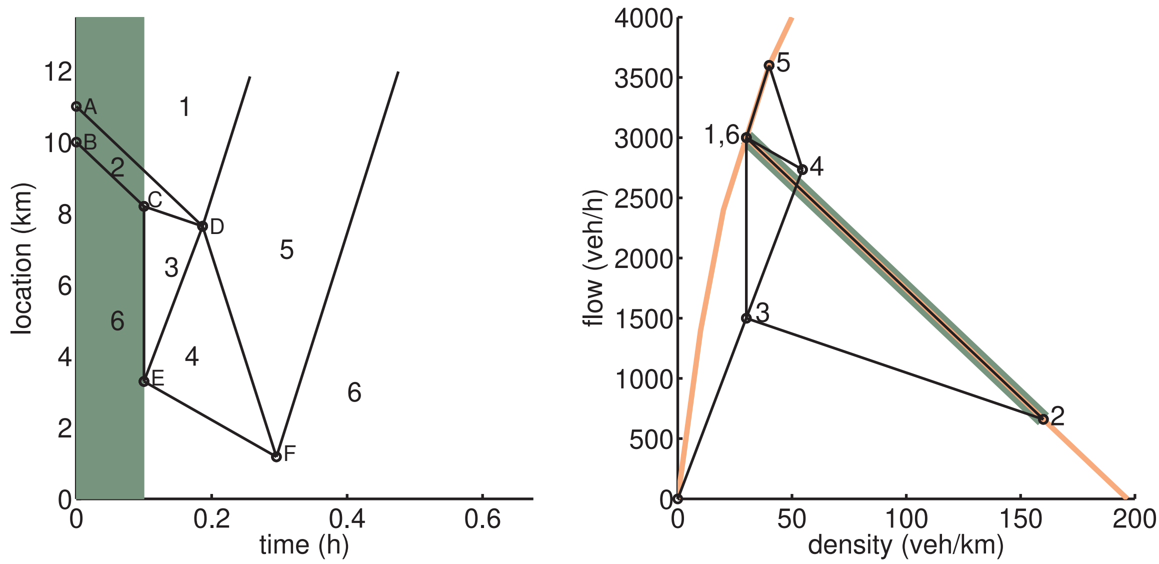


Figure 12-2: Phase 1: The jam wave is detected.

traffic leaving area 4 will have a higher flow and a higher speed than state 4, represented by state 5. This leads to a backward propagating front between states 4 and 5, which resolves state 4 as shown in Fig. 12-4.

Phase 4. What remains is state 5, and state 6 upstream and state 1 downstream of it. The fronts between states 1 and 5, and between states 6 and 5 propagate downstream, which means that eventually the backward propagating jam wave is converted into a forward propagating wave leading to a higher outflow of the link as shown in Fig. 12-5.

Obviously, not all traffic situations are suitable to construct the above control scheme. The exact requirements for such a scheme will be discussed in Section 12-4-2 when the resolvability assessment is discussed.

12-4 Algorithm development

In this section we translate the theory of Section 12-3 to a practically applicable algorithm. For this translation the following aspects are taken into account:

- The theory assumes that speed as well as flow and density measurements are available. In practice density is seldomly measured and needs to be estimated. We assume for the algorithm that only speed and flow measurements are available for each freeway section, and the density is approximated.

- The theory assumes that the measurements are available for all locations and all times. In practice the measurements are discrete in time and space. A typical spacing is every 500 or 600m and a typical sampling time is 30s, or 1, 2 or 5 minutes. The algorithm takes into account the discrete nature of the measurements.

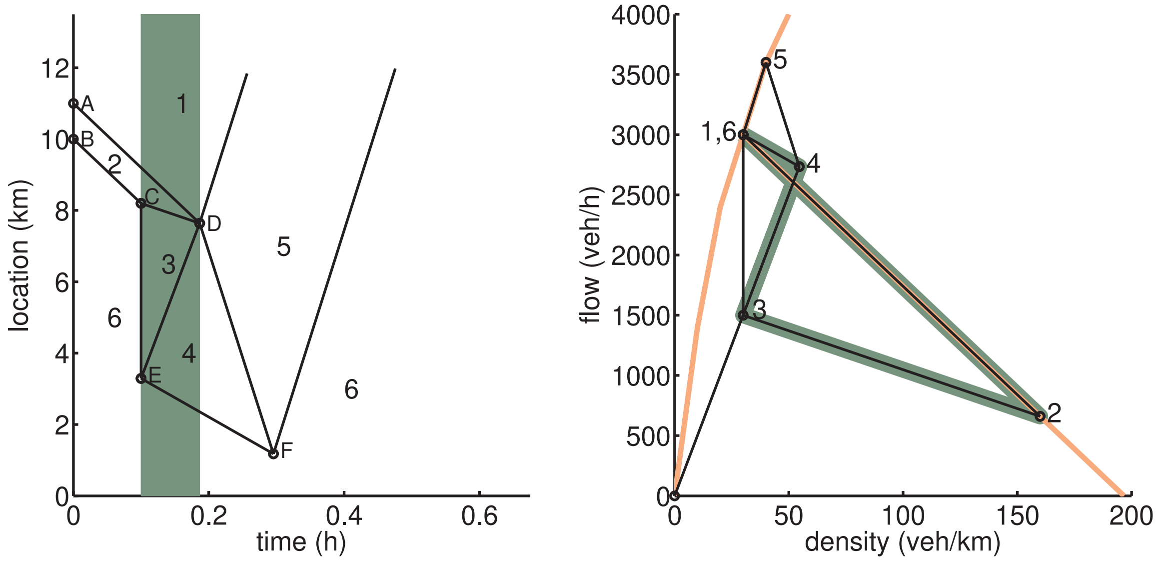


Figure 12-3: Phase 2: Speed limits are turned on. The jam wave dissolves.

- Similarly, the control actions (such as setting a speed limit) may be actuated by a discrete-time or an event-based system. The control actions may be event-based to achieve the shortest possible delay between the control decision and the appearance of the signal on the signaling device.

To cope with these issues the theory is reformulated as an algorithm. Before going into the details of the algorithm we describe the general steps of SPECIALIST:

1. **Jam wave detection.** When new flow and speed measurements arrive, test whether there is a jam wave on the considered stretch. If there is no jam wave present then wait for the next measurement and go to step 1, otherwise continue with step 2.

2. **Resolvability assessment.** Determine the solvability of the detected jam wave based on the measurements in the jam wave and upstream and downstream of it. If it is not resolvable then wait for the next measurement and go to step 1, otherwise continue with step 3.

3. **Control scheme generation.** Based on the traffic states calculate the control scheme according to the theory of Section 12-3. This scheme will apply until the jam wave and area 4 are resolved.

4. **Control scheme application.** Determine the applicable speed limits for the current moment based on the control scheme. The speed limits that are in areas 2, 3 and 4 are activated. Repeat this step regularly (e.g., every second) and ignore new measurements until no speed limits need to be set anymore according to the control scheme and thus this control scheme is finished (i.e., area 4 has been resolved). Wait for the next measurement and go to step 1.

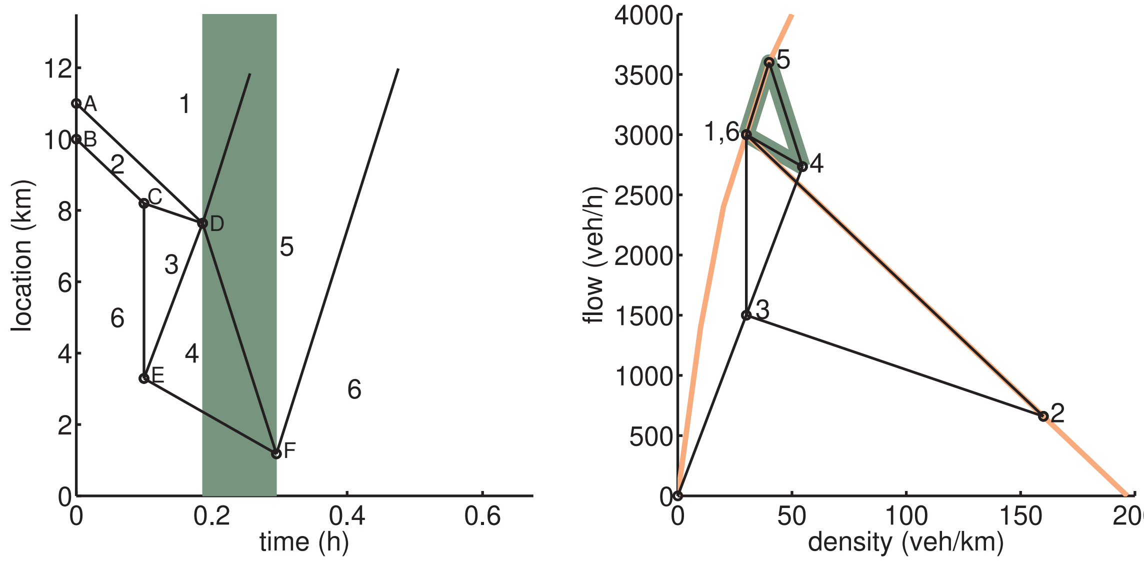


Figure 12-4: Phase 3: The speed-limited area resolves and flows out efficiently.

12-4-1 Jam wave detection

In the jam wave detection step the jam wave is detected by using thresholds for the speed and flow measurements of each segment, which determine the maximum speed $v_{\max}$ (km/h) and the maximum flow $q_{\max}$ (veh/h/lane) that the traffic in the jam wave may have. When there are no other types of jams on the considered stretch, this identifies the location of the jam wave.

12-4-2 Resolvability assessment

For the assessment of the resolvability first all the traffic states in the control scheme are determined. The areas directly upstream and downstream of the jam wave that are not falling below the thresholds are classified as free-flow, and the measurements from these areas are determined to calculate the average states upstream and downstream of the jam wave. The average flow $\bar{q}$ (veh/h/lane) and average density $\bar{\rho}$ (veh/km/lane) upstream (downstream) are approximated by $\bar{q} = (1/N)\sum_{i \in I} q_i$, and $\bar{\rho} = (1/N)\sum_{i \in I}(q_i/v_i)$, where q_i (veh/h/lane) and v_i (km/h) are respectively the flow and speed measurements of segment i, and I is the set of appropriately chosen segment indices and N the number of segments considered. This determines states 1 and 6.

As a very low speed is associated with state 2 and induction loop detectors are known to be inaccurate for low speeds, the density of state 2 is not determined using the speed measurement, but by using the density and flow of state 1, the flow of state 2, and the propagation speed of the head of the jam wave (front between states 1 and 2, which determines the slope of the line that connects states 1 and 2). The propagation speed of the head of the jam wave can be taken from off-line data, since it is well-known to be fairly constant (around -18 km/h).

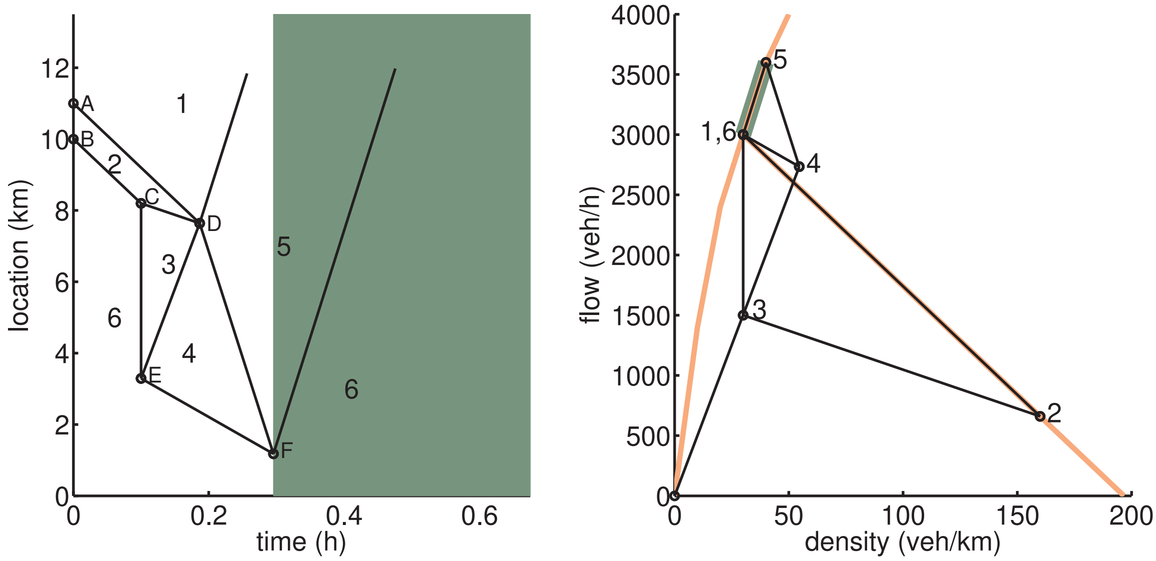


Figure 12-5: Phase 4: The remaining area 5 is a forward propagating high-speed high-flow wave.

State 3 directly follows from the density of state 6 and the used speed reduction. The speed of state 4 equals the speed of the speed limits. However, the density of state 4 does not follow from the jam wave theory and can considered as a design variable. Similarly the density and flow of state 5 is also a design variable. These variables will be discussed in Section 12-5. For now, it is sufficient to assume that they have a fixed value.

When all six states are determined, the control scheme can be constructed and the resolvability can be assessed. A jam wave is classified as resolvable if:

- the control scheme can be constructed according to the theory given the traffic states 1–6.

  The necessary and sufficient conditions for the constructions of the control scheme are:

  - the head and tail of area 2 should converge, otherwise the jam wave will not be resolved,
  - the same applies for area 4,
  - state 5 should have a higher flow and a higher density than state 1, otherwise there is no forward propagating front between these two states (this requirement will also be a guideline in the tuning of state 5),

- the speed in area 6 should be higher than the speed limit, otherwise the speed limits do not have any effect (even if the control scheme can be constructed),

- the necessary length of trajectory that is speed-limited should be smaller than length of the upstream free-flow area,

- if the speed-controlled area is limited by the physical availability of speed limit signs the controlled area following from the scheme should fall inside the physically available speed limits.

If these conditions are satisfied, the jam wave is classified as resolvable.

12-4-3 Control scheme generation

If the jam wave is resolvable, then the control scheme can be constructed. The construction involves the determination of the various fronts by resolving straightforward linear equations based on the traffic states and shock wave theory. These fronts also determine the necessary length of the speed-limited area, and the intersection of the fronts at the points where areas 2, 3, and 4 are resolved.

The feedforward control scheme only depends on the currently available measurements. Since it is the best to resolve the jam wave as early as possible (when is is as short as possible) the detected length covers usually only one or two detector locations. Consequently the estimation of the jam wave state may be inaccurate and the algorithm needs to be robust against these inaccuracies. Using these early but possibly inaccurate measurements is better than waiting longer and using more accurate state estimations, as at a later moment the jam wave may turn out to be not resolvable anymore. In Section 12-5 robustness will be discussed in the context of the tuning of the parameters of the algorithm.

12-4-4 Control scheme application

The control scheme application is straightforward: given the current time the speed limits in areas 2, 3 and 4 are activated (at their discrete locations). As the necessary speed limits change over time according to the control scheme, the locations are recalculated frequently (e.g., every second). While the control scheme is active, no new measurements are taken into account, since it is difficult to identify the areas 2, 3, 4 accurately during the control scheme. If areas 2, 3 and 4 are in the past then the control scheme is completed and the speed limits can be deactivated and the algorithm waits for new measurements.

12-5 Tuning parameters

In the algorithm there are several parameters that can be selected by the designer of the control system. Here we discuss the effects of the selection and the guidelines to select proper values.

The thresholds v_{max} and q_{max} should be chosen such that they are between the speed and flow associated with jam waves and the speed and flow associated with free-flow traffic. Empirical data (not presented here) shows that the gap between the two states is in general large for both speed and flow, and that the location of the identified jam wave is not very sensitive to the value of these thresholds. In general higher thresholds values lead to longer jam waves.

To compensate for the dependency of the head and tail location on the threshold values, and to cope with the fact that the jam wave can only be detected at discrete locations and discrete times additional margins m_{h} (km) and m_{t} (km) are defined for the estimation of the location of the head and tail. The location of the jam wave head $\hat{x}_{\mathrm{h}}$ (km) and the tail $\hat{x}_{\mathrm{t}}$ (km) that the algorithm uses to determine the control scheme are given by $\hat{x}_{\mathrm{h}} = x_{\mathrm{h}} + m_{\mathrm{h}}$, and

$\hat{x}_{\mathrm{t}} = x_{\mathrm{t}} - m_{\mathrm{t}}$, where x_{h} and x_{t} are respectively the detector locations where the head and the tail of the jam wave have been identified. So, by these margins the jam wave is assumed to be somewhat longer (typically a few hundreds of meters on both ends) than what follows from the measurements, and leads to a control scheme that is "on the safe side". In other words, larger margins lead to more robust control schemes. However, if the margins are chosen too large, the algorithm may classify all jam waves as unresolvable, or may create control schemes that limit the flow on a longer stretch than necessary.

One of the most important tuning parameters is the density associated with state 4. The speed of state 4 is determined by the speed limits, however the choice of the density is a design variable that influences the shape of the control scheme. If the density is chosen higher, the slope between states 4 and 6 will be less steep, which means that the tail of the speed limits will propagate less quickly backwards. This relationship also can be formulated the other way around: by letting the tail of the speed limits (the front between states 4 and 6) propagate faster backwards, the density in area 4 can be kept low. By selecting 4 properly (low enough), the stability of the traffic under the speed limits can be ensured. The density at which traffic becomes unstable can be identified by selecting an sufficiently low value where traffic is known to be stable and gradually increasing it until the traffic becomes unstable.

Other tuning parameters are the flow and density of state 5. The choice of state 5 determines the state that the control scheme will produce at the end, and should satisfy the following rules:

- state 5 should have a higher flow than state 1, otherwise there is no flow improvement by the control scheme,
- state 5 should have a higher density than state 1, otherwise it would have a density lower than or equal to the density of state 1, but a higher flow than state 1, which is not in accordance with the general shape of the free-flow branch of the fundamental diagram,
- state 5 should have a lower speed than state 1, for the same reason.

Obviously, the flow of state 5 cannot be unboundedly high, as it is limited by the autonomous behavior of drivers as they leave state 4.

Other considerations for the choice of state 5 may include potential bottlenecks downstream of the jam wave, which may cause a breakdown if the incoming flow is too high. In such cases the flow of state 5 can be chosen lower.

As a last remark we mention that depending on the enforcement level a certain displayed speed limit value may result in different realized traffic speeds. The effective (realized) value of the speed limit may be considered as another tuning variable, and the algorithm should use the effective speed limit in the control scheme generation, and the displayed value for the actuation signal.

12-6 Simulation

In this section we demonstrate the algorithm by simulation, where SPECIALIST is applied to a model of a freeway stretch and the results are evaluated in quantitative and qualitative

Figure 12-6: The considered freeway stretch: a part of the Dutch A12 from Bodegraven to Harmelen.

terms.

The freeway stretch that we model is a part of the Dutch A12 and has three lanes and a length of approximately 14 km going from the connection with the N11 at Bodegraven up to Harmelen, and is shown in Fig. 12-6. The stretch includes a few on-ramps, however the on-ramp volumes do typically not create jams at the on-ramps. Jam waves are often created around km 48 (in the figure behind the red A12 sign).

The stretch is equipped with double loop detectors with a typical distance of 500 to 600m, measuring the average speed and flow every minute. Above each detector there is a VMS panel that displays the speed limit.

The stretch is simulated with METANET model (Messmer and Papageorgiou, 2001) including the extensions proposed in (Hegyi et al., 2005), which is a second-order macroscopic traffic flow model with some extensions to reproduce shock waves better. The model was calibrated with real traffic data, and several traffic scenarios were considered, which were also taken from real traffic data.

Due to spacial limitations we present here only one scenario which contains one jam wave that enters the freeway and travels upstream without being resolved. The evolution of traffic in the uncontrolled case is shown in Fig. 12-7.

SPECIALIST was applied to this scenario with using the following parameters: displayed and effective speed limit of 60 km/h (assuming full compliance), density of state 4 of 33 veh/km/lane, density and flow of state 5 of 26 veh/km/lane and 6100 veh/h, jam wave propagation speed of -15 km/h, m_h = 0 km, m_t = -1.5 km, and the threshold parameters $q_{\max}$ = 1500 veh/h/lane, $v_{\max}$ = 50 km/h. The parameters were chosen after only a few

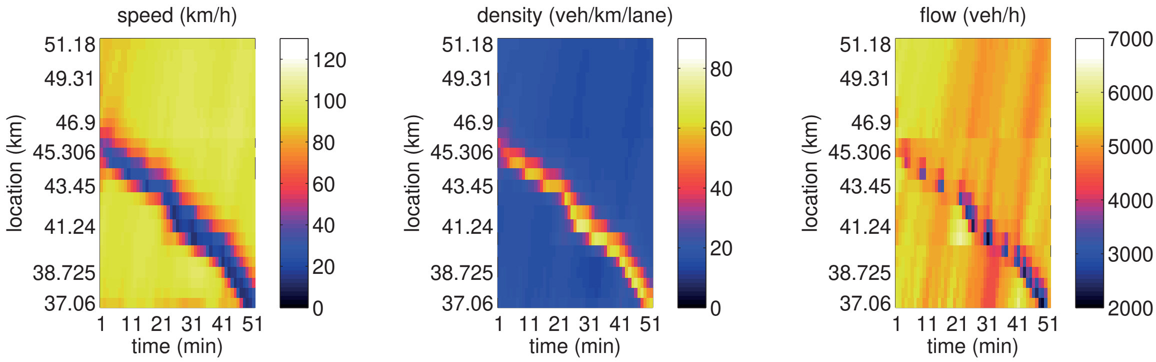


Figure 12-7: The time-space diagrams of the speed, density and flow for the uncontrolled case. The jam wave remains existent and propagates upstream.

Table 12-1: Comparison of the performance of the uncontrolled and the controlled cases.

	DVH (veh.h)	TTS (veh.h)	average flow (veh/h)
Uncontrolled	2389	853	6461
Controlled (improvement)	2471 (82)	775 (9.1%)	6738 (4.2%)

iterations of tuning.

The results of the control system are shown in Fig. 12-8 where the high density around km 45.306 is the beginning of the jam wave. The control scheme is indicated by the white lines, and the different area's can be clearly distinguished. When the jam wave is detected and classified as resolvable, the control system reduces the speed limits, creating area 3, which has a low density and low flow. This resolves the jam, and at the same time creates area 4, which has the same speed but a higher density than area 3. When area 4 resolves into area 5 it leads to a high flow, which nearly triggers a new jam at the on-ramp at 40.046 km. Comparing the outflows for the uncontrolled and the controlled case, the difference can be clearly seen. Note that no new jams have been created under or upstream of the speed limits, which suggests stability.

The uncontrolled and controlled cases have also been compared by numerical performance measures as shown in Table 12-1. The performance measure DVH (downstream vehicle hours) specifies the total time that vehicles spend downstream of the considered stretch. The higher this value, the better: the value is higher if vehicles leave the link earlier or with a higher flow. TTS is the total time that vehicles have spent in the stretch during the simulation period, and third performance measure was the average outflow of the stretch including the off-ramps. All three performance measures show that there is a significant improvement due to the speed limit control. The improvement of the TTS for the other scenarios was in the range of 10-19% and is of the same order as known from other publications (Popov et al., 2008).

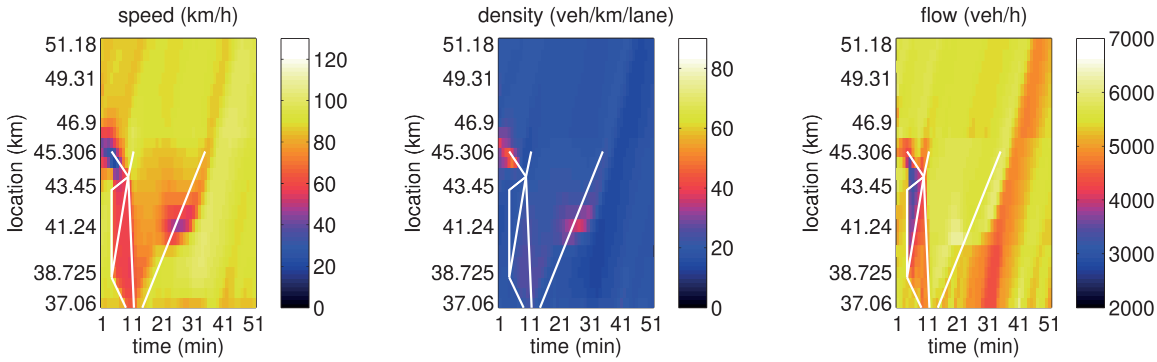


Figure 12-8: The time-space diagrams of the speed, density and flow for the controlled case. The jam wave is resolved after a few minutes.

12-7 Real-world evaluation approach

Both traffic performance and the operation of the algorithm is also evaluated in real traffic. We discuss here the evaluation questions and the approach for answering these questions. In the approach the visual assessment of the traffic scenario in the space-time plots of the speed, flow and density plays an important role. Although this method is not rigorous, the information in these plots was invaluable for the tuning and evaluation of the algorithm. Currently there are no systematic methods available that can reliably distinguish between several types of jams, such as jam waves (moving jams), jams at on-ramps, jams caused by accidents, etc.

12-7-1 Traffic performance evaluation

For the evaluation of the traffic performance the following performance indexes were considered:

- **The number of resolved jams.** For each speed limit intervention the targeted jam was evaluated whether is was resolved by the speed limits. Resolution can be identified by disappearance of the high density and very low flow that is associated with jam waves. An example of a successful resolution is given in Fig. 12-9 and of an unsuccessful resolution in Fig. 12-10.

- **The gain in veh-hours.** The gain in time that the vehicles spend in the considered section is calculated for each intervention in the following way. First, the basis of comparison is the assumption that in case of no intervention the jam would have lasted for an hour, and would have had an outflow of 1500 veh/h/lane. These assumptions were taken from other studies that were performed previously on this freeway stretch. Next, the flow for each intervention is measured at 2 km downstream of the line AD for the duration of the control scheme (from A to F). The distance of 2 km was chosen such that the vehicles leaving the jam are not accelerating anymore (during acceleration the full flow is not reached) but not too far downstream to prevent measuring flows corresponding to other jams. From the difference in flows of the assumed uncontrolled scenario and the measured flows the gain in vehicle-hours was determined.

- **New jams triggered.** Each intervention was judged whether new jams were created in or upstream of area 4 (which may be due wrong tuning of the density 4) and downstream (which may be due to the tuning of state 5). An example where three upstream jams were created is given in Fig 12-11.
- **Compliance.** The compliance of the traffic with the shown 60 km/h is determined by investigating the average and the standard deviation of the speed in areas 3 and 4 for each intervention.

12-7-2 Evaluation of the algorithm

The algorithm has been evaluated whether the general operation is as expected, and whether the scheme has the expected shape. Furthermore, by visual inspection of the space-time plots the type of jam was determined for which the speed limits became active. These types were classified as (1) jam wave, or (2) other type, where the other types of jams include jams at on-ramps (recognizable in the space-time plots by the fixed head at an on-ramp location), fixed jams at a bottlenecks (fixed head, but no on-ramp), jams due to accidents (non-structural location combined with very low flow), a few cases of jams caused by snowplows (forward moving bottleneck combined low flow), and jams of unclear type.

12-8 Experiment set-up

The freeway stretch for the test is a part of the Dutch A12 freeway and has three lanes and a length of approximately 14 km. The stretch is located between the connection with the N11 at Bodegraven up to Harmelen as shown in Fig. 12-6. The stretch includes a three on-ramps: at Bodegraven, at Nieuwerbrug, and at Woerden. From offline data analysis it was known that the jam waves are often created around km 48 (in the figure behind the green E25 sign).

The stretch is equipped variable message signs that displays the speed limit. The spacing of the gantries is 500 to 600m. Under each gantry there are double loop detectors (one pair for each lane), measuring the speed and flow. The speeds and flows were averaged over a period of one minute, but with a sampling time of 10s. At some locations there were additional loop detectors which were also used for the traffic measurements.

The displayed speed limit for this algorithm was 60 km/h, and for the lead-in 80 and 100 km/h were used.

12-9 Results

In this section we discuss the performance of the SPECIALIST algorithm and the tuning steps taken during the test. Since the tuning steps depended on the performance results they will be discussed together in Section 12-9-1. In Section 12-9-2 we discuss the findings regarding the operation of the algorithm.

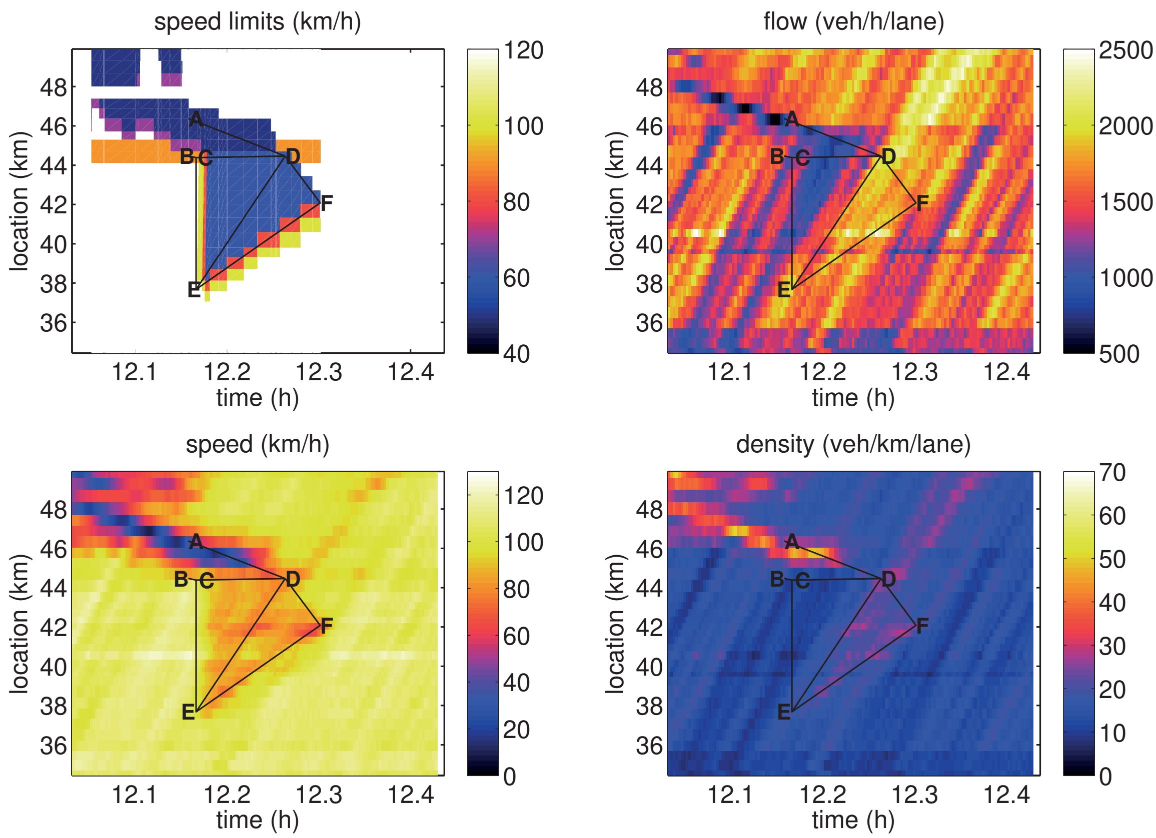


Figure 12-9: An example of a properly resolved jam wave (2010-02-26)

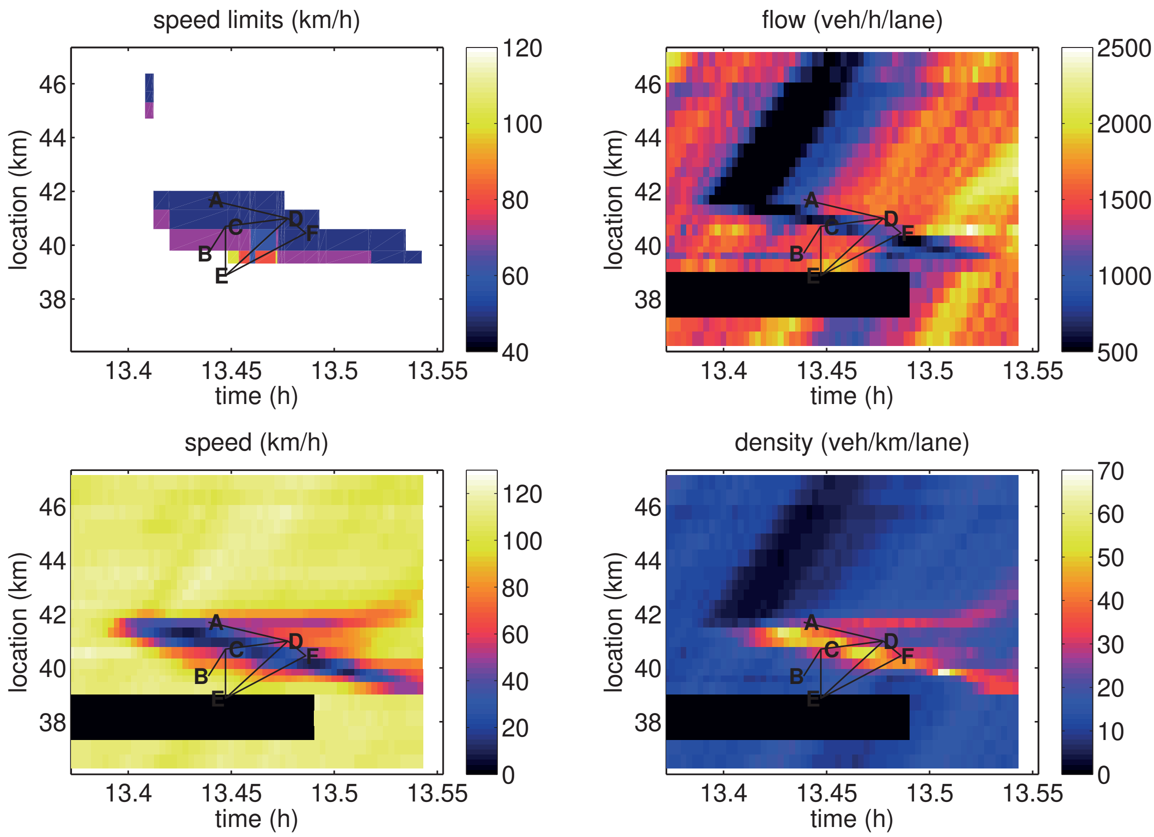


Figure 12-10: An example of a jam that was not resolved (2010-02-17). The black areas indicate missing data.

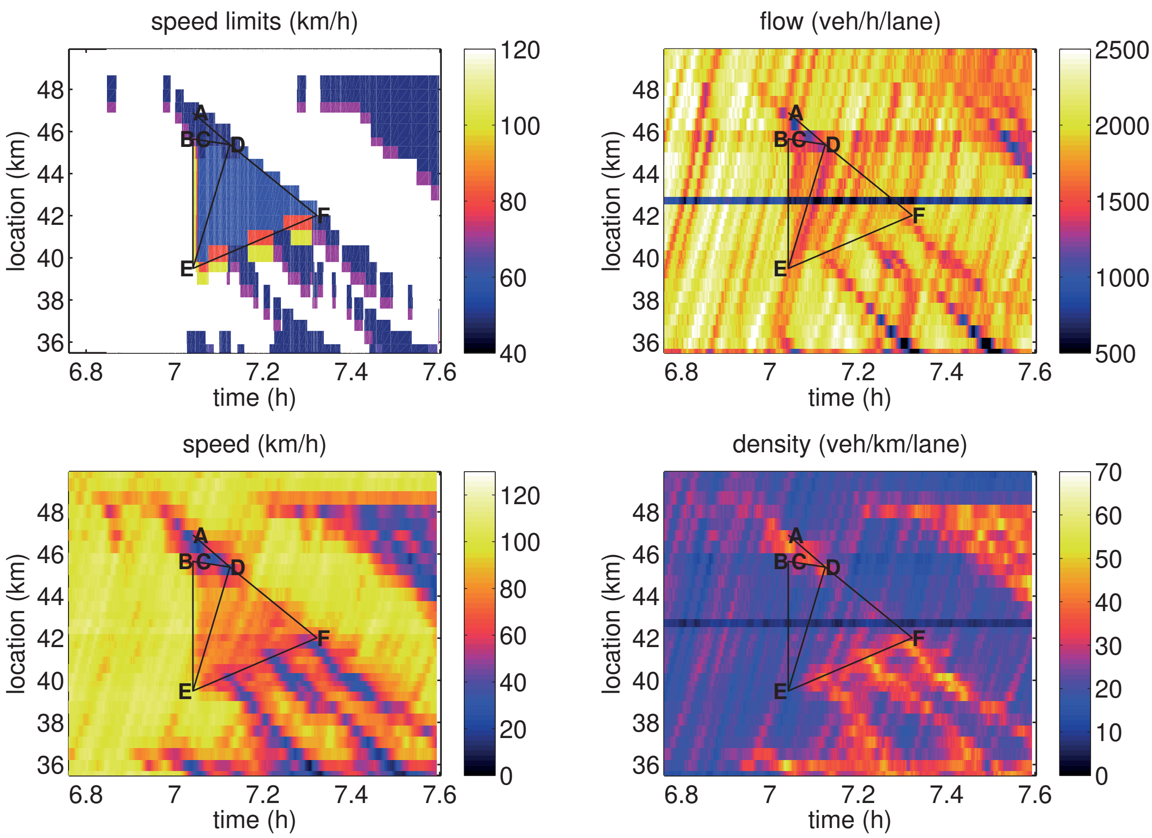


Figure 12-11: An example of new breakdowns due to a too high $\rho_{[4]}$ (2009-10-05)

12-9-1 Traffic performance and tuning

The initial tuning parameters were selected based on experience with the offline data and the simulation studies, and are shown in Table 12-2. Table 12-3 shows the performance indexes discussed in Section 12-7. Since the algorithm also activated for other types of jams than jam waves only, the table distinguishes the results according to jam type.

From Table 12-3 it is clear that in the period of the third parameter setting the performance was significantly lower than in the other periods. This is probably due to the extremely bad weather in this period, where there was a lot of snowfall. In this period often one or two lanes were closed and snowplows were used on a regular basis.

The algorithm activated in roughly the half of the cases for jam waves, and the other half for other jam types. For the jam waves the average gain in total travel time was 35 veh-h, while for the other jams it was close to 0 veh-h.

During the test the density $\rho_{[4]}$ was lowered gradually, in order to stabilize traffic and reduce the number of new jams appearing upstream of the speed-limited area. This led to a reduction from 0.33 to 0.07 new upstream jams.

The compliance during the whole test was roughly constant and was somewhat better (lower) in area 4 than in area 3. This might be due to the higher density in area 4, which may motivate drivers better to drive slower.

12-9-2 Algorithmic results

- The algorithm activated frequently for other jams than jam waves. This is mainly due to the simple way that jam waves are detected and the desire to detect jam waves as early as possible. Early detection of jam waves entails that there is not time to track the head of the jam and see if it is backwards propagating (which is the distinguishing feature of jam waves).

- In a relatively large number of cases the jam waves were classified as not resolvable due to the resolvability condition 12-4-2 in Section 12-4-2. The reason was that (the measured) $v_{[1]}$ was lower than $v_{[5]}$ or that the $\rho_{[1]}$ was higher than $\rho_{[5]}$. These cases were due to the random fluctuations in the downstream state. Since for a given state 4, there are several solutions for state 5 that result in the same slope 4-5 (all the solutions are on the same line), another setting was chosen for state 5, which had a lower speed and a higher density (but with the same slope 4-5) in order to minimize the number of jam waves that are assessed as not resolvable. The effect of this change on the number of activations per day can be seen in Table 12-3.

12-10 Conclusions

A speed limit control algorithm called SPECIALIST was presented that can resolve jam waves on freeways. The approach is based on theoretical considerations in terms of the shock wave

Table 12-2: The parameter settings for the previously performed data analysis, micro and macro simulations, and the four settings in the real-world test. The values that were changed during the test are printed in bold.

	Par. 1	Par. 2	Par. 3	Par. 4
start date (dd-mm-yy)	08-09-09	29-10-09	15-12-09	03-02-10
end date (dd-mm-yy)	29-10-09	15-12-09	03-02-10	28-02-10
v_{max} (km/h)	50	50	50	50
q_{max} (veh/h/lane)	1500	1500	1500	1500
v_{front} (km/h)	-18.1	-18.1	-18.1	-18.1
v_{eff} (km/h)	70	70	70	70
$\rho_{[4]}$ (veh/km/lane)	28	**27**	**26**	26
$v_{[5]}$ (km/h)	93	93	93	**81**
$q_{[5]}$ (veh/h/lane)	2060	2060	2060	**1945**
$x_{\text{head-offset}}$ (km)	0	0	0	0
$x_{\text{tail-offset}}$ (km)	-1,25	-1,25	-1,25	-1,25

Table 12-3: Statistics of the SPECIALIST activations.

		parameter setting				
		1	2	3	4	total
# days		49	42	35	23	149
# activ. total		84	60	31	67	242
av. activ. per day		1,7	1,4	0,9	2,9	1,6
resolved %		68%	60%	32%	58%	59%
av. new jams upstr. per activ.		0,33	0,08	0,13	0,07	0,17
av. new jams downstr. per activ.		0,01	0,03	0	0,15	0,05
av. gain veh-h per activ.		16	16	7	26	18
jam waves	# activ.	48	24	11	32	115
	resolved	83%	83%	55%	72%	77%
	av. gain veh-h per activ.	30	41	33	39	35
other jams	# activ.	36	36	20	35	127
	resolved	47%	44%	20%	46%	41%
	av. gain veh-h per activ.	-2	-1	-8	14	2
compliance area 3 [std] (km/h)		82[15]	83[20]	79[21]	84[13]	83[16]
compliance area 4 [std] (km/h)		74[15]	74[18]	75[19]	76[12]	75[15]

theory. These considerations were translated into a control approach that is suitable for on-line real-life application. The approach results in a speed limit control plan that is applied as a feedforward control signal to the traffic process. The algorithm only includes a few parameters which all have physical interpretation, which makes the tuning of the algorithm intuitive and insightful.

The SPECIALIST algorithm was demonstrated by a simulation example of a stretch on the the Dutch freeway A12. The simulation results show that the algorithm is capable of resolving jam waves on freeways.

Selected problems

For this chapter, consider problems: A-13-6

Chapter 13

Local intersection control – fundamentals

The objective of traffic control is to guarantee a safe and efficient use of “conflict areas”, areas which vehicles, approaching from different directions, want to use at the same time instant. Although traffic can be controlled by both fixed and variable signs, the terms “controlled intersection” and “traffic control” will be used for the application of traffic signals giving cyclic right of way. During a cycle, each conflicting movement gets a green period, this is permission to use the conflict area. Because of the use of red, green and yellow signals, this is also called signal group.

To distinguish the various traffic movements, they are coded according to the Dutch standardisation. These codes are also used to indicate the default conflicts between pairs of movements in the conflict matrix.

13-1 Standard codes

In the Netherlands standard codes of traffic signals are used to identify traffic movements and their standard conflicts.The codes used are, see figure 13-1:
01 – 12: for motor cars
21 – 28: for bicycles
31 – 38: pedestrians
41 – 52: trams, buses and emergency vehicles (and sometimes taxis)
61 – 72: following movements
81 – 99: reserve

Sometimes one traffic signal controls a combination of movements from one approach. The traffic signal is called the according to the through going movement, see figure 13-2

If a pedestrian or bike movement is not divided by an intermediate the total crossing is indicated by the even number, as in figure 13-3.

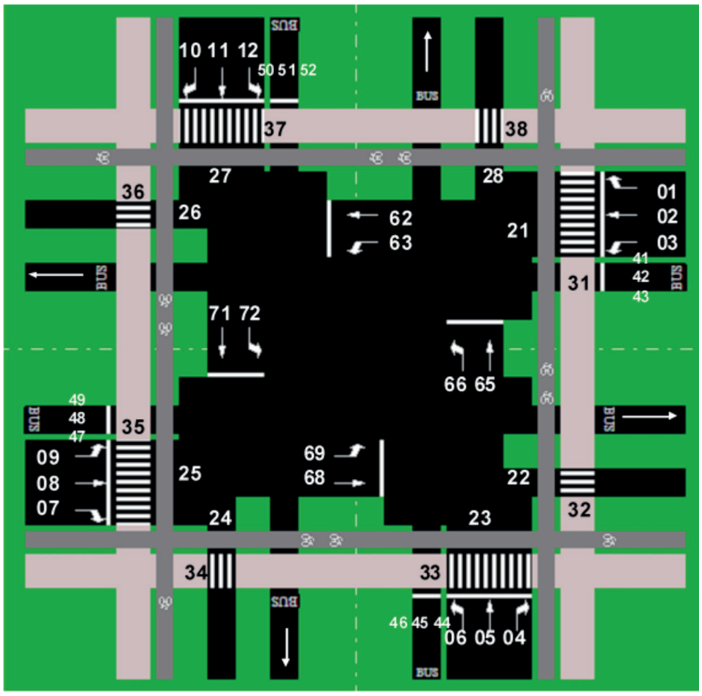


Figure 13-1: Standard code of traffic movements at an intersection

Through going and right turning movement

Through going, left and right turning movement

Through going and left turning movement

Figure 13-2: Combined movements

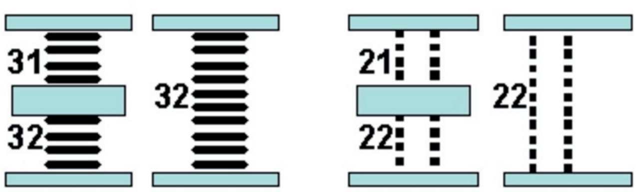


Figure 13-3: Connected movements

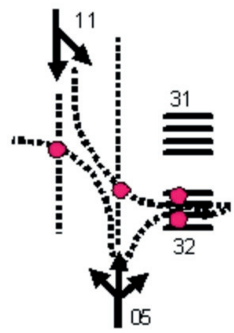


Figure 13-4: Permitted conflicts

13-2 Conflicts

There are two types of conflicts:

1. ***protected conflict***

   A protected conflict prohibits two movements having green and/or amber (yellow) at the same time. Every traffic signal that controls a single movement, and every pair of movements using a common part of the intersection (conflict area) is considered conflicting by default. Such conflicts are protected.

2. ***permitted conflict***

   Permitted conflicts occur at signals that are controlling combined movements that, according to the standard codes, would not conflict. The movement is coded according to parallel directions (e.g. stream 05 and 11). However, the movements are a combination of through going movements and turning movements that do have conflict (e.g. left turning traffic from 05 and through going traffic from 11, see figure 13-4). Since the complete movement is partly parallel and partly conflicting, a permitted conflict is sometimes known as "partial conflict". By default, parallel signals are not conflicting, but permitted conflicts are only permitted if there is sufficient space on the intersection to allow for handling the conflict according to the priority rules. If this space is lacking, or it is not safe to permit the conflict, the signals should be considered conflicting. In this case, for given example, signal 05 and 11 will conflict.

   The permitted conflicts found in figure 13-4:

   (a) Left turn movement of signal 05 and through going movement of signal 11

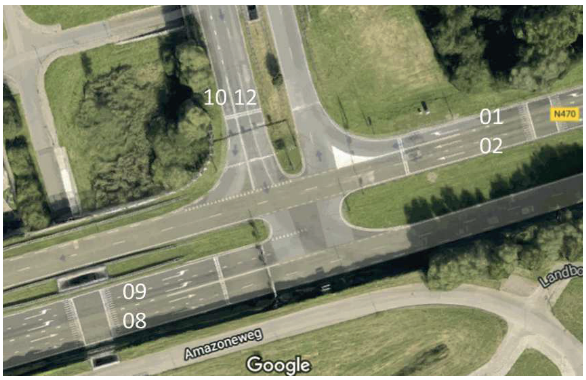
   (b) Left turn movement of signal 11 and through going movement of signal 05

   (c) Right turn movement of signal 05 and pedestrian signal 32

   (d) Left turn movement of signal 11 and pedestrian signal 31

13-3 Conflict matrix and clearance times

The conflict matrix shows the conflicting and non-conflicting pairs of signals at an intersection. As an example, figure 13-5 shows an intersection, the Kruithuisweg - Pr. Beatrixlaan intersection, with its conflict matrix. The conflicting pairs are indicated by "o" and non-conflicting movements by "-". The main axis is grey, since a signal never can have conflict with itself.



	01	02	08	09	10	12
01		-	-	o	-	-
02	-		-	o	o	o
08	-	-		-	-	o
09	o	o	-		-	o
10	-	o	-	-		-
12	-	o	o	o	-	

Figure 13-5: Intersection Kruithuisweg - Pr.Beatrixlaan, Google maps (left) and conflict matrix (right)

The "o" can be exchanged by clearance times (all red times). As the clearance times depend on the distances for entering and exiting the conflict area, the conflict matrix describes mathematically the layout of the intersection.

13-3-1 Clearance times

The purpose of the red clearance interval (between the start of red for one traffic signal and the start of green for the succeeding conflicting signal) is to ensure that traffic in the second movement can safely enter the intersection without colliding with the last vehicle from the first movement. While clearance times are important for safety, they also impact traffic operations: they contribute to lost time and affect delays, queuing, and necessary cycle length. In principle clearance times should be as small as possible while still allowing a traffic movement to safely follow a conflcting movement.

The Dutch practice of movement-based vehicle actuated traffic signal control, requires that before a traffic signal can get a green indication, it must satisfy clearance times for every proceeding conflicting traffic signal. Red clearance times $T_{clear}\,(i, j)$ must be supplied to the controller for each ordered pair of conflicting movements (i,j), where i is the index of the "exiting movement" and j the index of the "entering movement". For a given ordered pair of conflicting movements, clearance time is based on the travel time of vehicles in the exiting and entering movement to that movement pair's "conflict zone"; the area within the intersection where paths taken by vehicles in the two movements overlap, see figure 13-6 (Other examples are given in CROW (2014), this Handbook is in Dutch, but the figures are instructive).

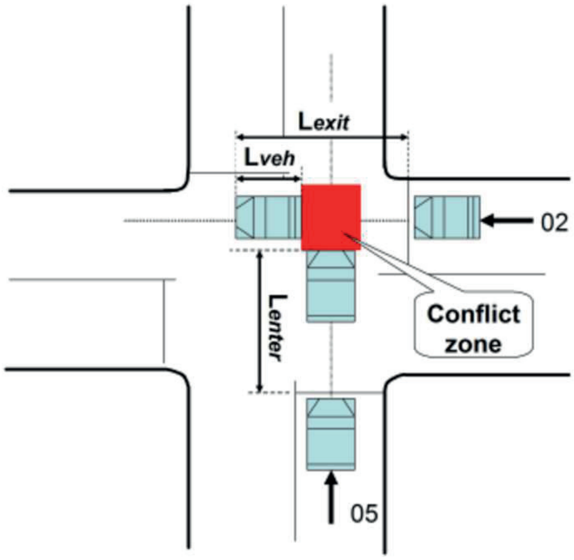


Figure 13-6: Clearance Distances

Let L_{exit} equal the distance a vehicle in the exiting movement must travel from the stop line to fully clear the conflict zone (including the vehicle length, commonly taken to be 12 m so as to represent a truck), and let T_{exit} equal an exiting vehicle's travel time from the stop line to just beyond the conflict zone; similarly, let L_{enter} equal the distance from the stop line of the entering movement to the conflict zone, and let T_{enter} equal the amount of time the first vehicle from the entering movement needs to reach the conflict zone.

To avoid a collision, the length of the red clearance interval, $T_{clear}\,(i,j)$ must be:

$$T_{clear}\,(i,j) = T_{exit}\,(i) - T_{enter}\,(j)\,|T_{clear}\,(i,j) \geq 0 \tag{13-1}$$

According to the Dutch traffic regulations, the clearance time should be equal or larger than zero.

Figure 13-6 illustrates the importance of determining clearance time as a function of an ordered pair of movements. Consider the conflict zone between conflicting movements 02 and 05. As one can see, 02's stop line is much closer to the conflict zone than is 05's. Consequently, if 02 is exiting while 05 is entering, little or no red clearance will be needed, because the last 02 vehicle has a considerable amount of time to clear the conflict zone before the first 05 vehicle arrives. On the other hand, if 05 runs first followed by 02, a considerable red clearance time will be needed, because the first 02 vehicle could arrive very quickly at the conflict zone, well before the last 05 vehicle had cleared if it entered the intersection just at the end of yellow. Stream 02 should be delayed by a red clearance interval. In the interest of safety, the exit time should concern a relatively slow vehicle, while the entrance time should pertain to a fast vehicle.

The exit time calculation is rather straightforward:

$$T_{exit} = L_{exit}/v_{exit} \tag{13-2}$$

where v_{exit} is the exit speed.

The exit speed is assumed the speed of a vehicle in the exit movement that crosses the stop line at the last moment of yellow. Because such a vehicle was presumably unable to stop during the yellow interval, its speed is unlikely to be below the average approach speed; nevertheless, a somewhat conservative value may still be used (say 40 km/h). In case of very short waiting queues and a high percentage of trucks a lower speed can be considered.

In comparison to the calculation of exit times, the calculation of entrance times is more complex. For vehicles that decelerate to a standstill at the stop line before the light turns green, entrance time can be determined easily enough based on an assumed acceleration trajectory. However, consider a vehicle approaching the stop line that has started to decelerate because the signal is red. Before it comes to a standstill, the signal turns green. That vehicle can then begin to accelerate and enter the intersection at some speed. Such a vehicle may well reach the conflict zone sooner than it would have if it had been standing at the stop line when the signal turned green. Depending on the position and speed of the entering vehicle when the traffic signal turns green, the entrance time could differ. For safety, clearance time should be based on the smallest possible entrance time.

To determine clearance time according to the complex situation described above, a model of driver behaviour is required. The following model is used:

- A driver (with no traffic ahead) approaches the intersection with an approach speed v_{app}
- Seeing the red signal, drivers decelerate as late as possible with a constant deceleration rate adec following a trajectory that, if uninterrupted, brings them to a standstill at the stop line
- When the signal turns to green, drivers accelerate with the constant acceleration aacc until they reach the maximum speed v_{max}
- A reaction time between when the light turns green and acceleration begins can be taken into account.

To obtain safe values for the clearance time, the parameters for the described model should represent a rather aggressive driver.

The indicative vehicle that determines the enter time, is the vehicle that was decelerating during red, but can start acceleration at the onset of green without having to stop. The entrance time depends on the distance to the stop line of the indicative vehicle d_s, and its acceleration distance, d_{acc}. After covering the acceleration distance, the speed of the vehicles is assumed to be constant, v_c. The acceleration distance is:

$$d_{acc} = \frac{v_c^2}{2(a_{acc} + a_{dec})} \tag{13-3}$$

where a_{acc} is the acceleration, and a_{dec} is the deceleration (a driver decelerates as long the signal is red, but starts acceleration if green is observed; the deceleration a_{dec} has a negative value.). If the distance to the stop line, d_s, is smaller than the acceleration distance, the enter time for vehicles entering the conflict area is:

$$T_{enter} = \sqrt{\frac{2d_s}{a_{acc} + a_{dec}}} \tag{13-4}$$

If applies $d_s > d_{acc}$, then the enter time is:

$$T_{enter} = \frac{d_s}{v_c} + \frac{v_c}{2(a_{acc} + a_{dec})} \tag{13-5}$$

In most cases applies the distance to the stop line is smaller than the acceleration distance, so equation 13-4 can be used. The analytical derivation of equation 13-4 is given in Muller et al. (2004).

13-4 Control structures

A control structure shows the order of realisation of green phases of the signals (permission to access the intersection).

As an example, figure 13-7 shows the control schemes for the intersection of the Kruithuisweg and the Prinses Beatrixlaan in Delft.

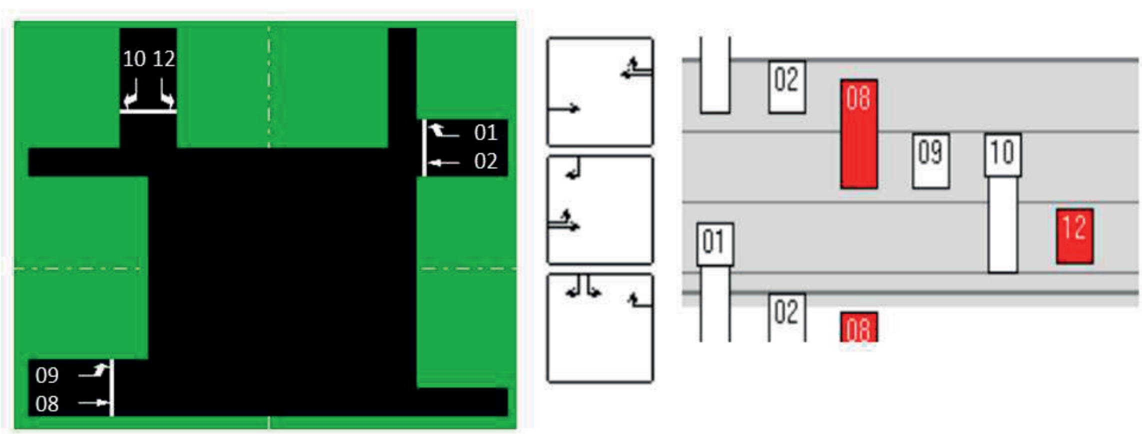


Figure 13-7: Schematic layout of the Kruithuisweg intersection (left), flow diagram (middle) and control structure (right) In the flow diagram and the control structure, the time is depicted top to bottom and movements that determine the cycle time are indicated red (08 and 12).

The control structure is derived from its conflict matrix, by determining the mutual conflicts. With these mutual conflicts the so-called "conflict groups" can be formed, this is the aggregation of mutual conflicting traffic movements. Conflict groups consists of movements that cannot have green at the same moment, because these movements want to use the same conflict area. In the control scheme these conflicts will occupy different "stages", also called "blocks", or "green phase combinations". Within a stage movements can combined that do not have conflict. In the flow diagram of 13-7 three stages are visible. In the first stage signals 01, 02 and 08 are combined, in the second stage signals 08, 09 and 10 are combined, in the

Table 13-1: A control structure of the Kruithuisweg intersection

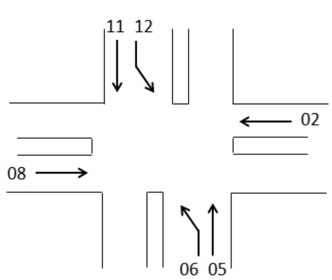
time	cg 1	cg 2	cg 3	cg 4
stage 1	02	01	02	08
stage 2	09	09	10	08
stage 3	12	01	02	12
stage 1	02	01	02	08

third stage signals 10, 12 and 01. As can be seen from this control scheme, movements can appear in multiple stages, signal 01 (stage 1 and 3), 08 in stage 1 and 2 and 10 (stage 2 and 3).

From the conflict matrix can be derived that signals 01-09 conflict, signals 02-09, 02-10, 02-12 conflict, signals 08-12 conflict and signals 09-12 conflict. It also appears that signal 02 has conflict with both 09 and 12, but also signals 09 and 12 have conflict. The "aggregation of mutual conflicts" will therefore lead to conflict group 02-09-12, which is for this particular control structure the maximum conflict group. All conflict groups are combined to make the control structure, with as starting point the maximum conflict group. The control structure is given in table 13-1. After the complete cycle of three stages, the next cycle starts with stage 1.

The sequence of the movements of the first conflict group can be changed. Instead of the order 02-09-12-02, the order can also be 02-12-09-02, these are different control structures. The order 09-12-02-09 is basically the same as 02-09-12-02, since the order of the stages is the same. This can be seen if multiple cycles are observed: 02-09-12-02-09-12-02-09-12-02-09-12-02-: starting at the first or second stage is of minor importance. So if the maximum conflict group size is n, at least $(n-1)!$ possible structures can be generated.

For the Kruithuisweg example, once the order of the movements in the first conflict group is determined, the control structure is determined, due to its conflicts. Given these conflicts, the number of structures that can be determined is $(2-1)! = 2$. Often more than $(n-1)!$ structures can be generated, this depends on the conflict groups. If there are multiple options to add a conflict group, multiple structures can be generated. Take for instance the intersection and conflict matrix of figure 13-8.



	02	05	06	08	11	12
02		o	o	-	o	o
05	o		-	o	-	o
06	o	-		o	o	-
08	-	o	o		o	o
11	o	-	o	o		-
12	o	o	-	o	-	

Figure 13-8: Example of an intersection with six car movements

The conflict groups that can be found given the conflict matrix, are: 02-05-12, 02-11-06, 08-05-12 and 08-11-06, so there are four conflict groups with all a length of three signals. Now all possible combinations of conflict groups have to be determined to find all possible control structures. The order of the first conflict group can be either 02-05-12-02 or 02-12-05-02, but movement 05 can be combined with either 06 or 11. So also the order of the second conflict group can be changed, while the order of the first conflict group remains the same. So, this leads to 2x2=4 possible control structures, as can be seen from table 13-2.

Table 13-2: Resulting control structures

conflict group	1	2	3	4
stage 1	02	02	08	08
stage 2	05	11	05	11
stage 3	12	06	12	06
stage 1	02	02	08	08

conflict group	1	2	3	4
stage 1	02	02	08	08
stage 2	05	06	05	06
stage 3	12	11	12	11
stage 1	02	02	08	08

conflict group	1	2	3	4
stage 1	02	02	08	08
stage 2	12	06	12	06
stage 3	05	11	05	11
stage 1	02	02	08	08

conflict group	1	2	3	4
stage 1	02	02	08	08
stage 2	12	11	12	11
stage 3	05	06	05	06
stage 1	02	02	08	08

13-5 Control programs

Control programs control the transition from one stage to another (stage control) or from one individual green phase to another (phase control). Also the duration of the stages or green phases is controlled by the program.

Stage control is commonly used in combination with ***fixed time control*** also called ***pre-timed control***. The sequence of the stages is predetermined and each stage has a fixed duration. The cycle time is fixed and calculated based on handling the average traffic volumes in the control scheme.

During the day, the cycle time and the sequences can be changed by using another pre-timed control program. This type of control is sometimes called ***traffic dependent, pre-timed control.***

Another alternative is ***semi-actuated control***. The minor approaches have detectors (inductive loops, infra red, push buttons, etc.). Unless a detector on the minor road is activated, the main movement remains green (wait in green). The main movement green phase is terminated only after a guaranteed minimum green time has passed. The green period on the minor road lasts until a gap between vehicles longer than a certain value (say 3.5 s) has been measured or until the maximum duration of green is reached.

Phase control focuses on the realisation and termination of the green, yellow and red phase of individual traffic movements. The realisation and duration of green phases depends on the actual presence of vehicles, pedestrians, and bikes in relation to the actual traffic of the conflicting movements. Phase control is commonly used in combination with ***vehicle***

actuated control also called ***fully actuated control***. Fully actuated control is the mode of operation where all approaches have detectors and all green phases are controlled by means of detector information.

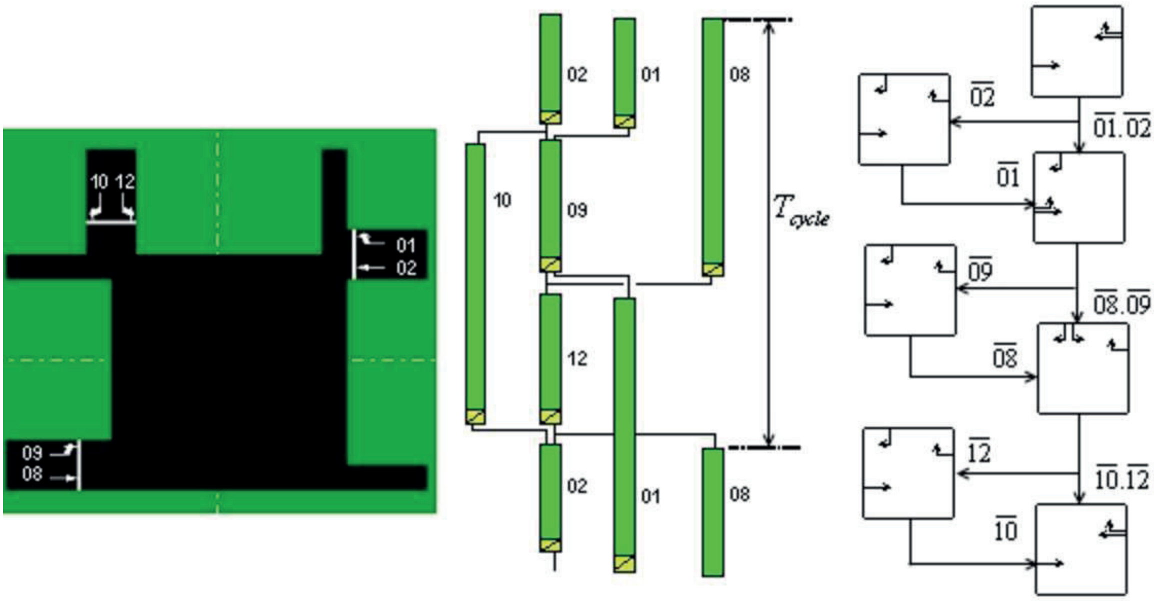


Figure 13-9: Phase control based on a structure diagram and flow diagram

In figure 13-9 phase control based on a structure diagram and a flow diagram is depicted.The green bars in the structure scheme represent the green periods during the cycle. The yellow blocks (with diagonal) show the yellow periods. A black connector represents the all-red period enabling the last vehicle to clear the conflict area before the first vehicle of the succeeding conflicting movement enters the conflict area.

If no traffic is present, the green phase can be skipped or terminated. Phase control provides more flexibility. As phase control controls the individual transitions from one green phase to another one, and each pair of traffic movements has its own clearance times, the internal lost times (where no vehicle can use the conflict area) can be reduced considerably.

Phase control results into an alternative flow diagram where the actual path can vary per cycle.

13-6 Capacity

The capacity of an intersection is determined by the shared use of conflict areas by conflicting traffic movements. The capacity is reached if one of the conflicting movements does not get enough opportunities to handle its traffic. So, the capacity depends on the ratio of flows and the time assigned to the various conflicting traffic movements.

The maximum throughput of a controlled intersection depends on the cycle time, the green splits, the yellow times, the all red times and the flow ratios in the critical path through the control scheme.

13-6-1 PCU values

The capacity of a single lane is the maximum number of vehicles that can pass the stop line of a lane. The basic saturation flow is 1800 pcu/h, where pcu stands for 'passenger car unit', or PCE for 'passenger car equivalent'. Since lorries, buses and vans need more time than ordinary cars, they have a higher pcu value than 1. Motorcycles and ordinary cycles have a pcu value smaller than 1, see table 13-3. In calculations of the lane capacity, the PCU values should be taken into account, but they are not intended to be used for delay calculations, parking space requirements, queue length etc.

Table 13-3: PCU values per vehicle category

Vehicle category	PCU value
Passenger car	1.0
Lorry	1.5
Articulated Lorry	2.3
Bus	2.0
Motor cycle	0.4
Bicycle	0.2

13-6-2 The saturation flow of lanes

The saturation flow is affected by geometric and traffic conditions such as lane width, parked vehicles, turning movements etc. The expression to be used for the saturation flow s is:

$$s = \beta_1 s_0 \tag{13-6}$$

where s_0 is the basic saturation flow (1800 pcu/h) and β_1 is the saturation adjustment factor for geometric conditions.

Factor β_1 is a product of the following factors:

$$\beta_1 = n_{lanes}\, f_w\, f_{HV}\, f_g\, f_p\, f_{bb}\, f_a\, f_{RT}\, f_{LT} \tag{13-7}$$

with the following adjustment factors: n_{lanes}: number of lanes, f_W: adjustment factor for lane width, f_{HV}: adjustment factor for heavy vehicles, f_g:adjustment factor for grade, f_p: adjustment factor for parking facilities, f_{bb}: adjustment factor for bus blockage, f_a: adjustment factor for area type, f_{RT}: adjustment factor for right turning movements, f_{LT}: adjustment factor for left turning movements.

For the values of the adjustment factors the American Highway Capacity Manual (HCM, see HCM (2000)) can be used, but the Swedish manual (CAPCAL) and the German method (Harders) also give useful estimations of these factors. The differences are small, see for instance table 13-4, the HCM data is recommended. In Appendix 13-13 tables with saturation adjustment values are given.

Table 13-4: Lane width adjustment factor

Lane width	2.6	3.0	3.5	4.0
f_W HCM	0.88	0.94	0.99	1.03
f_W CAPCAL	0.93	1.00	1.02	1.03

13-6-3 Adjustment factors for turning movement

If a turning movement is controlled without conflicts (a 'protected' green phase), the capacity is affected only by the turning of the cars. If the movement has a permitted conflict, the capacity is also affected by the fact that the pedestrians (or opposing traffic) have to be given right of way.

The calculation of the capacity for left turning traffic with a permitted conflict for the opposing traffic requires a special calculation. One has to determine how much time the queue of the opposing traffic needs to clear. That time can not be used by turning traffic. The remainder of the green phase can be partially used, because sometimes there is still a vehicle that has right of way.

The saturated green time for the opposing direction depends on the load ratio q/s of the lane of the opposing traffic, where q is the demand, and s the (adjusted) saturation flow.

The length Q of the queue of the opposing flow is, at the start of green ($t_{r_{eff}}$ is the effective red time):

$$Q = q\, t_{r_{eff}} \tag{13-8}$$

After the time $t_{empty}$the queue is empty.

$$t_{empty} = q\, t_{r,eff} / (s - q) \tag{13-9}$$

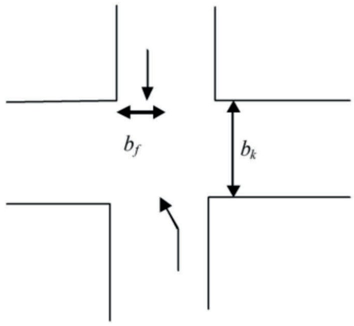
The remainder of the green time t_{unsat} is not saturated and gives possibilities for left turning traffic:

$$t_{unsat} = (T_C - t_{r,eff}) - (q\, t_{r,eff} / (s - q)) \tag{13-10}$$

In the saturated part of the green phase some vehicles may pass the intersection, if the lane is used by straight on and left turning vehicles. The number of vehicles that can depart in the saturated part of the green time depends on:

1. The left turning fraction of the traffic
2. The availability of queuing space on the intersection for left turning traffic

The number of vehicles that can depart during t_{unsat} depends on the volume of conflicting vehicles and the gap that is necessary. The length of this necessary gap depends on the geometry, see figure 13-10. For a necessary gap of 5.3 s the following figure gives the number of vehicles that can pass in the unsaturated green time as a function of the volume of conflicting opposing traffic and the fraction turning traffic. For non-shared lanes the fraction is, of course, 1.



		b_k	
		≤ 10	> 10 m
b_f	≤ 8 m	4.8 s	5.3
	> 8	5.3	5.8

Figure 13-10: Geometry of the intersection that determines the necessary gaps (in city centers the necessary gap is 0.5 s less)

The number of vehicles that can pass during t_{unsat} is:

$$N_u = s_u t_{unsat} \tag{13-11}$$

Finally one may suppose that left-turning vehicles that are queued on the intersection can clear during the yellow and clearance phase. Adding together all vehicles that can depart in the saturated and unsaturated green time($N_s + N_u$) and in the clearance time N_c gives a value for the saturation flow for the left turning lane:

$$S_{LT} = 3600 \left(N_s + N_u + N_c\right) / t_{g,eff} \tag{13-12}$$

13-7 Cycle time

The cycle time of a control structure depends on: the flow per direction, the adjusted saturation flow, the yellow time, and the internal lost time. After every ending of a green phase, there is lost time, which comprises of the the clearance time between successive conflicts, after start green: the start up delay, this is a lag start caused by reaction time and driver's actions and driving to the stop line, and the yellow use, this is using yellow time as green, because vehicles cannot stop for yellow.

The internal lost time of conflict group k is a summation of the lost time per movement i in conflict group k:

$$T_{LI,k} = \sum_i T_{start,i} + \sum_i \left(T_{yellow,i} - T_{yellow\_used,i}\right) + \sum_i T_{clear,(i,j)} \tag{13-13}$$

with: $T_{start,i}$: time lost after start green of i (start up delay, this is the reaction time of the driver), $T_{yellow,i}$: yellow time of movement i, $T_{yellow\_used,i}$: yellow time used as green, $T_{clear,(i,j)}$: clearance time of movement i to the succeeding movement j in the conflict group (for the last movement in the conflict group, this is the first movement of the next cycle, so the first movement of the conflict group).

In order to do a complete calculation of the cycle time, the clearance time of conflicts of the Kruithuisweg intersection is given in table 13-5, the timing and flow information about the start up delay, the yellow time and yellow use and minimum green, the flow and saturation flow is given in table 13-6.

Table 13-5: clearance time Kruithuisweg - Pr. Beatrixlaan

	01	02	08	09	10	12
01				0.0		
02				0.0	0.0	2.0
08						0.0
09	2.0	2.0				0.0
10		0.0				
12		1.0	4.0	1.0		

Table 13-6: timing and flow information Kruithuisweg - Pr. Beatrixlaan

	01	02	08	09	10	12
start up delay (s)	1.0	1.0	1.0	1.0	1.0	1.0
yellow time (s)	3.0	3.0	3.0	3.0	3.0	3.0
yellow use (s)	1.0	1.0	1.0	1.0	1.0	1.0
minimum green (s)	6.0	6.0	6.0	6.0	6.0	6.0
flow q_i (veh/h)	500	400	800	400	500	400
saturation flow s_i (veh/h)	1500	1800	1700	1700	1500	1700
load ratio $\frac{q_i}{s_i}$	0.333	0.222	0.470	0.235	0.333	0.235

In table 13-7 the internal lost time for conflict group 02-09-12 of the intersection Kruithuisweg – Pr. Beatrixlaan is given.

Table 13-7: Internal lost time Kruithuisweg– Pr. Beatrixlaan

Transition	02-09	09-12	12-02	02-09-12
Start up delay	1.0	1.0	1.0	
Yellow	3.0	3.0	3.0	
Yellow use	1.0	1.0	1.0	
Clearance	0.0	0.0	1.0	
Internal lost time	3.0	3.0	4.0	10.0

Another sequence of realisation of the movements in the conflict group can result into another value for the internal lost time. The smaller the internal lost time, the higher the capacity.

The Webster cycle time, the equation for the optimum cycle time calculation of conflict group k, $T_{Cwebster,k}$, is commonly used for fixed time control:

$$T_{Cwebster,k} = \frac{1.5T_{LI,k} + 5}{1 - \sum_i (q_i/s_i)}, \tag{13-14}$$

with: i: the movements of conflict group k, $T_{LI,k}$: the internal lost time of conflict group k, q_i: Flow rate of movement i (veh/s), s_i: Saturation flow rate of movement i (veh/s)

Webster derived this equation by assuming random arrivals of vehicles at the intersection, using the Poisson distribution, and this cycle time appeared to be optimum.

Assuming uniform arrivals and departures, the minimum cycle time for conflict group k can be calculated by:

$$T_{Cmin,k} = \frac{T_{LI,k}}{1 - \sum_i (q_i/s_i)} \tag{13-15}$$

The actual cycle time depends on the actual arrivals. For vehicle actuated control, the minimum cycle time is used to assess the control sequence. Considering the timing, the sequence with the smallest minimum cycle time is preferable.

In table 13-8 the internal lost time, the load ratio, the Webster cycle time and minimum cycle time per conflict group is given. The structure cycle time is the maximum cycle time of all conflict groups. Notwithstanding the fact that the first conflict group 02-09-12 has a size of 3, conflict group 08-12 is the critical conflict group with the largest conflict group cycle time, this due to the high load of 08, and the relative large clearance time 12-08. The Webster cycle time for this control structure is 68.0 s, the minimum cycle time is 34.0 s

Table 13-8: Conflict group cycle time Kruithuisweg– Pr. Beatrixlaan

Conflict group k	02-09-12	01-09-01	02-10-02	08-12-08	
T_{LIk} (s)	10.0	8.0	6.0	10.0	
$\sum \frac{q_k}{s_k}$	0.693	0.569	0.556	0.706	
$T_{Cwebster,k}$ (s)	65.1	39.4	31.5	68.0	$T_{Cwebster}$=68.0 s
$T_{Cmin,i}$ (s)	32.6	18.5	13.5	34.0	T_{Cmin} =34.0 s

If the cycle time is increased, because of increasing demand, the capacity is increased as well, because the relative share of the lost time in the cycle time reduces. However, very long cycle times increase the maximum waiting times and hardly contribute to the capacity. Figure 13-11 shows the percentage of time of effective use of a conflict area calculated for internal lost times of 10 and 20 s. Therefore, 120 s is commonly used as the maximum cycle time. If volumes are low, the cycle times will be shorter. Since signals can be skipped due to lack of traffic, the cycle time may be lower than the minimum cycle time.

13-8 Green times

The cycle time minus the internal lost time can be used for effective green. For the total amount of effective green time in the critical conflict group applies:

$$T_{g,eff,tot} = T_C - T_{LI}. \tag{13-16}$$

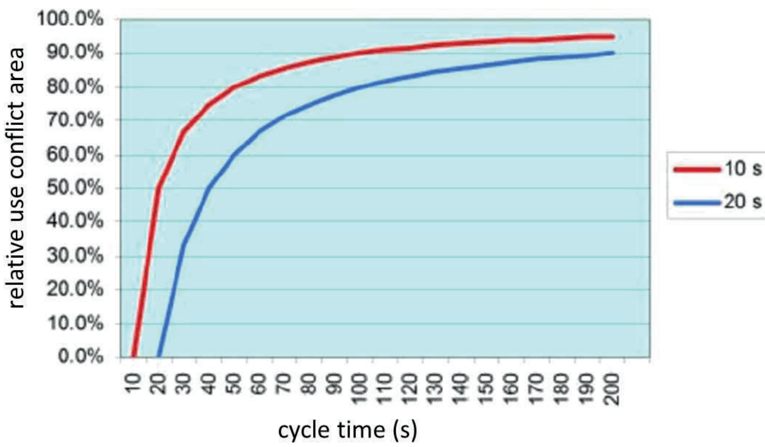


Figure 13-11: Relative use of the conflict area as function of the cycle time

Normally individual green times for movements depend on the "volume over saturation flow ratio", or "load" (q_i/s_i) of the concerned movement in relation to the sum of volume over saturation flow ratios. The green time in a cycle must be distributed over the signals in the critical conflict group, relative to their load $y_i = \frac{q_i}{s_i}$ with respect to the total load $Y = \sum_i y_i$:

$$T_{g,eff,i} = (y_i/Y)(T_C - T_{LI}), \tag{13-17}$$

for the actual green time, that is displayed in the cycle, applies:

$$T_{g,i} = T_{g,eff,i} + T_{start,i} - T_{yellow\_used,i}. \tag{13-18}$$

For the critical conflict group, the lost time can be rewritten with equation 13-15 to:

$$T_{LI} = (1 - Y)T_C, \tag{13-19}$$

combined with equation 13-17 this yields to:

$$T_{g,eff,i} = y_i T_{Cmin}, \tag{13-20}$$

so for the actual green time applies:

$$T_{g,i} = y_i T_{Cmin} + T_{start,i} - T_{yellow\_used,i} \tag{13-21}$$

If the resulting green time is shorter than a defined minimum duration, special attention should be paid on the cycle time and green time calculation (see Section 13-10).

In table 13-9 that correspond with the cycle time are given. For the Webster cycle time, first determine the green time of the movements in the critical conflict group, so 08 and 12. The green time for these movements are: T_{g08}=38.7 s, T_{g12}=19.3 s. Now for the first conflict group, the rest of the cycle time, minus the green time of 12 and the lost time can be divided between 09 and 02, this makes the green times: T_{g02}= 18.8 s, T_{g09}= 19.9 s. Stream 01 follows from conflict group 2, with given green time of 09, and lost time: 40.1 s, which is more than needed for the load ratio. Stream 10 follows the same procedure in conflict group 3, the cycle time, minus the internal lost time, minus the green time of 02, gives a green time of 43.2 s.

The green time for the minimum cycle time is for movement 02, 08, 09, and 12 have to be calculated with equation 13-16, but the green time for 01 and 10 can extend.

Table 13-9: Green times Kruithuisweg – Pr. Beatrixlaan

	01	02	08	09	10	12
Green time Webster (s)	40.1	18.8	38.6	19.9	43.2	19.3
Green time minimum cycle time (s)	18.0	7.6	16.0	8.0	20.4	8.0

13-9 Yellow times

The meaning of the yellow phase is defined as:
"Stop! If the driver is so close to the stopping line that he reasonably can not stop, than he may proceed".

For the Netherlands, this is described in art 68.1 of the RVV (regulations for traffic rules and traffic signs).

The duration of the yellow phase follows from this definition. As red light means that the driver has to stop, the yellow phase should be long enough to enable drivers to stop. The time needed depends on reaction time, the speed and the deceleration rate. If the signal changes from green to yellow, driver should decide either to stop or to go. The decision depends on the distance from the stop line (L_{stop}).

$$L_{stop} = v_0 T_{react} + \tfrac{1}{2}|a_{dec}|T_{dec}^2 \tag{13-22}$$

with:

L_{stop}	Distance from the stop line from where a driver can stop
v_0	Permitted speed
T_{react}	Reaction time
a_{dec}	Deceleration rate
T_{dec}	Deceleration time

As the deceleration time is:

$$T_{dec} = v_0/|a_{dec}| \tag{13-23}$$

The stop distance becomes:

$$L_{stop} = v_0 T_{react} + \tfrac{1}{2}v_0^2/|a_{dec}| \tag{13-24}$$

Vehicles closer to the stop line can not stop; vehicles further away can stop. The yellow time should be long enough to enable vehicles that can not stop to pass the stop line during yellow, this is L_{go}:

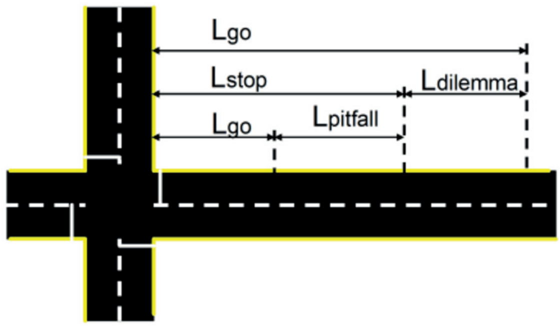


Figure 13-12: Decision point

Table 13-10: Yellow times with T_{react} =1s

	a_{dec}= 2.5 m/s^2	a_{dec}= 3.0 m/s^2	a_{dec}= 3.5 m/s^2	a_{dec}= 4.0 m/s^2
50 km/h	3.8 s	3.3 s	3.0 s	2.7 s
70 km/h	4.9 s	4.2 s	3.8 s	3.4 s
90 km/h	6.0 s	5.2 s	4.6 s	4.1 s

$$L_{go} = v_0 T_{yellow}, \tag{13-25}$$

where T_{yellow} is yellow time.

If the yellow time is too short L_{go} is shorter than L_{stop}, some of the vehicles can neither go nor stop. They are in the so-called a pitfall zone (see figure 13-12).

On the other hand, the yellow phase should not be too long. In that case L_{go} is longer than L_{stop} and drivers can decide either to go or to stop (dilemma zone). If a driver decides to stop and the following driver judges different, this can lead to rear end collisions.

So the optimum yellow time is found for $L_{go} = L_{stop}$, and with the equations 13-25 and 13-24 can be derived that applies:

$$T_{yellow} = T_{react} + ½ v_0/|a_{dec}| \tag{13-26}$$

Default values are: T_{react} =1s, a_{dec}= 3 m/s^2. A reaction time of 1 second is attainable for almost every driver.

The Dutch regulations require that cars should be able to decelerate with 3.85 m/s^2. In practice, cars can decelerate with 5 m/s^2 or even more. However, such high deceleration rates are not comfortable.

The default values for reaction time and deceleration rate result in rather long yellow times, see table 13-10. So, the "default driver" is not aggressive and L_{go} and L_{stop} are relative long preventing pitfall zones. A more aggressive following driver (most probably with a high desired speed), will have a high attention level (short reaction time), and will eventually accept a high deceleration rate. Such a driver needs a shorter L_{go} and L_{stop}. The designed decision point is further away from the stop line than some individual driver's decision point. Although the yellow time is calculated to prevent a design dilemma zone, a individual aggressive driver

Table 13-11: Decision point

	a_{dec}= 3.0 T_{react}=1.0	a_{dec}= 4.0 T_{react}=0.5	Aggressiveness $L_{dilemma}$
50 km/h	46.0 m	31.1 m	15.0 m
70 km/h	82.5 m	57.0 m	26.5 m
90 km/h	129.2 m	90.6 m	38.5 m

can still experience a dilemma zone (aggressiveness dilemma zone). Some road authorities want to prevent a signal change from green to yellow while a vehicle is in the dilemma zone. Assuming a aggressive driver with a t_{react}=0.5 s and a_{dec}= 4.0 m/s^2 the decision point for "stopping or going" will be closer to the stop line (see table 13-11). An inductive loop can detect vehicles in the aggressiveness dilemma zone and if present extend the green phase. This can be done by applying vehicle actuated control. Other road authorities do not want to award aggressive drivers who, anticipating on experience or knowledge that the light will remain green, even accelerate.

As postponing the transition to yellow will increase the cycle length, it decreases the effectiveness of the controlled intersection.

13-10 Minimum green times

To keep faith in signals (avoid disco light effects) a green phase should last at least a minimum green time, $T_{g,min}$. This can be the guaranteed time ($T_{g,min} = T_{gg}$). For cars, bikes and pedestrians the minimum duration is about 5 seconds. Buses phases can have a lower value of 4 seconds. For vehicle actuated control, the green phase will start with a short fixed duration (T_{gF}) to enable the vehicles in the waiting queue to start running in order to measure regular gaps to detect the end of the queue. Usually, the T_{gF} is longer than the guaranteed green time, about 6 seconds. If $T_{gF} > T_{gg}$, the minimum green is the fixed time green: $T_{g,min} = T_{gF}$.

If for a movement the calculated green time is less than the minimum green time, the calculation of cycle time and green times should be adjusted. This is done by removing the load of the green time concerned, and treating it as lost time. All the conflict group cycle times have to be recalculated, because due to the new green time, another conflict group might become critical. In this case, the Webster cycle time becomes:

$$T_{Cwebster} = \frac{1.5T_{LI+}5 + \sum_{i==i_{min}} T_{g,min}}{1 - \sum_{i \neq i_{min}} (q_i/s_i)} \tag{13-27}$$

The minimum cycle time becomes:

$$T_{Cmin} = \frac{T_{LI+} \sum_{i==i_{min}} T_{g,min}}{1 - \sum_{i \neq i_{min}} (q_i/s_i)} \tag{13-28}$$

The phase transition from green to yellow takes place for pre-timed control:

1. if the calculated green time expires.

2. The signal concerned is removed from the conflict group and so from $\sum_i (q_i/s_i)$ and the green period for the signal is considered as idle time. As the signal still has yellow and clearance times T_{LI} is not changed. Also in this case, recalculate the cycle time.

For vehicle actuated control the phase transition to yellow takes place:

1. if the minimum green time is reached and no traffic is present after the initial vehicle(s) have left after start of green.

2. if a detected priority vehicle forces conflicting signal to end their green phases. This is called "truncation". Also here, if the signal has just started the minimum green time will be regarded.

13-11 Flashing green times

In the Netherlands, instead of yellow, pedestrian signals have flashing green. The meaning of flashing green differs from the meaning of yellow. Flashing green means: "Go. However, the red phase starts soon". So, while flashing green, pedestrians can still start crossing; slow pedestrians better stop.

A slow pedestrian, who started crossing just before flashing green, should be able to finish crossing during flashing green and the clearance time; a fast pedestrian during the clearance time.

The pedestrian exiting time (Texit) clearance time should be:

$$T_{exit} = L_{exit}/v_{fast} \tag{13-29}$$

$$T_{exit} + T_{flash} = L_{exit}/v_{slow} \tag{13-30}$$

$$T_{flash} = L_{exit}/v_{slow} - L_{exit}/v_{fast} \tag{13-31}$$

where:

T_{exit}	Conflict area exiting time
T_{flash}	Flashing time
v_{fast}	Speed of fast pedestrian
v_{slow}	Speed of slow pedestrian
L_{exit}	Exiting distance

For convenience reasons, the minimum fixed green time (T_{gF}) should enable slow pedestrians to reach the middle of the crossing before the flashing green starts.

$$T_{gF} = (1/2)(L_{exit}/v_{slow}) \tag{13-32}$$

Example, with assumed speeds and exiting distance (see figure 13-13):

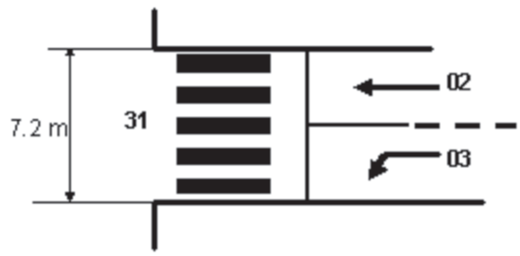


Figure 13-13: Pedestrian Crossing

v_{fast} = 1.2 m/s
v_{slow} = 1.0 m/s
L_{exit} = 7.2 m
T_{exit}= 7.2/1.2=6 s
T_{flash} = 7.2/1.0 – 7.2/1.2= 7.2 – 6.0 = 1.2 s
T_{gF} = (1/2)(7.2/1.2) = 3 s

13-12 Minimum red times

Like guarantee green time there is also a guarantee red time (T_{RG}), T_{RG} is usually 4 seconds.

The guarantee red time is needed in vehicle actuated control with a so called "Wait-in-Red" tactic. This tactic keeps signals on red if there is no traffic present. Now it can happen that, just after turning to red, a vehicle can request for green and if no conflicting movements are present, the green phase can return after a minimum red time.

Epilogue

When designing traffic control, always the safety should be considered. Often the safety is related to minimum delays (if the delay increases, drivers might take risks when they become impatient), hence the smallest possible cycle time is preferred, but also the design of the intersection and the behaviour of the traffic participants should be considered. If the minimum cycle time leads to a control that does not suit the intersection for a certain reason, another structure with a larger cycle time should be considered.

In Part II of this course, vehicle actuated control will be discussed in more detail: how detectors function, how the phase cycle is applied, and what tactics can be used to optimise the use of detectors and green phases.

Selected problems

For this chapter, consider problems: 293, 294, 295, 296

13-13 Appendix Saturation Adjustment tables

In this appendix adjustments factors are given, according to HCM.

Table 13-12: Adjustment factor for heavy vehicles

Percentage heavy vehicles	0	2	4	6	8	10	15	20	25
f_{HV}	1.00	0.99	0.98	0.97	0.96	0.95	0.93	0.91	0.89

Table 13-13: Adjustment factor for grade

	Downhill				Uphill		
Grade %	-6	-4	-2	0	+2	+4	+6
f_g HCM	1.03	1.02	1.01	1.00	0.99	0.98	0.97
f_g CAPCAL	1.06	1.04	1.02	1.00	0.98	0.96	0.94

Table 13-14: Adjustment factor for parking (HCM)

	# parking maneuvers per hour					
n_{lanes}	Not allowed	0	10	20	30	40
1	1.00	0.90	0.85	0.80	0.75	0.70
2	1.00	0.95	0.92	0.89	0.87	0.85
3	1.00	0.97	0.95	0.93	0.91	0.89

Table 13-15: Adjustment factor for bus blockage

	# stopping buses per hour				
# lanes in lane group	0	10	20	30	40
1	1.00	0.96	0.92	0.88	0.83
2	1.00	0.98	0.96	0.94	0.92
3	1.00	0.99	0.97	0.96	0.94

Table 13-16: Adjustment factor for area type (HCM)

Area type	Central business district	All other areas
f_a	0.90	1.00

Table 13-17: Adjustment factors for right turn

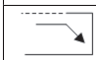
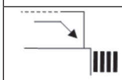
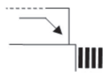
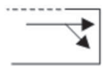
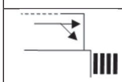
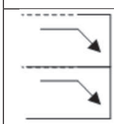
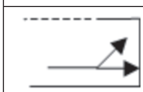
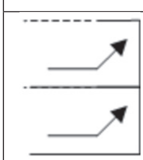
Lane group	Control	Formula
	Protected	$f_{RT} = 0.85$
	Permitted	$f_{RT} = max\{(0.85 - N_p/2100), 0.05\}$, N_p= # crossing pedestrians per hour
	Protected during part of the green phase, afterwards permitted	$f_{RT} = 0.85 - (N_p/2100)(1 - P_{RTA})$, P_{RTA} is fraction of right turns during protected phase
	Protected	$f_{RT} = 1.00 - 0.15P_{RT}$, P_{RT} is fraction right turns in lane
	Permitted	$f_{RT} = max\{(1.00 - P_{RT}(0.15 + N_p/2100)), 0.05\}$
	Protected	$f_{RT} = 0.85$

Table 13-18: Adjustment factor for protected left turns

Lane group	Control	Formula
	Protected	$f_{LT} = 0.95$
	Protected	$f_{LT} = 1.00/(1.00 + 0.5P_{LT})$, where P_{LT} is the fraction left turns
	Protected	$f_{LT} = 0.92$

Selected problems

For this chapter, consider problems: A-13-8

Alphabetical Index

Bibliography

Ahn, S. and Cassidy, M. (2007). Freeway traffic oscillations and vehicle lane-change maneuvers. In Allsop, R. E., Bell, M. G. H., and Heydecker, B. G., editors, *Proceedigs of the International Symposium of Transportation and Traffic Theory*, pages 691–710. Elsevier, Amsterdam.

Alessandri, A., Di Febbraro, A., Ferrara, A., and Punta, E. (1999). Nonlinear optimization for freeway control using variable-speed signaling. *IEEE Transactions on Vehicular Technology*, 48(6):2042–2052.

Åström, K. and Wittenmark, B. (1997). *Computer controlled systems*. 3rd ed. Prentice Hall.

Aw, A. and Rascle (2000). Resurrection of the "second order" models of traffic flow. *SIAM Journal on Applied Mathematics*, 60:916–938.

Bando, M., Hasebe, K., Nakayama, A., Shibata, A., and Sugiyama, Y. (1995). Dynamical model of traffic congestion and numerical simulation. *Physical Review E*, 51(2):1035. Copyright (C) 2007 The American Physical Society Please report any problems to prola@aps.org PRE.

Banks, J. H. (2005). Metering ramps to divert traffic around bottlenecks: Some elementary theory. *Transportation Research Record*, 1925:12–19.

Bliemer, M. (2007). Dynamic queuing and spillback in analytical multiclass dynamic network loading model. *Transportation Research Record: Journal of the Transportation Research Board*, (2029):14–21.

Board, T. R. (1998). Managing speed: review of current practice for setting and enforcing speed limits. Technical report, Transportatin Research Board, Committee for Guidance on Setting and Enforcing Speed Limits, Washington, D.C. National Academy Press, ISBN 0-309-06502-X.

Brackstone, M. and McDonald, M. (1999). Car-Following: A Historical Review. *Transportation Research F*, 2:181 – 186.

Cambridge Systematics, Inc. (2001). *Twin Cities Ramp Meter Evaluation – Final Report*. Cambridge Systematics, Inc., Oakland, California. Prepared for the Minnesota Department of Transportation.

Cappendijk, C. (2007). Verkeerskundige specificatie gecoördineerde toeritdosering. Technical report, Rijkswaterstaat Dienst Verkeer en Scheepvaart (Dutch ministry of transport: Transport Research Center).

Carlson, R. C., Papamichail, I., Papageorgiou, M., and Messmer, A. (2010). Optimal motorway traffic flow control involving variable speed limits and ramp metering. *Transportation Science*, 44(2):238–253.

Cassidy, M. and Bertini, R. (1999a). Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological*, 33(1):25 – 42.

Cassidy, M., Jang, K., and Daganzo, C. (2011). Macroscopic fundamental diagram for freeway networks: Theory and observation. In *Proceedings of the 90th Annual Meeting of the Transportation Research Board*.

Cassidy, M. J. and Bertini, R. L. (1999b). Some traffic features at freeway bottlenecks. *Transportation Research Part B*, B33:25 – 42.

Cassidy, M. J. and Rudjanakanoknad, J. (2005). Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B: Methodological*, 39(10):896 – 913. ISSN 0191-2615.

Chang, H., Wang, Y., Zhang, J., and Ioannou, P. A. (2007). An integrated roadway controller and its evaluation by microscopic simulator VISSIM. In *Proceedings of the European Control Conference*, pages 2436 – 2441. Kos, Greece.

Chaudhary, N. A. and Messer, C. J. (2000). Ramp metering technology and practice. Technical Report 2121-1, Texas Transportation Institute, The Texas A&M University System, College Station, Texas 77843-3135.

Chiabaut, N., Leclercq, L., and Buisson, C. (2010). From heterogeneous drivers to macroscopic patterns in congestion. *Transportation Research Part B: Methodological*, 44(2):299 – 308. ISSN 0191-2615.

Chien, C.-C., Zhang, Y., and Ioannou, P. A. (1997). Traffic density control for automated highway systems. *Automatica*, 33(7):1273–1285.

Chung, K., Rudjanakanoknad, J., and Cassidy, M. J. (2007). Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B: Methodological*, 41(1):82–95.

Corthout, R., Flötteröd, G., Viti, F., and Tampère, C. M. (2012). Non-unique flows in macroscopic first-order intersection models. *Transportation Research Part B: Methodological*, 46(3):343–359.

Courant, R., Friedrichs, K., and Lewy, H. (1928). Über die partiellen differenzengleichungen der mathematischen physik. *Mathematische annalen*, 100(1):32–74.

CROW (2014). *Handboek verkeerlichtenregelingen 2014*. CROW.

DACCORD (1996). Network level state estimation. Technical Report DACCORD TR1017 - deliverable 5.2.

Daganzo, C. (2007). Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, 41(1):49–62.

Daganzo, C. F. (1994). The Cell Transmission Model: a Dynamic Representation of Highway Traffic Consistent With the Hydrodynamic Theory. *Transportation research part B*, 28B(4):269–287.

Daganzo, C. F. (1995a). The cell transmission model, part ii: network traffic. *Transportation Research Part B: Methodological*, 29(2):79–93.

Daganzo, C. F. (1995b). Requiem for second-order fluid approximations of traffic flow. *Tranportation Research Part B: Methodological*, 29:277–286.

Daganzo, C. F. (1996). The nature of freeway gridlock and how to prevent it. In Lesort, J., editor, *Proceedings of the 13th International Symposium of Transportation and Traffic Theory*. Lyon, France.

Daganzo, C. F. (1997). *Fundamentals of Transportation and Traffic Operations*. Pergamon.

Daganzo, C. F. (2002a). A behavioral theory of multi-lane traffic flow. part i: Long homogeneous freeway sections. *Transportation Research Part B: Methodological*, 36(2):131 – 158. ISSN 0191-2615.

Daganzo, C. F. (2002b). A behavioral theory of multi-lane traffic flow. part ii: Merges and the onset of congestion. *Transportation Research Part B: Methodological*, 36(2):159 – 169. ISSN 0191-2615.

Daganzo, C. F. (2005). A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B: Methodological*, 39(2):187–196.

Drake, J. S., Schöfer, J. L., and May, A. (1967). A statistical analysis of speed density hypotheses. In Edie, L. C., Herman, R., and Rothery, R., editors, *Proceedings of the Third International Symposium on the Theory of Traffic Flow*. Elsevier North-Holland, New York.

Driels, M. R. and Shin, Y. S. (2004). Determining the number of iterations for monte carlo simulations of weapon effectiveness.

Dynasmart (2003). DYNAMSMART-X – User's Guide. Technical report, University of Maryland.

Edie, L. (1965). Discussion of traffic stream measurements and definitions. In *Proceedings of the Second International Symposium on the Theory of Traffic Flow*. OECD, Paris, France.

Flötteröd, G. and Rohde, J. (2011). Operational macroscopic modeling of complex urban road intersections. *Transportation Research Part B: Methodological*, 45(6):903–922.

Fortuijn, L. G. H. and P., H. S. (2015). Turbo roundabouts: Comparing capacity estimation on gap and flow level. In *Proceedings of the 94th Annual Meeting of the Transportation Research Board*.

Geroliminis, N. and Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9):759–770.

Gettman, D. and Head, L. (2003). Surrogate safety measures from traffic simulation models, Final report. Technical Report FHWA-RD-03-050, Federal Highway Administration.

Gipps, G. P. (1986). A model for the structure of lane-changing descisions. *Tranportation Research Part B: Methodological*, 20:403–414.

Godunov, S. K. (1959). A difference scheme for numerical computation of of discontinuous solutions of equations of fluid dynamics. *Math. Sb.*, 47:271 – 290.

Greenshields, B. D. (1934). A Study of Traffic Capacity. *Proceedings Highway Research Board*, 14:448–477.

Hajiahmadi, M., Knoop, V. L., De Schutter, B., and Hellendoorn, H. (2013). Optimal dynamic route guidance: A model predictive approach using the macroscopic fundamental diagram. In *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*, pages 1022–1028. IEEE.

Hall, F. L. and Agyemang-Duah, K. (1991). Freeway Capacity Drop and the Definition of Capacity. *Transportation Research Record: Journal of the Transportation Research Board No.1320*, pages 91–98.

Hall, F. L. and Agyemang-Duah, K. (1991). Freeway capacity drop and the definition of capacity. *Transportation Research Record*, (1320):91–98.

Hardman, E. J. (1996). Motorway speed control strategies using SISTM. In *Road Traffic Monitoring and Control*, number 422 in Conference Publication, pages 169–172. IEE.

HCM (2000). *Highway Capacity Manual*. National Research Council, Transportation Research Board, Washington, D.C.

Hegyi, A., De Schutter, B., and Hellendoorn, J. (2005). Optimal coordination of variable speed limits to suppress shock waves. *IEEE Transactions on Intelligent Transportation Systems*, 6(1):102–112.

Hegyi, A. and Hoogendoorn, S. (2010). Dynamic speed limit control to resolve shock waves on freeways – Field test results of the SPECIALIST algorithm. In *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems (ITSC 2010)*, pages 519–524. Madeira, Portugal.

Hegyi, A., Hoogendoorn, S., Schreuder, M., Stoelhorst, H., and Viti, F. (2008). SPECIALIST: a dynamic speed limit control algorithm based on shock wave theory. In *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems*, pages 827 – 832. Beijing, China.

Heikoop, H., editor (2011). *Capaciteitswaarden Infrastructuur Autosnelwegen*. Dienst Verkeer en Scheepvaart.

Helbing, D. (2003). A section-based queueing-theoretical traffic model for congestion and travel time analysis in networks. *Journal of Physics A: Mathematical and General*, 36(46):L593.

Helbing, D., Hennecke, A., and Treiber, M. (1999). Phase diagram of traffic states in the presence of inhomogeneities. *Physical Review Letters*, 82(21):4360.

Helly, W. . (1959). Simulation of bottlenecks in single lane traffic flow. In *Proceedings of the Symposium on Theory of Traffic Flow*, pages 207– 238. General Motors Research Laboratories, Elsevier, New York.

Hoogendoorn, S. and Bovy, P. (1996). Optimal control of freeway traffic by means of dynamic lane allocations. In *Proceedings of the 3rd World Congress on Intelligent Transportation Systems (ITS)*. Orlando, Florida.

Hoogendoorn, S. and Bovy, P. (1997). Optimal routing control by means of variable message signs. In *Proceedings of the 3rd TRAIL PhD Congress*. TRAIL Research School.

Hoogendoorn, S., Hoogendoorn, R. G., and Daamen, W. (2011). Wiedemann revisited. *Transportation Research Record: Journal of the Transportation Research Board*, 2260(1):152–162.

Hoogendoorn, S. P. (2005). Unified approach to estimating free speed distributions. *Transportation Research Part B: Methodological*, 39(8):709–727.

Ji, Y. and Geroliminis, N. (2012). On the spatial partitioning of urban transportation networks. *Transportation Research Part B: Methodological*, 46(10):1639–1656.

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

Kerner, B. S. (2004a). *The Physics of Traffic*. Understanding Complex Systems. Springer. ISBN: 978-3-540-20716-0.

Kerner, B. S. (2004b). *The Physics Of Traffic: Empirical Freeway Pattern Features, Engineering Applications, And Theory*. Springer, Berlin.

Kerner, B. S. and Rehborn, H. (1996). Experimental features and characteristics of traffic jams. *Physical Review E*, 53(2):R1297–R1300.

Keyvan-Ekbatani, M., Daamen, W., and Knoop, V. (2016). Categorization of the lane-change decision process on freeways. *Transportation Research part C*.

Knoop, V., Tamminga, G., , and Leclercq, L. (2016). Network transmission model: Application to a real world city. In *proceedings of the 95th Annual Meeting of the Transportation Reseach Board*.

Knoop, V. L., Duret, A., Buisson, C., and Van Arem, B. (2010). Lane distribution of traffic near merging zones – influence of variable speed limits. In *Proceedings of IEEE Intelligent Transportation Systems*.

Knoop, V. L. and Hoogendoorn, S. P. (2015). An area-aggregated dynamic traffic simulation model. *European Journal of Transport and Infrastructure Research (EJTIR), 15 (2), 2015*.

Knoop, V. L., Hoogendoorn, S. P., and Van Zuylen, H. J. (2007). Empirical Differences between Time Mean Speed and Space Mean Speed. In *Proceedings of Traffic and Granular Flow 07*. Springer, Paris, France.

Knoop, V. L., van Lint, H., and Hoogendoorn, S. P. (2015). Traffic dynamics: Its impact on the macroscopic fundamental diagram. *Physica A: Statistical Mechanics and its Applications*, 438:236–250.

Knoop, V. L., Van Zuylen, H. J., and Hoogendoorn, S. P. (2009). *Proceeding of the International Symposium of Transportation and Traffic Theory*, chapter Microscopic Traffic Behavior near Accidents. Springer, New York.

Koshi, M., Iwasaki, M., and Ohkura, I. (1981). Some findings and an overview on vehicular flow characteristics. In *Proceedings of the 8th International Symposium on Transportation and Traffic Theory*, pages 403–426. Univ. of Toronto Press, Toronto.

Kotsialos, A., Papageorgiou, M., Diakaki, C., Pavlis, Y., and Middelham, F. (2002). Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool metanet. *Intelligent Transportation Systems, IEEE Transactions on*, 3(4):282–292.

Kühne, R. D. (1991). Freeway control using a dynamic traffic flow model and vehicle reidentification techniques. *Transportation Research Record*, 1320:251–259.

Landman, R. L., Hegyi, A., and Hoogendoorn, S. P. (2015). Coordinated ramp metering based on on-ramp saturation time synchronization. *Transportation Research Record: Journal of the Transportation Research Board*, 2484(1):50–59.

Laval, J. and Castrillón, F. (2015). Stochastic approximations for the macroscopic fundamental diagram of urban networks. In *Transportation Research Procedia*, volume 7, page 615–630.

Laval, J. A. (2011). Hysteresis in traffic flow revisited: An improved measurement method. *Transportation Research Part B: Methodological*, 45(2):385 – 391.

Laval, J. A. and Daganzo, C. F. (2006). Lane-changing in traffic streams. *Transportation Research Part B: Methodological*, 40(3):251–264.

Laval, J. A. and Leclercq, L. (2013). The hamilton–jacobi partial differential equation and the three representations of traffic flow. *Transportation Research Part B: Methodological*, 52:17–30.

Lebacque, J. P. (1996). The gudunov scheme and what it means for first order traffic flow models. In Lesort, J. B., editor, *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, pages 647–677. Lyon, France.

Lebacque, J.-P. (2005). First-order macroscopic traffic flow models: Intersection modeling, network modeling. In *Transportation and Traffic Theory. Flow, Dynamics and Human Interaction. 16th International Symposium on Transportation and Traffic Theory.*

Leclercq, L., Chiabaut, N., Laval, J. A., and Buisson, C. (2007). Relaxation Phenomenon After Lane Changing. *Transportation Research Record: Journal of the Transportation Research Board, No. 1999*, pages 79–85.

Leclercq, L. and Geroliminis, N. (2013). Estimating mfds in simple networks with route choice. *Transportation Research Part B: Methodological*, 57:468–484.

Leclercq, L., Parzani, C., Knoop, V. L., Amourette, J., and Hoogendoorn, S. P. (2015). Macroscopic traffic dynamics with heterogeneous route patterns. *Transportation Research Part C: Emerging Technologies*, 59:292–307.

Lenz, H., Sollacher, R., and Lang, M. (2001). Standing waves and the influence of speed limits. In *Proceedings of the European Control Conference 2001*, pages 1228–1232. Porto, Portugal.

Levinson, D. and Zhang, L. (2006). Ramp meters on trial: Evidence from the Twin Cities metering holiday. *Transportation Research Part A*, 40A:810–828.

Li, P. Y., Horowitz, R., Alvarez, L., Frankel, J., and Robertson, A. M. (1995). Traffic flow stabilization. In *Proceedings of the American Control Conference*, pages 144–149. Seattle, Washington.

Lighthill, M. J. and Whitham, G. B. (1955a). On kinematic waves, II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society*, 229A(1178):317–345.

Lighthill, M. J. and Whitham, G. B. (1955b). On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads,. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 229(1178):317 – 345.

Messmer, A. and Papageorgiou, M. (2001). *METANET – A simulation program for motorway networks*. Technical University of Crete.

Middelham, F. (1999). A synthetic study of the network effects of ramp metering. Technical report, Transport Research Centre (AVV), Dutch Ministry of Transport, Public Works and Water Management.

Muller, T. H., Dijker, T., and Furth, P. G. (2004). Red clearance intervals: Theory and practice. *Transportation Research Record*, 1867:132–143.

Nagalur Subraveti, H. H. S., Knoop, V. L., and van Arem, B. (2019). First order multi-lane traffic flow model–an incentive based macroscopic model to represent lane change dynamics. *Transportmetrica B: Transport Dynamics*, 7(1):1758–1779.

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

Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic, part ii: Queueing at freeway bottlenecks. *Transportation Research Part B: Methodological*, 27(4):289 – 303.

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

Oh, S. and Yeo, H. (2012). Estimation of capacity drop in highway merging sections. *Transportation Research Record: Journal of the Transportation Research Board*, 2286(1):111–121.

Ossen, S. J. L. (2008). *Longitudinal Driving Behavior: Theory and Empirics*. Trail thesis series, Delft University of Technology.

Papageorgiou, M. (1980). A new approach to time-of-day control based on a dynamic freeway traffic model. *Transportation Research Part B*, 14B:349–360.

Papageorgiou, M., Blosseville, J.-M., and Haj-Salem, H. (1990). Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris: Part II: coordinated on-ramp metering. *Transportation Research Part A*, 24A(5):361–370.

Papageorgiou, M., Hadj-Salem, H., and Middelham, F. (1997a). ALINEA local ramp metering — summary of field results. *Transportation Research Record*, 1603:90–98.

Papageorgiou, M., Haj-Salem, H., and Middelham, F. (1997b). ALINEA local ramp metering: Summary of field result. In *Proceedings of the 76th Annual Meeting of the Transportation Research Board*. Washington, D.C.

Papageorgiou, M. and Kotsialos, A. (2000). Freeway ramp metering: An overview. In *Proceedings of the 3rd Annual IEEE Conference on Intelligent Transportation Systems (ITSC 2000)*, pages 228–239. Dearborn, Michigan, USA.

Papageorgiou, M. and Kotsialos, A. (2002). Freeway ramp metering: An overview. *IEEE Transactions on Intelligent Transportation Systems*, 3(4):271–280.

Papageorgiou, M. and Papamichail, I. (2007). Handbook of ramp metering. Document for european ramp metering project (euramp), Document for EUropean Ramp Metering Project (EURAMP).

Papageorigiou, M., Hadj-Salem, H., and Blosseville, J.-M. (1991). ALINEA: a local feedback control law for on-ramp metering. *Transportation Research Record*, (1320):58–64.

Papamichail, I., Kotsialos, A., Margonis, I., and Papageorgiou, M. (2010). Coordinated ramp metering for freeway networks - a model-predictive hierarchical control approach. *Transportation Research Part C: Emerging Technologies*, 18(3):311–331.

Papamichail, I. and Papageorgiou, M. (2008). Traffic-responsive linked ramp metering control. *IEEE Transactions on Intelligent Transportation Systems*, 9(1):111–121.

Payne, H. J. (1971). Models of freeway traffic and control, mathematical models of public systems. *Simulation Council Proceedings Series*, 28(1):51–61.

Popov, A., Hegyi, A., Babuška, R., and Werner, H. (2008). Distributed controller design approach to dynamic speed limit control against shockwaves on freeways. *Transportation Research Record*, 2086:93–99.

Pueboobpaphan, R. and van Arem, B. (2010). Understanding the relation between driver/vehicle characteristics and platoon/traffic flow stability for the design and assessment of cooperative cruise control. In *Proceedings of the 89th Annual Meeting of the Transportation Research Board*.

Rees, T., Harbord, B., Dixon, C., and Abou-Rhame, N. (2004). Speed-control and incident detection on the M25 controlled motorway (summary of results 1995-2002). Technical Report PPR033, TRL.

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

Richards, P. I. (1956b). Shock waves on the highway. *Operations Research 4*, 4:42 – 51.

Robinson, M., McGowen, P., Habets, A., and Strong, C. (2002). Safety applications of its in rural areas. Technical report, Science Applications International Corporation.

Schakel, W. J., Van Arem, B., and Knoop, V. (2012). Lmrs: An integrated lane change model with relaxation and synchronization. In *Proceedings of the 91st Annual Meeting of the Transportation Research Board*.

Schik, P. (2003). *Einfluss von Streckenbeeinflussungsanlagen auf die Kapazität von Autobahnabschnutten sowie die Stabilität des Verkehrsflusses*. Ph.D. thesis, Universität Stuttgart.

Schreiter, T., Van Lint, H., Yuan, Y., and P., H. S. (2010). Propagation wave speed estimation of freeway traffic with image processing tools. In *Proceedings of the 89th Annual Meeting*.

Sisiopiku, V. P. (2001). Variable speed control: technologies and practice. In *Proceedings of the 11th Annual Meeting of ITS America*, pages 1–11. Miami Beach, Florida.

Smits, E.-S., Bliemer, M. C., Pel, A. J., and van Arem, B. (2015). A family of macroscopic node models. *Transportation Research Part B: Methodological*, 74:20–39.

Smulders, S. (1989a). *Control of Freeway Traffic*. Ph.D. thesis, Delft University of Technology.

Smulders, S. (1989b). Control of freeway traffic flow. *CWI Tract*, 80.

Smulders, S. (1990). Control of freeway traffic flow by variable speed signs. *Transportation Research Part B*, 24B(2):111–132.

Smulders, S. A. and Helleman, D. E. (1998). Variable speed control: State-of-the-art and synthesis. In *Road Transport Information and Control*, number 454 in Conference Publication, pages 155–159. IEE.

Sontag, E. (1990). *Mathematical control theory, deterministic finite dimensional systems*. Springer-Verlag.

Srivastava, A. and Geroliminis, N. (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C: Emerging Technologies*, 30:161–177.

Tampère, C. M., Corthout, R., Cattrysse, D., and Immers, L. H. (2011). A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B: Methodological*, 45(1):289–309.

Toledo, T., Koutsopoulos, H. N., and Ben-Akiva, M. (2007). Integrated driving behavior modeling. *Transportation Research Part C: Emerging Technologies*, 15(2):96 – 112. ISSN 0968-090X.

Toledo, T., Koutsopoulos, H. N., and Ben-Akiva, M. (2009). Estimation of an integrated driving behavior model. *Transportation Research Part C: Emerging Technologies*, 17(4):365 – 380. ISSN 0968-090X.

Toorenburg, J. v. and Linden, R. v. d. (1996). Predictive control in traffic-management, an investigation of opportunities for application of predictive control in traffic-management on networks of major routes with the assumption that a short-term traffic-forecast is available. Technical report, Transpute & Dutch Ministry of Transport, Public Works and Water Management, The Netherlands.

Transportation Research Board, (2000). Highway Capacity Manual. Technical report.

Treiber, M. and Helbing, D. (2002a). Reconstructing the Spatio-Temporal Traffc Dynamics from Stationary Detector Data. *Cooper@tive Tr@nsport@tion Dyn@mics*, 1:3.1–3.24.

Treiber, M. and Helbing, D. (2002b). Reconstructing the spatio-temporal traffic dynamics from stationary detector data. *Cooper@tive Tr@nsport@tion Dyn@mics*, 1:3.1–3.24.

Treiber, M., Hennecke, A., and Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, 62(2):1805.

Treiber, M. and Kesting, A. (2013). *Traffic Flow Dynamics; Data, Models and Simulation*. Springer.

Treiber, M., Kesting, A., and Helbing, D. (2006). Delays, inaccuracies and anticipation in microscopic traffic models. *Physica A: Statistical Mechanics and its Applications*, 360(1):71–88.

van den Hoogen, E. and Smulders, S. (1994). Control by variable speed signs: results of the Dutch experiment. In *Proceedings of the 7th International Conference on Road Traffic Monitoring and Control*, IEE Conference Publication No. 391, pages 145–149. London, England.

Van der Gun, J. P., Pel, A. J., and Van Arem, B. (2015). A general activity-based methodology for simulating multimodal transportation networks during emergencies. In *ICEM 2015: 3rd International Conference on Evacuation Modeling and Management, Tainan, Taiwan, 1-3 June 2015*.

van der Weg, M. and van Run, H. (2014). Meer groen met 'intergroen'. *NM Magazine*, 2014-3:30–31.

Van Lint, J., Bertini, R. L., and Hoogendoorn, S. P. (2014). Data fusion solutions to compute performance measures for urban arterials. In *Symposium Celebrating 50 years of Traffic Flow Theory 2014 FTF Summer Meeting, Portland (USA), 11-13 August, 2014*. TRB.

van Lint, J. W., Hoogendoorn, S. P., and Schreuder, M. (2008a). Fastlane: New multiclass first-order traffic flow model. *Transportation Research Record: Journal of the Transportation Research Board*, 2088(1):177–187.

Van Lint, J. W. C. (2010). Empirical evaluation of new robust travel time estimation algorithms. *Transportation Research Record*, (2160):50–59.

van Lint, J. W. C. and Hoogendoorn, S. P. (2010). A robust and efficient method for fusing heterogeneous data from traffic sensors on freeways. *Computer-Aided Civil and Infrastructure Engineering*, 25(8):596–612.

van Lint, J. W. C., Hoogendoorn, S. P., and Schreuder, M. (2008b). Fastlane new multiclass first-order traffic flow model. *Transportation Research Record*, 2088:177–187.

van Wageningen-Kessels, F., van Lint, H., Hoogendoorn, S., and Vuik, K. (2010). Lagrangian formulation of multiclass kinematic wave model. *Transportation Research Record: Journal of the Transportation Research Board*, (2188):29–36.

van Wageningen-Kessels, F., van Lint, J., Hoogendoorn, S., and Vuik, C. (2008). Numerical time stepping schemes for macroscopic traffic flow models. In *TRAIL Congres DVD*.

Van Wageningen-Kessels, F. L. M. (2013). *Multi-class continuum traffic flow models: Analysis and simulation methods*. Ph.D. thesis, Delft University of Technology.

Wang, Y., Papageorgiou, M., and Messmer, A. (2002). A predictive feedback routing control strategy for freeway network traffic. In *Proceedings of the American Control Conference*, pages 3606–3611. Anchorage, Alaska.

Wattleworth, J. A. (1965). Peak-period analysis and control of a freeway system. *Highway Research Record*, 157:1–21.

Wiedemann, R. (1974). *Stimulation des Strassenverkehrsflusses*. Heft 8 der schriftenreihe des ifv, Universitâ€šÃ„â˜žt Karlsruhe.

Wilkie, J. K. (1997). Using variable speed limit signs to mitigate speed differentials upstream of reduced flow locations. Technical report, Department of Civil Engineering, Texas A& M University, College Station, Texas 77843. Prepared for CVEN 677 Advanced Surface Transportation Systems.

Wu, N. (2002). A new approach for modeling of Fundamental Diagrams. *Transportation Research Part A: Policy and Practice*, 36(10):867–884. Doi: DOI: 10.1016/S0965-8564(01)00043-X.

Yperman, I. (2007). The link transmission model for dynamic network loading.

Yperman, I., Logghe, S., Tampère, C. M. J., and Immers, B. (2006). Multicommodity link transmission model for dynamic network loading. In *Proceedings of the 85th Annual Meeting of the Transportation Research Board*.

Yuan, K., Knoop, V., and Hoogendoorn, S. (2015a). Capacity drop: A relation between the speed in congestion and the queue discharge rate. In *Proceedings of the 94th Annual Meeting of the Transportation Research Board*.

Yuan, Y. (2013). *Lagranian Multi-Class traffic State Estimation*. Ph.D. thesis, Delft University of Technology.

Yuan, Y., Daamen, W., Hoogendoorn, S., and Vrancken, J. (2009). Coordination concepts for ramp metering control in a freeway network. In *Proceedings of the 3rd TRAIL PhD Congress*. Redondo Beach, USA.

Yuan, Y., Van Wageningen-Kessels, F., Van Lint, H., and Hoogendoorn, S. (2015b). *Proceedings of Traffic and Granular Flow 2011*, chapter Two modeling and discretization choices for Lagrangian multi-class first-order traffic flow model and their related (dis-)advantages. Springer.

Zhang, H. M. (1999). A mathematical theory of traffic hysteresis. *Transportation Research Part B: Methodological*, 33(1):1 – 23.

Zhang, J., Boitor, A., and Ioannou, P. (2005). Design and evaluation of a roadway controller for freeway traffic. In *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, pages 543 – 548. Vienna, Austria.

Zhang, L. (2007). Traffic diversion effect of ramp metering at the individual and system levels. In *Proceedings of the 86th Annual Meeting of the Transportation Research Board*. Washington, D.C. CDROM paper 07-2087.

Appendix A

Test questions

The following pages contain exams with relevant questions. Problem set 1-11 are sets for a full exam of "traffic flow theory and simulation". The level of the questions is the same as can be expected for the course "Traffic Flow Modelling and Control"; note that some of the content of these questions falls in part 1, and some of the content falls in part 2. There is an example quesiton on Urban intersection control. From problem set 12, the exams are full exams of Traffic Flow Modelling and Control part 1. Overall, a 3 hour exam is expected to have approximately 55 to 60 points.

A-1 Problem set 1

A-1-1 Short questions

A road has a capacity of 2000 veh/h, a free flow speed of 120 km/h, a critical density of 25 veh/km and a jam density of 150 veh/km.

Exercise 1. *Draw a realistic fundamental diagram in the with the above properties, both in the flow-density plane and in the speed-density plane; indicate how speed can be found in the flow-density diagram. (4 points)*

Exercise 2. *Describe briefly (100 words) the assumptions in Daganzo's theory of slugs and rabbits (3 points)*

Exercise 3. *What are cumulative curves, and how are they constructed? (1 point)*

Exercise 4. *What is a vertical queuing model and how is this related to cumulative curves (2 points)*

Exercise 5. *What is higher, the space mean speed of the time mean speed? Why? (1 point)*

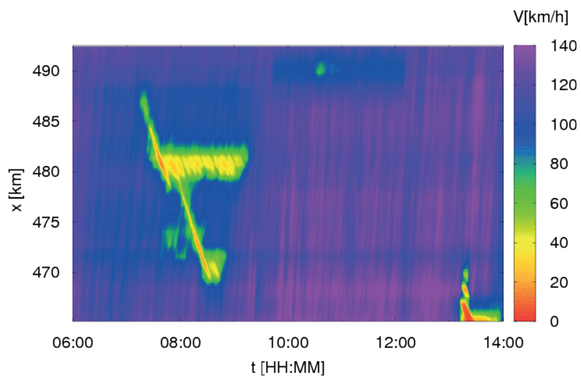
Exercise 6. *How can the space mean speed be calculated from individual local speed observations v_i. Give an equation. (1 point)*

A-1-2 State recognition

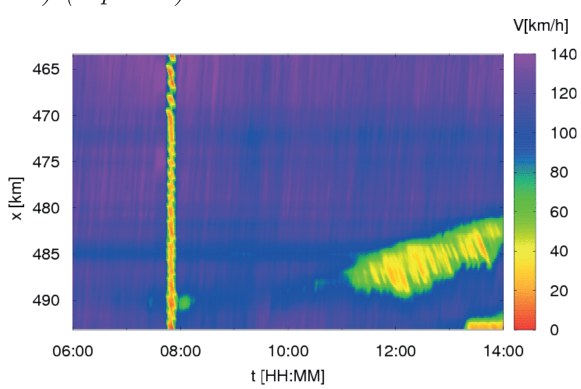
Exercise 7. *What are the phases according to three-phase traffic flow theory? How are they characterised? (2 points)*

Below, you find some traffic state figures from www.traffic-states.com. For each figure, indicate:

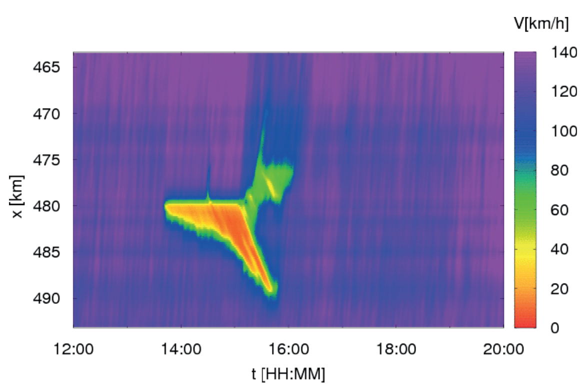
1. The driving direction (top-down or bottom-up) and explain why based on traffic flow theory (0.5pt per figure, spread over b-d) and
2. the traffic characteristics present, and the most likely causes for these (6 pt, spread over b-d)



Exercise 8.
(up to 10.00 am) (2 points)



Exercise 9.
(3 points)



Exercise 10.
(3.5 points)

A-1-3 Simulation model

Exercise 11. *Explain briefly (indication: 100 words) Newells car-following model. Also plot a space-time diagram indicating these properties. (3 points)*

Exercise 12. *Name the three types of (in)stability in traffic flow, and explain briefly their effect (50 words and 1 graph per type) (3 points)*

Exercise 13. *Can multi-anticipation improve stability? Which type and how? (3 points)*

Exercise 14. *How can lane-changing cause a breakdown? (50 words) (1 point)*

Exercise 15. *How can lane-changing prevent a breakdown? (Hint: use the lane distribution.) (2 points)*

A-1-4 Moving bottleneck

Imagine a truck drivers strike in France on a 3-lane motorway from time t=t0. For the road you may assume a triangular fundamental diagram with a capacity of 2000 veh/h/lane, a critical density of 25 veh/km/lane and a jam density of 125 veh/h/lane. Suppose there is no traffic jam at t<t0, and a constant demand of 5000 veh/h. Assume the truck drivers drive at 30 km/h, not allowing any vehicle to pass.

Exercise 16. *Calculate is the maximum flow on the road behind the trucks? (3 points)*

Exercise 17. *Sketch the traffic flow operations in a space-time diagram. Pay attention to the direction of the shockwave, and note the propagation speeds of the other waves. (3 points)*

After a while, the trucks leave the road.

Exercise 18. *Sketch again the traffic operations in the space time diagram. Do this until the traffic jam is solved (if applicable), or until the traffic states propagate linearly if demand does not change. (3 points)*

Exercise 19. *Calculate the speed at which the tail of the queue propagate backwards? (1 point)*

Exercise 20. *In the plot of question 2, draw the trajectory of the vehicle arriving at the tail of the queue at the moment the strike ends (Use a different colour or style than the shockwaves). (1 point)*

Exercise 21. *What is the speed of the vehicle in the jam? (1 point)*

A-1-5 Marathon Delft

Imagine a marathon organised in Delft. There are 30.000 participants, starting at a roadway section which is 25 meters wide.

Exercise 22. *Give a realistic estimate for the capacity of the roadway in runners per hour. Base your answer on the width and headway of a runner*

Exercise 23. *Calculate how long it would take for all runners to start? (1 point)*

Suppose the runners have uniform distribution of running times from 2.5 to 4.5 hours, and all run at a constant speed.

Exercise 24. *Given your earlier assumptions, calculate the minimum required width of the road halfway the track to avoid congestion? For reasons of simplicity, you may assume that they all start at the same moment (instead of your answer at question 5b). (5 points)*

A-2 Problem set 2

A-2-1 Short open questions

Exercise 25. *Explain why Kerner claims that in his three phase theory there is no fundamental diagram (indication: 25 words) (1 point)*

Exercise 26. *What is a Macroscopic Fundamental Diagram (indication: 25 words)? (1 point)*

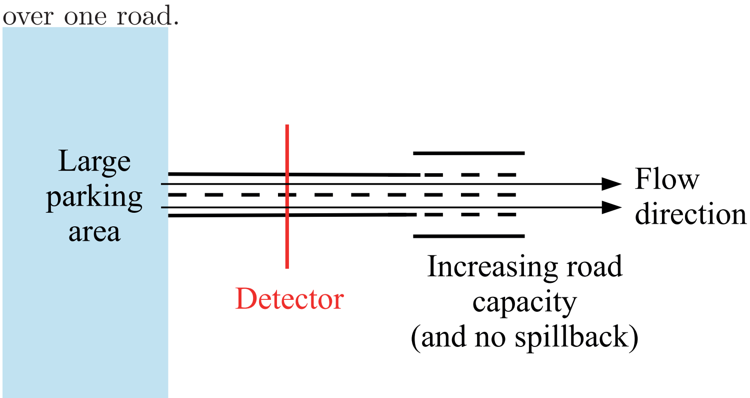
Exercise 27. *What are similarities and differences of a Macroscopic Fundamental Diagram compared to a normal fundamental diagram? Explain based on the underlying phenomena (indication: 75 words). (2 points)*

Exercise 28. *Explain in words the how the flow in from one cell to the next is calculated according to the Cell Transmission Model (indication: 100 words; equations can be useful, but need to be explained). (2 points)*

Exercise 29. *What do we mean with Lagrangian coordinates in macroscopic traffic flow simulation? Which variables are used? Why is this an advantage over other systems? (indication: total 100 words) (3 points)*

A-2-2 Leaving the parking lot

Consider the situation where many cars are gathered at one place, and they can only leave over one road.



The above figure shows a simplified representation of the road layout. There is a detector at the red line, in the two-lane road stretch. Downstream of the detector the road widens and there are no further downstream bottlenecks.

Exercise 30. *Explain what the capacity drop is (be precise in your wording!). (1 point)*

Exercise 31. *How large is the capacity drop? Give a typical interval bound. (1 point)*

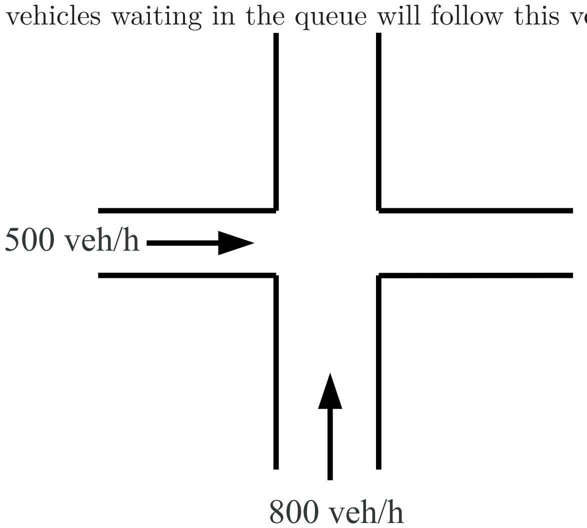
Exercise 32. *Sketch the traffic flow at the detector (indicated in the figure) as function of the demand (in veh/h, ranging from 0 to three times the road capacity) from the parking lot. (2 points)*

Exercise 33. *Explain the general shape you draw in the previous question. (1 point)*

Exercise 34. *Give some rough estimates for values in your graph. (1 point)*

A-2-3 Traffic lights

Consider a junction with a traffic light with equal green time g, and a clearance time of 2 seconds (i.e., the time needed to clear the junction; during this time, both directions are red). If the traffic light turns green, the first vehicle needs 3 seconds to cross the line. Afterwards, vehicles waiting in the queue will follow this vehicle with a 2 second headway.



Exercise 35. *What is the fraction of time traffic is flowing over the stop line per direction as function of the cycle time c? (2 points)*

Exercise 36. *What is the maximum flow per direction as function of the cycle time? In your equation, what are the units for the variables? (3 points)*

Suppose traffic flow from direction 1 is 500 veh/h and traffic from direction 2 is 800 veh/h, and suppose a uniform arrival pattern.

Exercise 37. *What is the minimum cycle time to ensure that no queue remains at the end of the cycle? Does an approach with vertical queuing models yield the same result as shockwave theory? Why? (3 points)*

This cycle time is fixed now at 120 seconds. We relax the assumption of uniform distribution to a more realistic exponential arrival pattern.

Exercise 38. *To which distribution function for the number of arrivals per cycle (N) does this lead? Name this distribution (1 pt). (1 point)*

The probability distribution function of X is given by:

$$f(k;\lambda) = \Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{A-1}$$

with $e \sim 2.71828...$) and $k!$ is the factorial of k. The positive real number λ is equal to the expected value of X, which also equals the variance.
Assume there is no traffic waiting when the traffic light turns red at the beginning of the cycle.

Exercise 39. *Express the probability p that there are vehicles remaining in the queue for direction 2 when the traffic light turns red at the end of the cycle in an equation (p = ...). Write your answer as mathematical expression in which you specify the variables. Avoid infinite series. There is no need to calculate the final answer as a number (3 points)*

Exercise 40. *Argue whether this probability is higher or lower than if a uniform arrival process is assumed (1 point)*

A-2-4 Car-following model

Consider the IDM car-following model, prescribing the following acceleration:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = a_0 \left(1 - \left(\frac{v}{v_0}\right)^4 - \left(\frac{s^*(v,\Delta v)}{s}\right)^2\right) \tag{A-2}$$

with the desired distance s^* as function of speed v and speed difference Δv:

$$s^*(v,\Delta v) = s_0 + v\,T + \frac{v\,\Delta v}{2\sqrt{a\,b}} \tag{A-3}$$

Exercise 41. *Explain in words the working of this car-following; i.e. comment on the acceleration. (3 points)*

Another car-following model is the relatively simple car-following of Newell.

Exercise 42. *How does the Newell car-following model work? (indication: 50 words) (1 point)*

Exercise 43. *Name two weak points of both car-following models (IDM and Newell) (2 points)*

Exercise 44. *Give two reasons to choose the IDM model over Newell's model (2 points)*

Exercise 45. *Give two reasons to choose Newell's model over the IDM model (2 points)*

A-2-5 Pedestrians in a narrow tunnel

We consider a pedestrian flow through a narrow tunnel, which forms the bottleneck in the network under high demand. The tunnel is 4 meters wide and 20 meters long. You may assume that the pedestrians are distributed evenly across the width of the tunnel. The pedestrian flow characteristics are described by a triangular fundamental diagram, with free speed $v_0 = 1.5m/s$, critical density $k_c = 1P/m^2$ and jam density $k_j = 6P/m^2$ (P stands for pedestrian).

Exercise 46. *Draw the fundamental diagram for the tunnel. What is the capacity of the tunnel expressed in P/min? (2 points)*

Between 11-12 am, the average flow through the tunnel is q = 180 P/min.

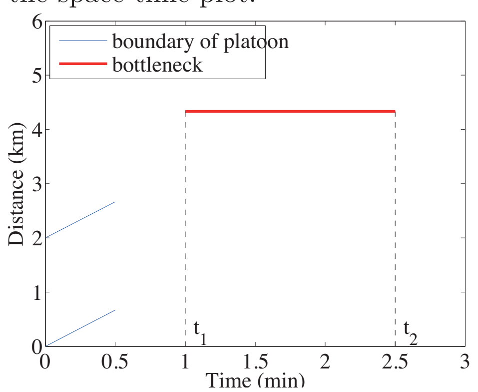
Exercise 47. *Assuming stationary conditions, what is the density in the tunnel expressed in P/m^2 ? (2 points)*

A-2-6 Moving bottleneck with different speeds

Consider a two-lane road with traffic in opposing directions. Assume a triangular fundamental diagram with a free speed of 80 km/h, a critical density of 20 veh/km and a jam density of 120 veh/h. At t=0 there is a platoon of 30 vehicles on the road with equal spacing in the section x=0 to x=2km. There are no other vehicles on the road. In this question, we will consider the effect of a speed reduction to 15 km/h.

Exercise 48. *Calculate the density on the road in the platoon. Indicate the conditions in the fundamental diagram. Make a clear, large drawing of the fundamental diagram such that you can reuse it for indicating states in the remaining of the question. (1 point)*

In each of the following subquestions, consider these initial conditions and no other bottlenecks than introduced in that subquestion – i.e., there never is more than one bottleneck.
Suppose there is a local, stationary bottleneck where drivers have to pass at 15 km/h from time $t1$ to $t2$. The figure below shows the position and the duration compared to the platoon in the space time plot.



Exercise 49. *Sketch the resulting traffic operations in a space-time plot. Also sketch some trajectories in the plot, using a different color or style.* Indicate how you construct the diagram

(i.e., indicate how you get the lines, angles, and distances in the plot). There is no need to calculate the exact traffic states. Additional information: the combination of speeds and the fundamental diagram will lead to congestion. (4 points)

Suppose there is a temporal speed reduction to 15 km/h for a short period of time over the whole length of the road at the same time. You may assume the platoon to be completely on the road.

Exercise 50. *Sketch the resulting traffic operations in a space-time plot. Also sketch some trajectories in the plot, using a different color or style.* Indicate how you construct the diagram *(i.e., indicate how you get the lines, angles, and distances in the plot). (2 points)*

Suppose there is a large tractor moving in the opposite direction of the traffic at 5 km/h from x1 to x2. This wide vehicle causes vehicles that are next to it to reduce speed to 15 km/h. The speed (5 km/h) is lower than the speed of the shock wave at the tail of the queue in question b

Exercise 51. *Sketch the resulting traffic operations in a space-time plot. Also sketch some trajectories in the plot, using a different color or style.* Indicate how you construct the diagram *(i.e., indicate how you get the lines, angles, and distances in the plot). (4 points)*

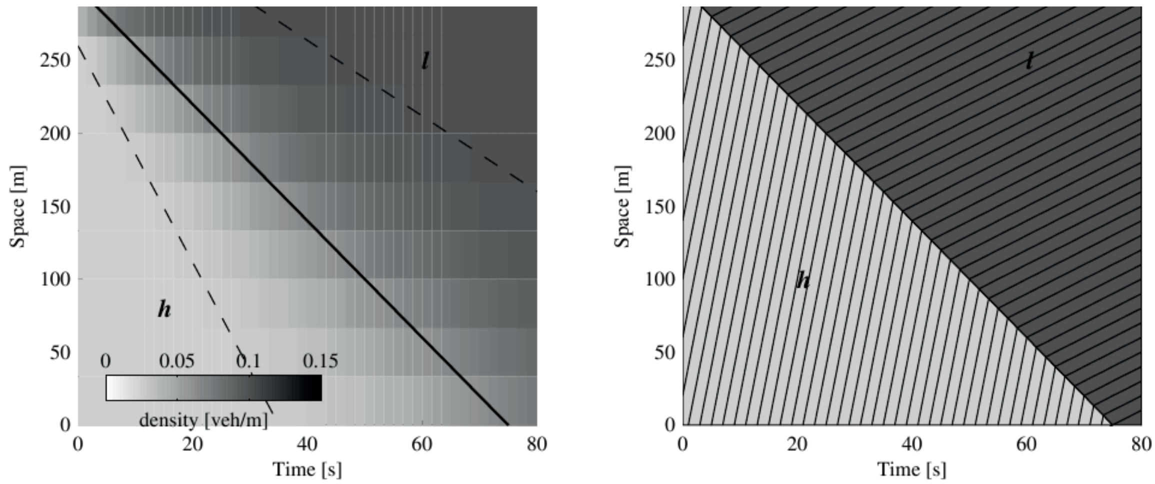
Suppose there is a large tractor moving in the opposite direction of the traffic at 50 km/h. This wide vehicle causes vehicles that are next to it to reduce speed to 15 km/h. The tractor speed (50 km/h) is higher than the speed of the shock wave of the tail of the queue in question b.

Exercise 52. *Sketch the resulting traffic operations in a space-time plot. Also sketch some trajectories in the plot, using a different color or style.* Indicate how you construct the diagram *(i.e., indicate how you get the lines, angles, and distances in the plot). (Hint: it can help to consider a finite tractor length, but it is not necessary) (5 points)*

A-3 Problem set 3

A-3-1 Short open questions

Below, two simulation results of a homogeneous congested area spilling back upstream are plotted. You see a high density, congested area, indicated with h downstream of x=0 at t=0, and a low density, uncongested area, indicated by l. The driving direction is up. One of the figures plots the results of a simulation in Eulearian coordinates and one of a simulation in Lagrangian coordinates.



Exercise 53. *Explain which of the two is Eulerian and which one Lagrangian.(1 point)*

Exercise 54. *Give the two main advantages of using Lagrangian coordinates. (2 points)*

Exercise 55. *Explain what a pce value is. (30 words) (1 point)*

Exercise 56. *How does the pce value of trucks depend on the speed, qualitatively (i.e., does it increase, decrease, ...). Give your reasoning*

Exercise 57. *How does the pce value of trucks depend on the speed, qualitatively (i.e., does it increase, decrease, ...). Give your reasoning (2 points)*

A-3-2 Multi-leader car-following models

Consider the IDM car-following model, prescribing the following acceleration:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = a_0 \left(1 - \left(\frac{v}{v_0}\right)^4 - \left(\frac{s^*(v, \Delta v)}{s}\right)^2\right) \tag{A-4}$$

In this equation, v is the speed of the following vehicle, v_0 a reference speed, s the distance headway and s^* a reference distance headway as function of speed v and the difference in speed with the leader, Δv:

$$s^*(v, \Delta v) = s_0 + v\,T + \frac{v\,\Delta v}{2\sqrt{a\,b}} \tag{A-5}$$

In this equation s_0 is a desired distance at standstill, and a is a reference acceleration and b is the maximum comfortable braking.

Exercise 58. *Explain in words the working of this car-following model; i.e. comment on the acceleration. Only comment on equation A-4. (indication: 50 words) (2 points)*

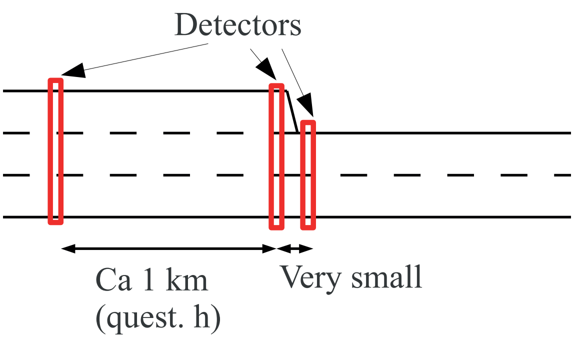
Exercise 59. *Explain in words what is meant with multiple-leader car-following models. Only comment on the multi-leader part. (indication: 50 words) (1 point)*

Exercise 60. *Reformulate the IDM model into a multiple-leader car-following model by including two leaders. Give your reasoning, and, for all points, formulate those reasonings into equations. (3 points)*

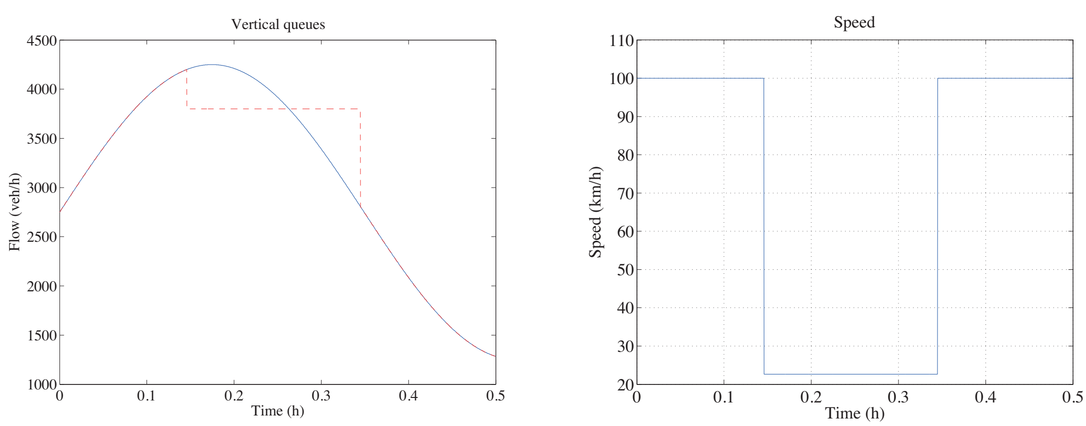
Exercise 61. *Formulate the generalised car-following IDM model considering n leaders (3 points)*

A-3-3 Measuring the speed at a cross section

Consider the following road layout with a lane drop from 3 to 2 lanes; traffic is flowing left to right.



The demand, as well as the flow at the downstream detector, is given in the figure below. The speed at the detector, located just upstream (you may assume no spacing between the bottleneck and the detector) of the bottleneck, is as follows.



(a) subfigure[Flow and demand **(b)** subfigure[Speed

In the question, you might need values from the graph. Slightly misreading the graph is not a problem, but *mention the values you directly read from the graph and where you find these.*

Exercise 62. *In figure a with the flow and the demand, which line is the demand and which line the flow. Argue why (1 point)*

Exercise 63. *Give the free flow capacity (and the reasoning how you find it) (2 points)*

Exercise 64. *Give the queue discharge rate. (and the reasoning how you find it) (2 points)*

Assume the fundamental diagram per lane is the same for all lanes, at all locations in this setting.

Exercise 65. *Explain why can this situation not be described with a triangular fundamental diagram. (1 point)*

Exercise 66. *Draw the simplest fundamental diagram possible for the three lane section (aggregated over all lanes). Explain how you find the values for the relevant points, and give calculate them. (5 points)*

Exercise 67. *On you answer sheet, sketch the demand as function of time (i.e., copy the above figure, indicate the times of speed change in the graph - no points given) and in the same graph, sketch the resulting flow at the detector. Pay attention to the times at which the flow changes compared to the time of the speed changes. (2 points)*

Now the delay is analysed by slanted cumulative curves.

Exercise 68. *Explain the offset you choose. (1 point)*

Exercise 69. *Sketch the slanted cumulative curve of the flow and the demand for the situation at hand. (2 points)*

Exercise 70. *Indicate in the figure how you can determine the delay. (1 point)*

Now a second detector is constructed approximately 1 km upstream of the bottleneck.

Exercise 71. *Sketch the demand as function of time (i.e., copy the second figure of this question, indicate the times of speed change in the graph - no points given) and in the same graph, sketch the resulting flow at this second detector. Pay attention to the times at which the flow changes compared to the time of the speed changes. (2 points)*

We now relax the assumption of all traffic states at the fundamental diagram to a more realistic situation (i.e., including demand and supply variations). In this situation, We measure the one-minute aggregated (harmonically averaged) speeds during one month.

Exercise 72. *Sketch a probability distribution or histogram of these speeds. Explain the shape (2 points)*

A-3-4 Moving bottleneck

Consider a two-lane motorway. Assume a triangular fundamental diagram with a free flow speed of 80 km/h, a critical density of 25 veh/h/lane and a jam density of 150 veh/km/lane. The inflow is stationary at 2000 veh/h.

Exercise 73. *Draw the fundamental diagram and calculate the capacity of the road stretch2*

Consider a wide vehicle (special transport) driving slowly is entering the road. There are no overtaking possibilities. This creates congestion, *of which the tail happens to stay at the same position.*

Exercise 74. *Calculate the flow in the jam. (Hint: you can use the speed of the shockwave at the tail of the queue)1*

Exercise 75. *Calculate the density in the jam.2*

Exercise 76. *Calculate the speed of the special transport.2*

After 5 km, the truck leaves the road.

Exercise 77. *Construct the space-time diagram of this situation, from before the moment the truck enters the road to after the moment the traffic situation is stationary. Explain how you find the speed of the solving of congestion; you may re-use the figure you created in a.4*

Exercise 78. *At the maximum queue length, how many vehicles are in congestion1*

A-3-5 Multi-lane traffic flow

Exercise 79. *Give the names of two regimes according to Daganzo's theory of slugs and rabbits (only names are required). (1 point)*

Below, you find a Google Earth image of the A1 motorway near Bathmen. Consider the right to left (east to west) direction

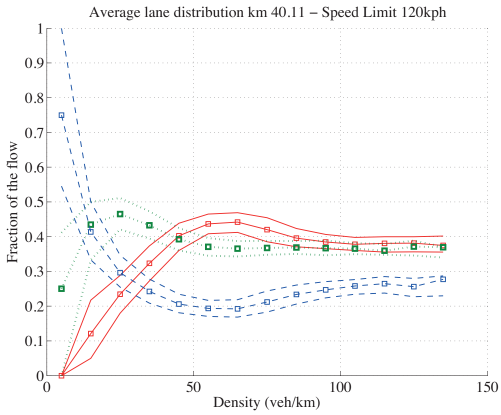
Exercise 80. *To which of the regimes of question a) does the situation in the figure resemble most. Explain why (2 points)*

There are 4 trucks and 7 passenger cars (including one delivery van) in the image. The fraction of passenger cars in the density is hence 7/11.

Exercise 81. *What would you expect from the fraction of passenger cars in the flow: is this higher, lower or equal to 7/11. Motivate your answer 2 points*

Exercise 82. *Calculate the fraction of the is passenger cars in the flow. Assume reasonable speeds for the different vehicle types in your answer – state your assumptions clearly. (4 points)*

Below, a figure of a lane flow distribution for a three lane road in the Netherlands is given.



Exercise 83. *In the figure, which color matches which lane. Argue why (1 point)*

Exercise 84. *Describe briefly (100 words) the assumptions in Daganzo's theory of slugs and rabbits (3 points)*

Exercise 85. *Using Daganzo's theory of slugs and rabbits, explain the differences in flow between most used lane (blue) and the least used lane (red) at a density of 75 veh/km. (2 points)*

A-4 Problem set 4

A-4-1 Short questions

Exercise 86. *What does a Macroscopic Fundamental Diagram describe? Make sure your answer shows clearly the differences between an MFD and a fundamental diagram.(2 points)*

Exercise 87. *What is qualitatively the effect of inhomogeneity in the network on the MFD? (No explanation needed) (1 points)*

Exercise 88. *In Lagrangian coordinates, the fundamental diagram is often expressed in spacing (horizontal) - speed (vertical) form. Sketch a fundamental diagram in the spacing-speed plane (2 points)*

A-4-2 From car-following to a fundamental diagram

The optimal velocity model is a car-following model specifying the acceleration a as follows:

$$a = a_0(v^* - v) \tag{A-6}$$

In this equation, v is the speed of the vehicle, and a_0 a reference acceleration (tunable parameter, constant for a specific vehicle-driver combination). v^* is determined as follows:

$$v* = 16.8(\tanh(0.086(\Delta x - 25) + 0.913)) \tag{A-7}$$

In this equation, Δx is the distance (in meters) between the vehicle and its leader.

Exercise 89. *Explain qualitatively the working of this car-following model; i.e. comment on these two equations. (2 points)*

Exercise 90. *What are the conditions for which a fundamental diagram holds? (1 point)*

Exercise 91. *Derive the expression a fundamental diagram (flow as function of density) for these conditions using the OVM model4*

Exercise 92. *What are the values capacity, free speed and jam density (either derive the value or use you graphical calculator to determine this – you may round numbers to the precision of your liking...) (1 point)*

Exercise 93. *Argue whether the values in the previous question are reasonable for the spacing/speed relationship? If so, for what conditions (urban, freeway...?) If not, can you change the parameters (numbers) of the model (A-6 and A-7) to get a proper fundamental diagram (if so: what are reasonable numbers). Show in your answer that you understand the traffic operations, and/or the model working.*

A-4-3 MOBIL lane change model

Consider the mobil lane change model. A driver c has several options: change lanes to the left, to the right or stay in his current lane. For each of the options a total utility (denoted U_{tot}) can be calculated.

$$U_{\text{tot}} = U_c + p \sum_{i \in \text{other drivers}} U_i \tag{A-8}$$

The utility for the driver i U_i is assumed to be its instantaneous acceleration, as computed using a car-following model (the Intelligent Driver Model – although the specific model is not relevant for the question). The driver is assumed to take the option with the highest utility.

Exercise 94. *Explain the working of the model in words (2 points)*

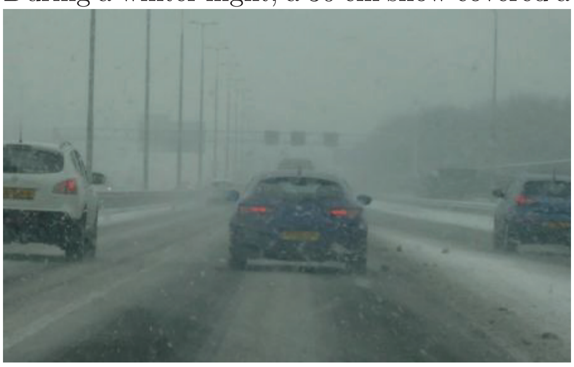
Exercise 95. *What value for p can be expected (1 point)*

A situation like this can be observed in practice

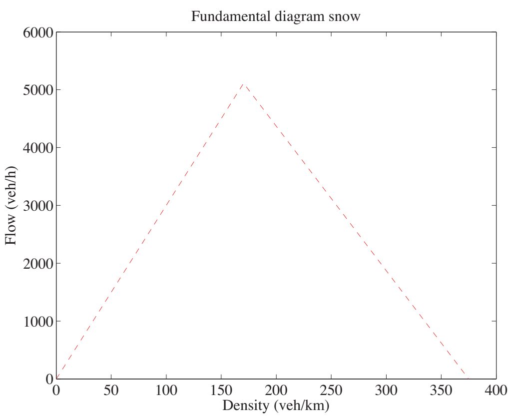
Exercise 96. *Does the MOBIL lane change model explain this? Reasoning from the model, explain how (if it does explain) or why (if it does not explain). (2 points)*

A-4-4 Snow plow

During a winter night, a 30 cm snow covered a three lane motorway fell. Traffic is still moving.



This changes the traffic operations. Assume a triangular fundamental diagram. The free flow speed reduces to 30 km/h, the jam density to 125 veh/km/lane. The capacity is 5000 veh/h.



A truck spins and cannot move further, thereby blocking the road completely, not allowing other vehicles to pass. This leads to a traffic jam. The inflow is 1000 veh/h.

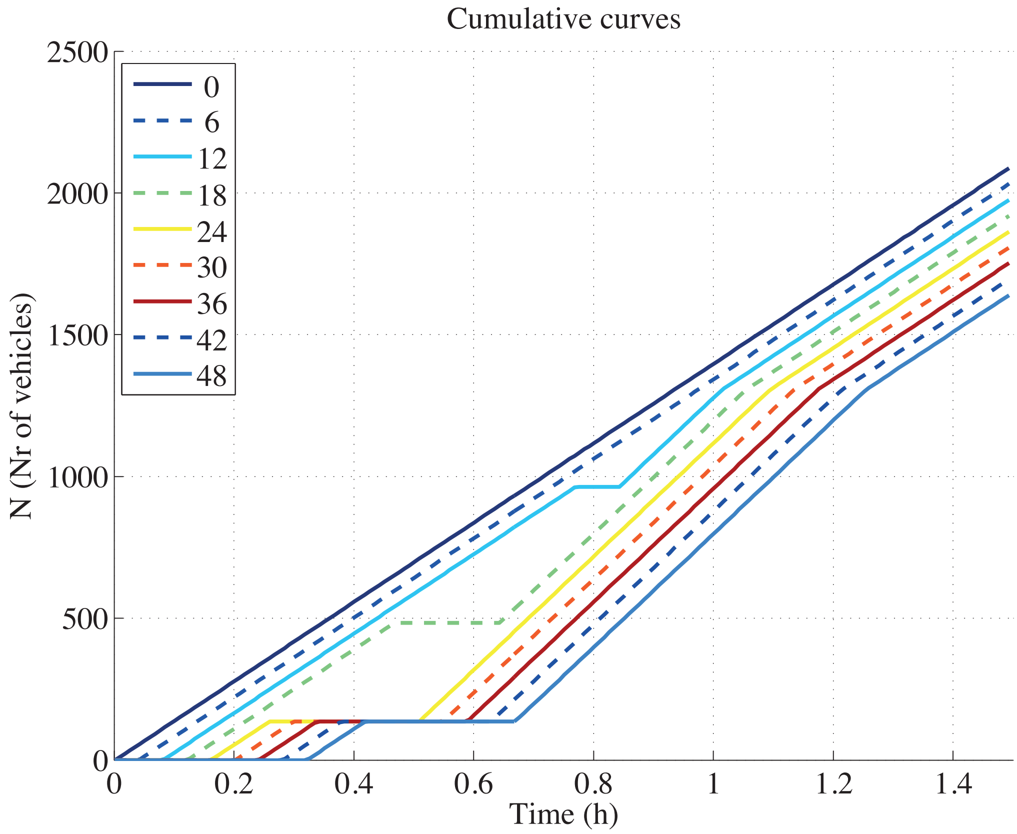
Exercise 97. *Draw the resulting traffic operations in a space time plot and in the fundamental diagram. Calculate all shock wave speeds. (4 points)*

A snow plow comes to free the vehicles that are stuck. After freeing the vehicles, the snow plows clear two of the three lanes of the motorway. Thereby, they drive at 5 km/h on the motorway. The capacity of vehicles passing the snow plough on the left lane is 2000 veh/h. The inflow on the road is 1000 veh/h.

Exercise 98. *Sketch the traffic operations in a space-time plot and in the fundamental diagram. Explain how you construct the graphs. (No point given here for the beginning, discussed in question b) (5 points)*

A-4-5 Cumulative curves

The graphs show cumulative curves for different locations along the road. The legends shows the distances in km from the beginning (i.e., upstream end) of a road. For the remainder of the question, reasoning is more important than exact readouts from the graph. When using graph readouts, please state so explicitly and *note the values you read from the graph.*



Exercise 99. *Explain the traffic state, mention a possible cause (e.g., "different speed limits for different sections", "peak hour jam") and explain why. (3 points)*

Exercise 100. *Estimate is the total delay encountered here (3 points)*

Exercise 101. *Sketch the traffic situation in space-time (shock waves – no trajectories needed).* Estimate the the location of changes in traffic states. *(3 points)*

Assume a triangular fundamental diagram.

Exercise 102. *Estimate, from the given curves, give the free speed, capacity, critical density and jam density (4 points)*

A-4-6 Crown jewels in the tower

We consider an exhibition with the most important piece an object in a small glass show case in the middle of the room. Visitors do not have a preference to see a particular side of the piece. The room is 15 meters long and 6 meters wide.

Exercise 103. *Explain why a larger glass show case can increase the capacity of the exhibition room, in terms of visitors per unit of time2*

The exhibitor has chosen for a show case of 3x1 meters.

Exercise 104. *Estimate the capacity of the room, measured in visitors per minute; base your answer on a reasonable watching time for the piece of art and state the assumptions explicitly4*

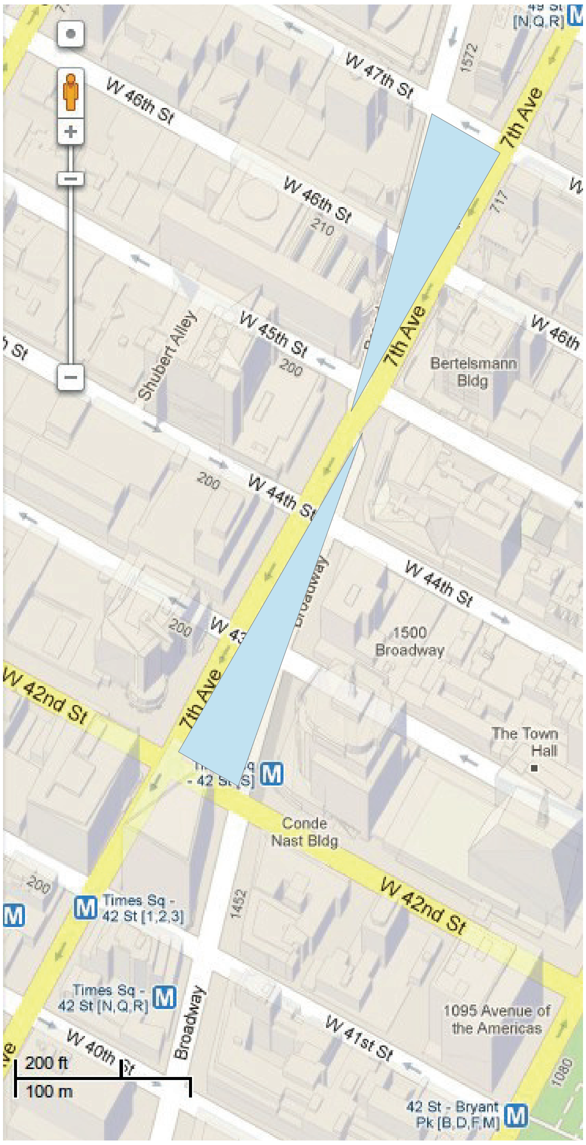
An alternative design is considered. Instead of the visitors walking by the art, the visitors can stand on a moving walkway (like in the airport), which is constructed at each side of the glass. Visitors step at the moving walkway at the beginning of the room, and step off at the end. They are not allowed to walk backwards on the moving walkway.

Exercise 105. *What is the capacity of the room in this case. Base your answer on (explicitly stated) reasonable assumptions on distance and speed.3*

A-5 Problem set 5

A-5-1 Short questions

Below, you find a google maps image of New York; Times Square is shaded in blue. The municipality of NY claimed that 1,000,000 people were at Times Square at the start of the new year (1 January 0h00).



Exercise 106. *From a traffic flow perspective, comment on this claim. (2 points)*

Exercise 107. *How can synchronized flow be recognized? (1 point)*

Exercise 108. *Sketch the three phases of traffic flow in a flow-density diagram, and give the names (2 points)*

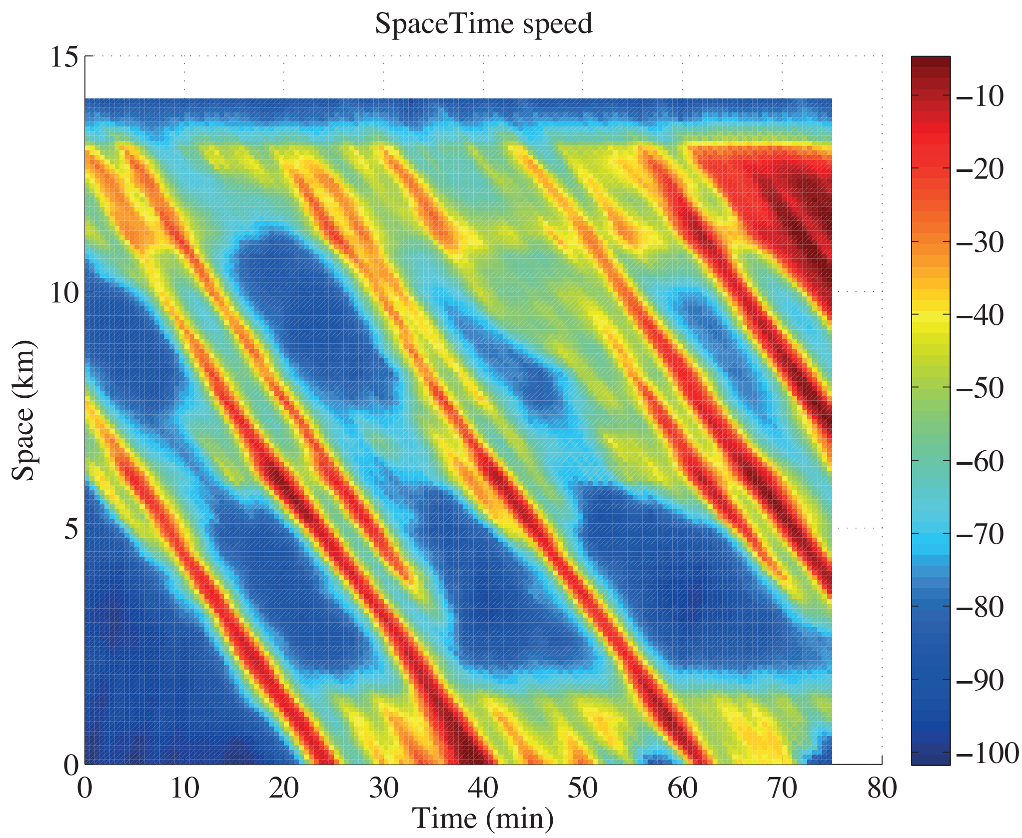
Exercise 109. *Comment on the concept of capacity in the three phase traffic flow theory (1 point)*

Exercise 110. *How are Lagragian coordinates for traffic defined (i.e., what is special on the the definition of the coordinates; how do they differ from Eulerian coordinates?) (1 point)*

Exercise 111. *What the main mathematical advantage of the Lagrangian coordinate system? (1 point)*

A-5-2 Motorway traffic operations

Below, you find a speed contour plot of traffic operations at a motorway (the negative speeds are plotted for color representation: units are in km/h, so blue is 100 km/h and red is (almost) standstill). (if you are color blind and cannot see the colors – please ask!)



Exercise 112. *Explain which direction traffic is flowing from the graphs1*

Exercise 113. *Name the two main traffic phenomena you see - and briefly describe what is happening. (2 points)*

Exercise 114. *What are the three different levels of stability, and what do they mean (2 points)*

Exercise 115. *Comment on the stability of traffic for this day (3 points)*

Exercise 116. *What traffic states can occur according to the theory of slugs and rabbits. Give the names and a short description (10-20 words per state) (2 points)*

Exercise 117. *Modelling a multi lane Dutch motorway, how does this theory in resulting traffic differ from the Mobil lane change model for moderately dense (undercritical) traffic? (2 points)*

A-5-3 Car-following

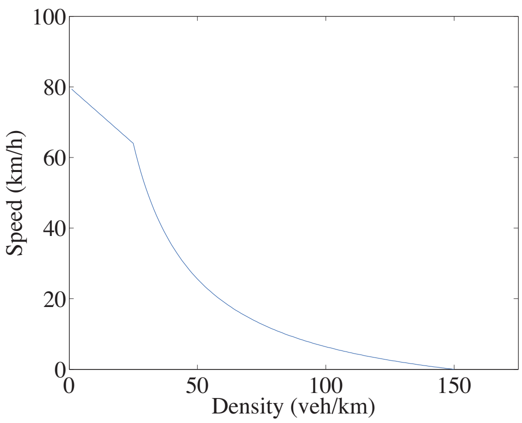
A fundamental diagram only holds in stationary and homogeneous traffic conditions.

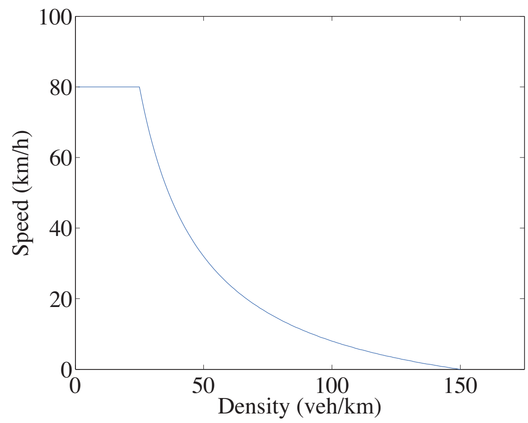
Exercise 118. *Explain the terms stationary and homogeneous (indication: 20 words each) (2 points)*

From a car-following model the fundamental diagram can be derived.

Exercise 119. *Can a car-following model be derived from the fundamental diagram? If so, explain how. If not, explain why not. (2 points)*

The two speed-density relationships shown below relate to a triangular fundamental diagram and Smulders fundamental diagram





Exercise 120. *Which figure is which fundamental diagram. Why? (1 point)*

Exercise 121. *Draw the corresponding flow-density and speed-flow curves for Smulders fundamental diagram (in your drawing, put attention to (or indicate) the values or slopes at several critical points at the ends of lines or at turning points)3*

A-5-4 Partial road blocking near a signalised intersection

Upstream of a signalised intersection the road is partially blocked by a crane (see below figure for illustration). Assume that between the crane and the intersection, all four lanes can (and will) be used if needed. *That means that the length between the crane and intersection is longer than in the images.*

During the red phase, the tail of the queue spills back further than the crane. During the green phase, the queue dissipates completely. Throughout this question, assume a fundamental diagram which is concave (not triangular).

Exercise 122. *Sketch the fundamental diagrams. Give typical values for the characterizing points. (2 points)*

Use this diagram throughout this question.

Exercise 123. *Sketch the traffic operations in a space time plot, starting at the moment the traffic light turns red and ending at the moment the traffic light turns red again. Also indicate all states in the fundamental diagram. Only consider traffic states upstream of the traffic light.8*

Exercise 124. *Draw the (moved) cumulative curves upstream of the traffic jam and downstream of the traffic light (2 points)*

Exercise 125. *How could you get the total delay from this for all vehicles together (explain, but no need for calculations) (1 point)*

Exercise 126. *Comment qualitatively on the effect of the distance between the intersection and the crane on the delay (2 points)*

The workers move the crane downstream in order to reach a building further downstream. They can only do so when there is no queue downstream. Assume they start when the queue is dissolving and traffic flows past the crane, but the queue is not dissolved completely yet, and assume a speed for the movement. The fundamental diagram does not change due to the movement (i.e., you may use the FDs of question a)

Exercise 127. *Sketch the traffic states in the flow-density plane, from the moment the crane moves until the queue is completely dissipated (you may assume the traffic light stays green, and assume the distance to the traffic light is large enough that the crane can keep moving until the traffic jam is solved). (4 points)*

A-5-5 Network Fundamental Diagram

Exercise 128. *What variables are related to each other in the Network Fundamental Diagram. Give the names and explain the terms (more than one answer possible) (2 points)*

Exercise 129. *Sketch a typical NFD (1 point)*

Exercise 130. *How is this shape different from a regular fundamental diagram. Explain briefly why (20 words). (1 point) –*

Exercise 131. *Comment on the traffic mechanism behind the right hand side. Show in your answer that you understand the difference in traffic mechanism with the fundamental diagram (1 point)*

A traffic engineer suggests to limit the access to the central business district, a part of the road network with many companies.

Exercise 132. *Explain using the NFD why this can be beneficial even though the stopped vehicles need to wait and experience delay. (3 points)*

A-6 Problem set 6

Exercise 133. *Explain what a car-following model does in general (1 point)*

A-6-1 Short questions

The pce value of a truck is claimed to be 2 for a road without queue warning systems.

Exercise 134. *Explain what a pce value is (indication: 30 words). (1 point)*

A study shows that by installing queue tail warning systems the capacity of the road increases. The factor by which it increases is 1.05 for passenger cars and 1.09 for trucks, meaning it can handle 5% more cars and 9% more trucks

Exercise 135. *Based on the information above, calculate the new pce value for trucks for the road with queue warning systems. (3 points)*

Exercise 136. *Traffic simulation in Lagrangian coordinates has the advantage that there is less numerical diffusion. Explain in your own words what this means – indication 50 words. (1 point)*

The total utility U for the mobil lane change model is given by:

$$U_{\text{tot}} = U_{\text{own vehicle}} + p \sum_{i \in \text{other drivers}} U_i \tag{A-9}$$

In this equation, the utility for each driver is determined by its instanteneous acceleration, given by a car following model. This utility can be calculated for the situation with lane change ($U^{(\text{lane change})}$), as well as without lane change, ($U^{(\text{no lane change})}$). For European driving, vehicle will change to the right if $U^{(\text{change to the right})} > (U^{(\text{no lane change})} - a_{th}$ and to the left if $U^{(\text{change to the left})} > (U^{(\text{no lane change})} + a_{th}$. In these equations, $a_t h$ is a threshold which should be exceeded to perform a lane change.

Exercise 137. *What is a reasonable value for a_{th} – explain why. (1 point)*

Exercise 138. *What would be a reasonable range for the parameter p – explain why. (1 point)*

Consider a fundamental diagram for pedestrian flow in the flow-density plane.

Exercise 139. *What are the units, as well as the minimum and maximum values at the axes? (2 points)*

A-6-2 Bridge opening

Consider a two lane motorway, with a capacity of 5000 veh/h and a queue discharge rate of 4500 veh/h. On this motorway, the free flow speed differs from the critical speed.

Exercise 140. *Sketch a fundamental diagram for the motorway in the flow-speed (!) plane.2*

Exercise 141. *Indicate reasonable values for the (i) free flow speed, (ii) critical speed, (iii) critical density and (iv) jam density.3*

Assume the fundamental diagram holds. The demand is constant at 3500 veh/h. The bridge opens at t=1h and is opened for 15 minutes.

Exercise 142. *Sketch the cumulative curves for the inflow and outflow.2*

Exercise 143. *Calculate the maximum number of vehicles in the queue.1*

Exercise 144. *Calculate the total delay of all vehicles combined.3*

In the remainder of the question, assume a Greenshields fundamental diagram with a capacity of 5000 veh/h, and a critical density of 60 veh/km.

A further detailed description of the traffic situation uses a so called *acceleration fan*.

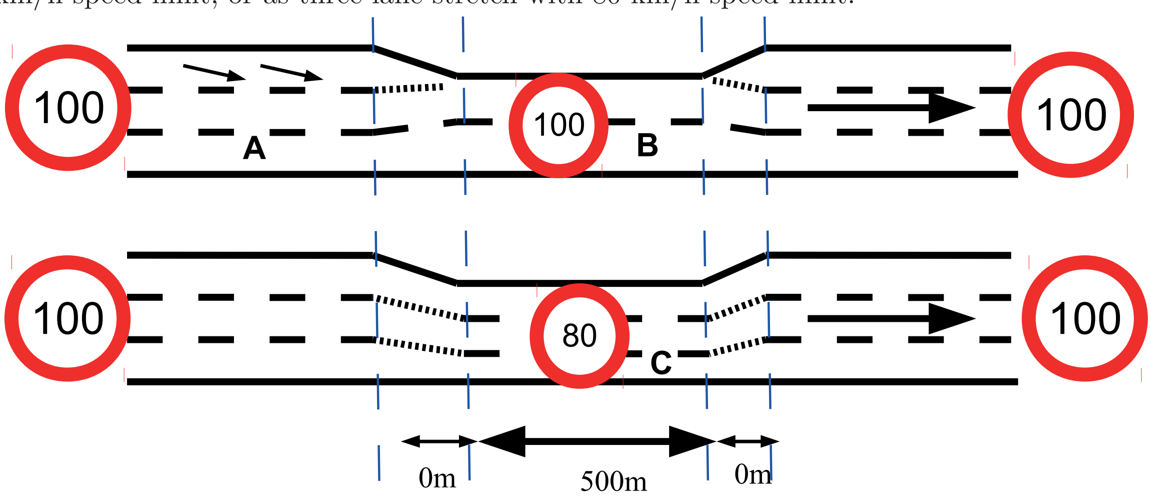
Exercise 145. *Explain in your own words what is an acceleration fan.1*

Exercise 146. *Under this assumption, sketch the flow 500 meters downstream of the bridge as function of time from the moment of the bridge closing (i.e., when the vehicles can flow again); also provide a reference value at the vertical axis.3*

Exercise 147. *Calculate when the flow starts to change in the graph above (hard question: 2 instead of 4 points).2*

A-6-3 Variable road layout

Consider a 3 lane motorway with a 100 km/h speed limit. The road has a narrower part of a limited length (a bridge, 500 meters long), which can be used as two lane stretch with 100 km/h speed limit, or as three lane stretch with 80 km/h speed limit.



Throughout the question you may assume a triangular fundamental diagram for all road types. For the wide lanes, the critical density is 25 veh/km/lane and the jam density is 125 veh/km/lane. Assume the width of the lanes does not influence the congested branch.

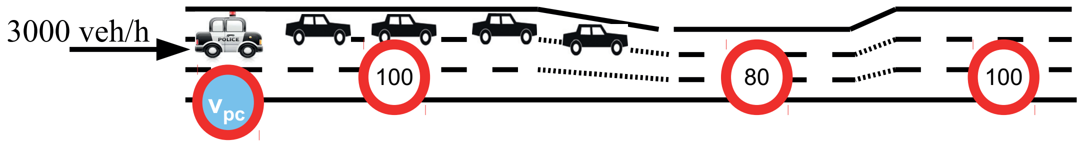
Exercise 148. *Sketch the fundamental diagrams (aggregated over all lanes) for the three parts (all in one figure): three wide lanes, two wide lanes and three narrow lanes: parts A, B and C in the figure above. (2 points)*

Exercise 149. *Calculate the capacity of the three lane narrow road section. (2 points)*

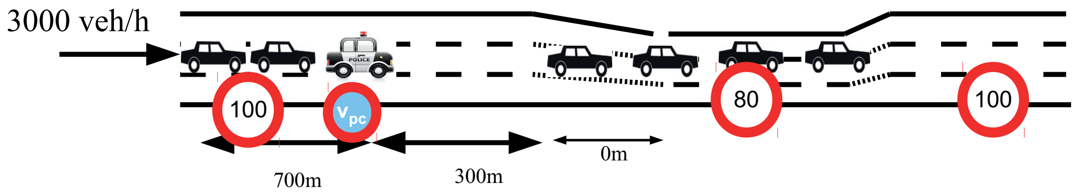
The road authority switches dynamically between the three narrow lanes and the two wide lanes.

Exercise 150. *Explain why this dynamic switching is beneficial for the delays (i.e., why does this give less delays than choosing one layout permanently) (indication: 50 words). (1 point)*

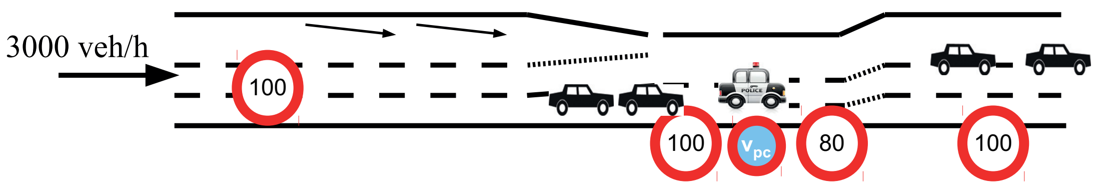
During switching, the road needs to have an empty stretch of 300 meters (because one cannot change the lanes at which vehicles are driving). The road authority wants to do so by insert a *pace car* into the traffic stream starting 1000m upstream of the switching part. This should drive at an appropriate speed (v_{pc}) to create this gap (the gap should be downstream of the pace car after this has travelled 700m). Consider the situation of switching from 3 lanes to 2 lanes. The demand is 3000 veh/h. For illustration purposes, the process is also shown in figures below.



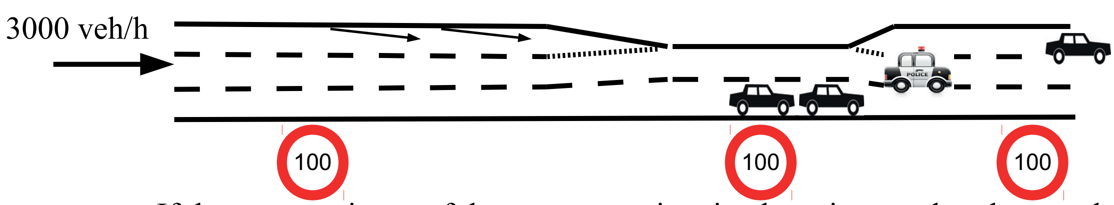
In normal conditions a pace car is introduced, driving at a speed v_{pc}



After 700 m driving, the pace car has a gap in front of 300 m. The last vehicle of the platoon downstream of the pace car is 300 downstream of the pace car, at the location of the change of layout



Upstream of the pace car, the road switches to a 2 lane road with a 100 km/h speed limit



If the pace car is out of the narrow section, it adapts its speed to the speed limit if not limited by traffic. Otherwise, it adapts it speed to the vehicles in front.

Exercise 151. *Calculate the speed at which the pace car needs to drive. (1 point)*

The pace car acts hence as moving bottleneck, and will continue driving at this speed throughout the bridge section, and then (instanteneously) speed up to the free speed.

Exercise 152. *Indicate the traffic states in the flow density diagram and sketch traffic states in the space-time diagram (and relate them to each other using shock wave theory). (no vehicle trajectories needed) (hard question: 6 points awarded for 9.5 points). Tip for sketching: choose the demand low in your graph (for simplicity you may assume other flows to be higher). If you did not find an answer for the question above, assume it is driving 60 km/h. (6 points)*

A-6-4 Traffic stability

Exercise 153. *Give the name of the mild congestion located at a bottleneck location in Kerner's three phase traffic flow theory (1 point)*

In the flow-density plan, this is an area rather than a line.

Exercise 154. *Link this to the Wiedeman car-following principle. (indication: 100 words) (2 points)*

A-6-5 Levels of description

One can argue that there are three levels at which traffic can be described: microscopic, macroscopic, and network-level. On the microscopic level, traffic behaviour can be indicated by a spacing-speed diagram.

Exercise 155. *Sketch this diagram. (2 points)*

Exercise 156. *Sketch the matching diagram in the flow-density plane; indicate how the values for this graph can be obtained from the previous graph.(2 points)*

Exercise 157. *Give is the name for the flow aggregated on network level (1 point)*

Exercise 158. *Give the name for the density aggregated on network level (1 point)*

Exercise 159. *Sketch the matching Network Fundamental Diagram for the network level in the same graph as the fundamental diagram of question b, and briefly comment on the difference (50 words).2*

A-7 Problem set 7

A-7-1 Short open questions

Upstream of an junction is a controlled intersection.

Exercise 160. *Which distribution is likely to describe the number of arrivals per time interval (i.e., flow) on the main road near this junction (and downstream of the controlled intersection)? Name and explain briefly why (no need to comment on the equations of the distribution functions). (2 points)*

Exercise 161. *Argue why traffic instability can only occur under platoon instability. (2 points)*

Traffic dynamics can be described by linking three variables, usually indicated by x, N, and t.

Exercise 162. *Describe the physical meaning of N(x,t). (1 point)*

Exercise 163. *Describe the difference between calibration and validation of a traffic model. (Indication: 40 words) (2 points)*

Exercise 164. *Which relationship is described by the macroscopic fundamental diagram. In your answer, make sure to be specific enough that the answer does not apply for the fundamental diagram. (There is no need to comment on the shape, traffic process leading to the shape or to name specific terms – indication: 25 words) (1 point)*

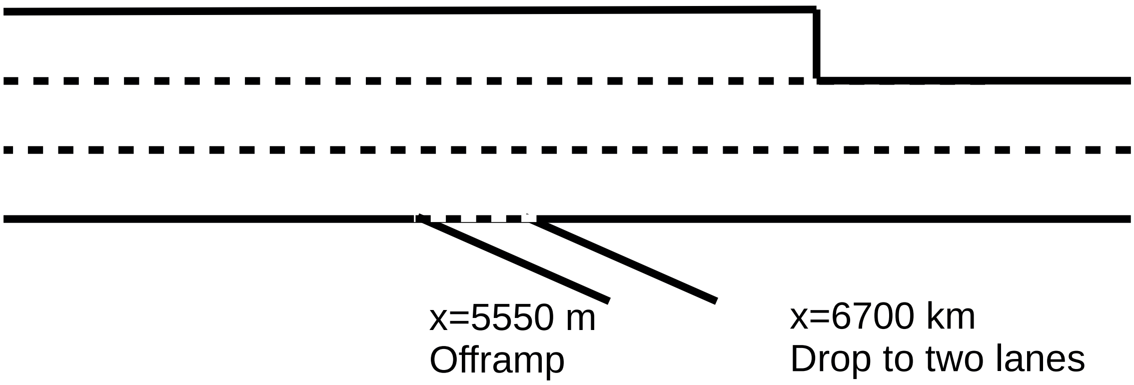
For pedestrian traffic, a typical value for the capacity which can be found in a handbook is 1.22.

Exercise 165. *What are the units for this capacity. (1 point)*

Exercise 166. *What is the physical meaning of these units? (1 point)*

A-7-2 Cumulative curves

Consider the following road layout:

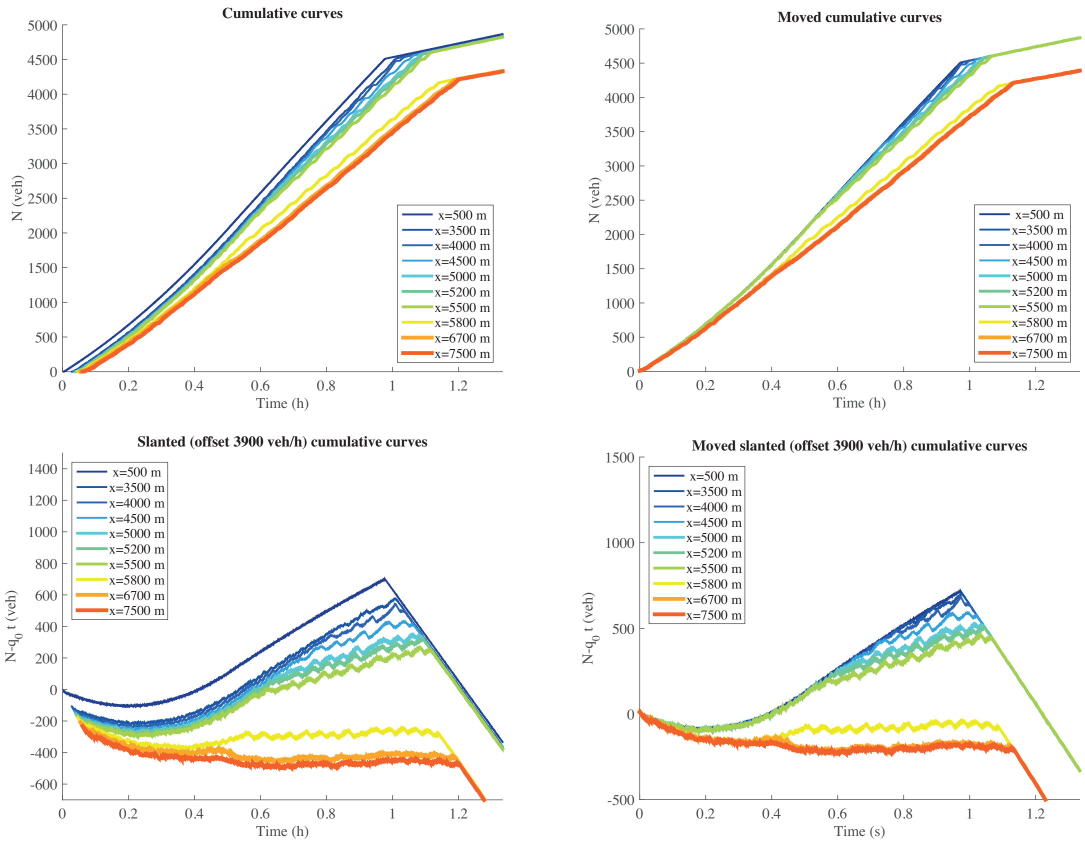


A three lane road has an offramp just downstream of 5500m, and a lane drop at approximately 6700m.

Cumulative counts are obtained from Fosim.

Exercise 167. *Name the two main limitations of using cumulative curves for delay determination in real world conditions. (2 points)*

On the next page, you find the cumulative curves, the moved cumulative curves (i.e., corrected for the free flow travel time), the slanted cumulative curves (offset 3900 veh/h), and the slanted moved cumulative curve (i.e., corrected for the free flow travel time). (If you cannot differentiate colors, raise your hand to get explanation on the lines!)



Exercise 168. *Describe the traffic operations (or sketch the traffic operations in space-time): This should show: is there an temporal bottleneck? Why does congestion end (demand or supply)? Where is the head and tail of the queue approximately. Also indicate what information you use to get to this conclusions. (3 points)*

Exercise 169. *Estimate the (free flow) capacity of the bottleneck in veh/h from the curves. (2 points)*

Exercise 170. *Estimate the capacity drop in veh/h from the curves. (2 points)*

The total delay due to the queuing is determined by summing the delays for all vehicles.

Exercise 171. *Argue whether the delay would differ if a vertical queuing model for the bottleneck was used, and if so, would the delay be less or more with a vertical queuing model. (2 points)*

A-7-3 Macroscopic traffic variables in case of different vehicle classes

Suppose there are two classes of vehicles, fast vehicles driving 120 km/h and slow vehicles moving at 90 km/h. At the entrance to the road (x=0), an equal number of each type of vehicles is measured under a total flow of 1000 veh/h (so 500 veh/h each). *For some subquestions you might need answers of earlier subquestions, which you could not answer. In that case, assume an answer, state so on your answer sheet and continue with the assumed answers.*

Exercise 172. *Calculate the density on the road. Base your answer on the density of the slow vehicles and the fast vehicles. (3 points)*

Exercise 173. *Calculate the space mean speed. (2 points)*

Exercise 174. *Calculate the time mean speed. (2 points)*

Exercise 175. *Give the equation for Edie's generalised definition of density. (1 point)*

There is an observer moving at 60 km/h.

Exercise 176. *Calculate the flow of fast moving vehicles passing this observer. (3 points)*

Exercise 177. *What is the time mean speed (absolute to the road, not relative to the observer) of the vehicles moving this observer (hard question: 3.5 points, yielding 2 points) (2 points)*

Exercise 178. *Will the flow of slow moving vehicles passing the moving observer be larger than, equal to or lower than the flow of the fast moving vehicles passing the moving observer. Argue (rather than calculate) why. (2 points)*

The capacity of the road is, independent of the vehicle class, 2000 veh/h, and the jam density is, independent of the vehicle class, 125 veh/km. Assume the car-following behavior of both classes can be described by Newell's car-following model.

Exercise 179. *Draw (not sketch) the fundamental diagrams for the two vehicle classes (2 points)*

Exercise 180. *If we take the fast vehicles as reference, what is the pce value (or fce, "fast car equivalent") of the slow vehicle (so expressed in units of fast vehicles) (1 point)*

Exercise 181. *Give the car-following model equation for the fast vehicle, including the values of the parameters for the fast vehicles (only the free flow branch is needed). (3 points)*

Consider these traffic operations at a two lane road. According to a lane selection model, all fast vehicles stay left and all slow vehicles stay right.

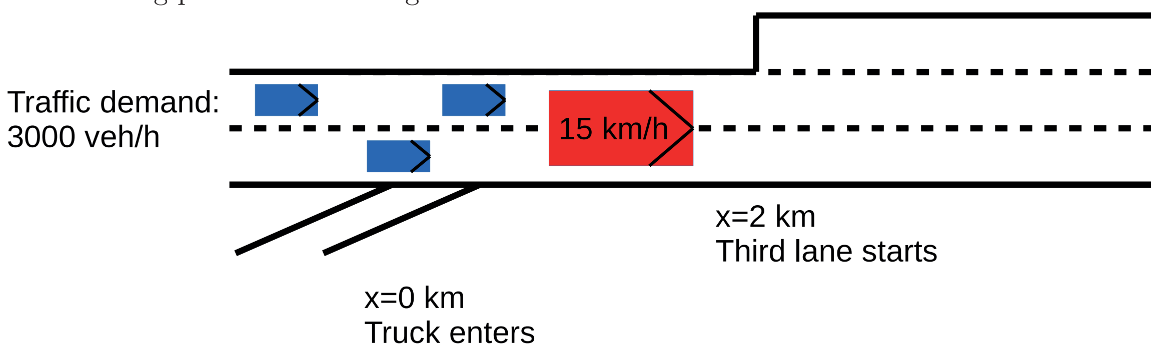
Exercise 182. *What is the name of this model (only the name is needed – no reasoning). (1 point)*

Now consider a different traffic state. Once traffic goes into the congested state, the fast cars move to the fastest lane.

Exercise 183. *Give the name of this state called according to* three phase traffic flow theory*.* *(1 point)*

A-7-4 Slow truck on the motorway

Consider a road for which on all of the lanes a triangular fundamental diagram holds with a critical density of 25 veh/km, a capacity of 2000 veh/h and a jam density of 125 veh/km. There is a two-lane road for a length of x<2km, and with an extra third lane for x>2km. The traffic demand is 3000 veh/h. At t=0 a wide truck (special transport) joins the road from x=0, driving at 15 km/h. This truck blocks two lanes, so in the two lane section, there are no overtaking possibilities: see figure.



Exercise 184. *Sketch the traffic operations (both in space-time and in flow-density) for the time up to the moment the truck reaches the three lane section. (5 points)*

After the truck reaches the downstream section, vehicles can overtake. The capacity point lies at k=60veh/km and q=2600 veh/h for the whole roadway.

Exercise 185. *Sketch the traffic operations (both in space-time as in flow-density) from the moment the truck reaches the three lane section. Sketch the slopes of the waves which occur from that moment, but* do not consider possible waves which start later.*(5 points)*

A-8 Problem set 8

A-8-1 Short open questions

Exercise 186. *Give the name of the distribution describing the headway distribution if the arrival process of vehicles is independent. (1 point)*

A particular distribution function describes the number of vehicles per aggregation interval well if the flow on the road is high.

Exercise 187. *Give the name of this distribution. (1 point)*

The lane flow distribution is the ratio of the flow in a lane over the total flow.

Exercise 188. *Express the relative flow in lane i f_i as function of the average headways $< h_i >$, which are given for all lanes i on a motorway (3 points)*

The theory of slugs and rabbits states there are two types of cars/drivers on the road. Assume they both have a pce value of 1.

Exercise 189. *Give the meaning of the letters pce (give the full words). (1 point)*

Consider the theory of slugs and rabbits. The demand (in veh/h) of slugs is equal to the demand of rabbits. Assume a two-lane road operating below critical density at all lanes.

Exercise 190. *What is the distribution of the flow over the lanes in free flow conditions (qualitatively – no numbers are needed)? Give your reasoning. (2 points)*

Exercise 191. *Explain how stop-and-go waves emerge. Use, name, and describe the two most relevant levels of stability in your answer. (3 points)*

Exercise 192. *Explain why in the three phase theory of Kerner there is not one value for capacity. (1 point)*

One of the advantages of using Lagrangian coordinates is that traffic predictions do not have "numerical diffusion".

Exercise 193. *Explain in your own words what "numerical diffusion" is. (1 point)*

A-8-2 Microscopic effects on the fundamental diagram

Exercise 194. *Give the conditions under which the fundamental diagram holds. (1 point)*

When aggregating traffic, a better fundamental diagram can be found if instead of a rectangular area in space-time a parallelogram is used.

Exercise 195. *Explain why, discussing free flow conditions as well as congested conditions. In your answer, also indicate the slopes of the edges of the parallelogram. (3 points)*

Traditionally, the density is calculated at one moment in time over a road section, and the flow is calculated over a time period at one location.

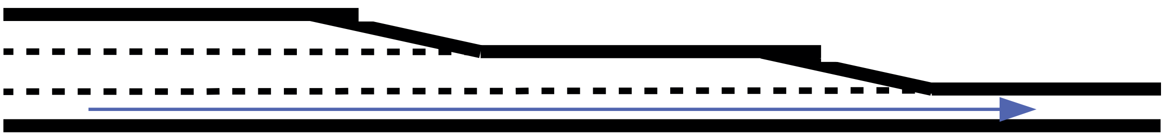
Exercise 196. *Give the name for the definitions used to express flow and density for an area in space-time. Equations are not needed. (1 point)*

During acceleration, one could observe deviations from the fundamental diagram. Consider multi-anticipation, i.e. drivers adapting their speed not only on the leader, but also on vehicles further downstream. This influences amongst others transitions from a congested state to capacity.

Exercise 197. *Describe for this transition the effect of multi-anticipation on hysteresis and sketch the effect of these changes in the flow-density plane (also show the normal FD in this graph as reference).4*

A-8-3 Prediction of downstream conditions

The number of lanes on a freeway reduces from 3 to 2 to 1, as sketched below.



The one lane section forms an active bottleneck. Assume throughout this question a triangular fundamental diagram, equal for each of the lanes, with a free flow speed of 100 km/h.

In the three lane part, a driver is driving in congestion due to the reduction to one lane. His speed is 8 km/h and he estimates the gross spacing to his predecessor in the same lane at 11m. *(If you are at one part unable to find the correct answer, assume one, state so clearly, and continue with the next part of the question).*

Exercise 198. *Assuming homogeneous conditions, calculate the flow through the bottleneck. (4 points)*

Exercise 199. *Calculate the jam density for the three lane part. Hint: use the capacity of the three lane part. (4 points)*

Exercise 200. *Calculate the speed of the vehicles in the two lane section. (3 points)*

Changes of inflow can lead to a changing traffic state. Adapting inflow can then make the traffic state stationary (i.e. constant over time). Consider one of these stationary situations, with congestion spilling back to somewhere in the three lane section

Exercise 201. *Sketch the flow (aggregated over all lanes) as function of space for this situation. (1 point)*

Consider all possible stationary situations, which have a different accumulation (i.e., nr of vehicles in the area).

Exercise 202. *Sketch the traffic production (here, average internal flow in veh/h) as function of the accumulation. Values are not needed in the sketch. (2 points)*

A-8-4 Demonstration of police cars

The police are planning a demonstration by which they drive 30 km/h on all lanes of the motorway. In practice, the police cars travel as normal vehicles in the traffic stream before the demonstration. And at t=0, they slow down instantaneously to 30 km/h. During the demonstration (the slow driving), there are no overtaking possibilities. In this question, the traffic state upstream of the police cars is referred to as "congestion". See the figure below for an idea.

In this question we consider the following variables:

1. The duration of the demonstration d
2. the demand (assumed constant throughout the demonstration), q_{in}.

After the demonstration, the police cars instantaneously accelerate to the free flow speed.

Assume a triangular fundamental diagram with a free flow speed of 80 km/h, a capacity of 2000 veh/h/lane (indicated C) and a jam density of 150 veh/h/lane. Consider a two lane road with no on and off ramps.

Exercise 203. *Sketch the traffic states in space-time and in the flow density plane for a high demand (demand is higher than the flow in congestion). (6 points)*

Exercise 204. *Calculate the flow in the congestion. (3 points)*

Exercise 205. *Sketch the moved cumulative curves for the inflow and the moved cumulative curves for the outflow with the demonstration. Consider the inflow point upstream of any queue, and the outflow point downstream of any queue. Consider time frame from well before the demonstration to well after any queue. (4 points)*

Exercise 206. *Show that the total delay (D) is given by the equation* $D = (1/2) * d^2 * q_{in} * C/(C - q_{in})$. *(4 points)*

Exercise 207. *What happens to the delay if the demand approaches capacity? Explain from a traffic flow perspective why. (3 points)*

Exercise 208. *Sketch the traffic states in space-time and in the flow density plane for a demand lower than the flow in congestion (2 points)*

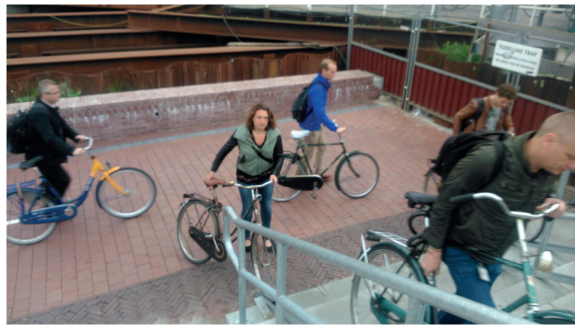
Exercise 209. *Argue whether the equation for the delay still holds in this case (1 point)*

Exercise 210. *Give the reasoning why the delay is quadratic as function of duration of the demonstration (3 points)*

A-8-5 Stairs for cyclists at Delft train station

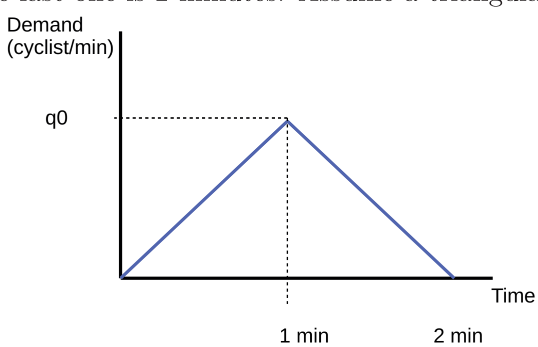
Currently, the station in Delft has an underground bike parking. To get to the street level, the cyclists need to dismount, and walk their bicycle up the stairs. See the picture below.

Assume all cyclists do so with the bike at their right. The figures below give an impression of the site, although not at peak demand



Exercise 211. *Calculate the capacity for this site based on an assumption of headway between people walking their bike. (3 points)*

In the peak hour, a train arrives. Suppose 300 passengers leave the train station by bike using these stairs. Suppose the difference between the first passenger leaving the bike parking and the last one is 2 minutes. Assume a triangular demand pattern at the bottom of the stairs:



Exercise 212. *Calculate the value for q0 in this pattern. (3 points)*

Assume the capacity C is lower than q0. Number the cyclists 1...N in order of leaving the bike parking. The cyclist with the highest delay due to queuing is N_E.

Exercise 213. *Derive the equation for N_E as function of C and q0. Give your reasoning. (If you are unable to give the equation, points can be awarded for a description of the construction.) (4 points)*

A-9 Problem set 9

A-9-1 Short questions

The Macroscopic Fundamental Diagram relates the average density to the average flow.

Exercise 214. *Give the name for the term used to indicate the average flow. (1 point)*

Exercise 215. *On a particular road the Greenshields fundamental diagram holds, with a free flow speed of 80 km/h and a jam density of 100 veh/km. Does this determine the capacity with no further assumptions? If so, calculate. If not, why not. (3 points)*

Exercise 216. *Describe what an acceleration fan is. (2 points)*

On a road, traffic is measured using loop detectors at a cross section. For each aggregation period the flow is given as output, as well as the time mean speed of traffic passing that detector. Then, densities are computed from the speed and flow, and a relation between speed and flow is plotted.

Exercise 217. *Explain whether there is a bias in the densities; if yes, explain which and why; if no, explain why not. (3 points)*

Exercise 218. *Traffic in a model is locally stable. Can the platoons be unstable? Explain why or why not. (2 points)*

Exercise 219. *Explain when and why "random seeds" are needed (approximately 100 words, no equations needed). (2 points)*

A-9-2 Car-following in fog

Helly's car-following model is given by the following equations:

$$a = \alpha \Delta v + \beta (s - s^*) \tag{A-10}$$

$$s^* = x_0 + vT \tag{A-11}$$

In this equation, a is the acceleration, v denotes the speed, Δv the difference in speed between the leader and its predecessor, s the spacing. Parameters in the model are α, β, x_0 and T. The interpretation of the parameters, and the meaning of s^* is left to the students. For reasons of simplicity, the time dependence and reaction time are ignored in this question.

Exercise 220. *Derive the equation of the fundamental diagram, q=q(k), from the above equation* for the congested branch. *Hint: use the relation v=v(s) in the derivation.4*

The model can be calibrated and validated.

Exercise 221. *Explain what is the value of validating the model (i.e., what can a validated model be used for, which cannot be done with a non-validated model).(1 point)*

The following parameters are found to fit the traffic best:
$\alpha = 0.3\text{s}^{-1}, \beta = 0.08\text{s}^{-2}, x_0 = 20\text{m}, T = 1\text{s}$. (In this line, s means second and m meter). Besides the car-following part, also a free speed is required to derive a full fundamental diagram. This is found to be 30 m/s.

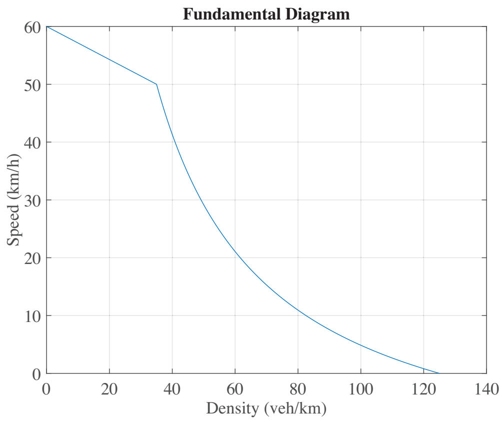
Exercise 222. *Compute the capacity of the road in veh/h* (hard question)*(2 points)*

The Dutch government recommends drivers in fog: "Half the speed, double the space headway to the predecessor". Suppose everyone takes this advice.

Exercise 223. *Compute the capacity of the road in the fog. You may (but do not need to) relate this to question 2c; in case you did not find an answer to that question, assume an answer and state so explicitly.(3 points)*

A-9-3 Cleaning the road

Consider a one-lane road for which the following fundamental diagram holds:



This is a well-known fundamental diagram.

Exercise 224. *Give the name of the shape of this fundamental diagram.(1 point)*

Exercise 225. *Compute the capacity of the road.(2 points)*

Exercise 226. *Sketch the fundamental diagram in the flow-density plane.(1 point)*

A cleaning car enters the road to clean the road, driving at 15 km/h.

Exercise 227. *Compute the maximum flow upstream of the cleaning car using the given fundamental diagram.(2 points)*

The inflow to the road is 1200 veh/h; the cleaning car enters the road at x=1km and leaves at x=2 km. Overtaking is not possible.

Exercise 228. *Construct the traffic state using shockwave theory: i.e., sketch the traffic operations in the flow-density plan and in space-time(6 points)*

A-9-4 Mobil lane changing model

The Mobil lane change model describes whether drivers change lanes or not. The equations for a European system (keep right unless overtaking) are given by the following equations. The utility for a driver is given by:

$$U_{\text{tot}} = U_c + \mathcal{P}\left(U_o + U_n\right) \tag{A-12}$$

The utility is calculated for each lane. Acceleration is considered as utility, and subscripts indicate the current vehicle (c), the old follower (o) and the (potential) new follower (n). A lane change to the right is performed if

$$U_{\text{tot}}^{\text{right}} - U_{\text{tot}}^{\text{left}} \geq a_{\text{th}} - a_{\text{bias}} \tag{A-13}$$

A lane change to the left is performed if

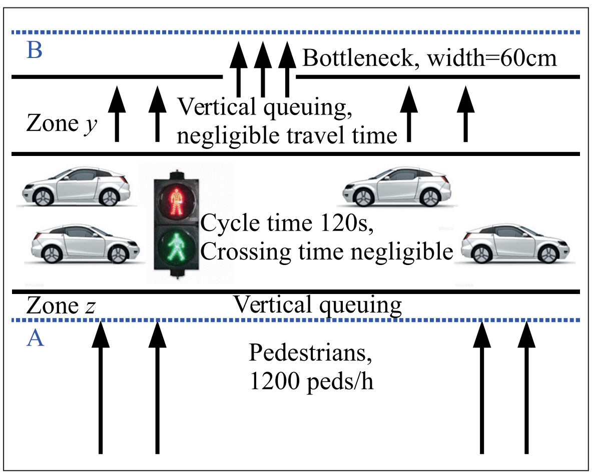
$$U_{\text{tot}}^{\text{left}} - U_{\text{tot}}^{\text{right}} \geq a_{\text{th}} + a_{\text{bias}} \tag{A-14}$$

Exercise 229. *Explain the working of the Mobil lane change model in words (approx 100 words); in your answer, include an interpretation for $\mathcal{P}$. (3 points)*

Exercise 230. *Give approximate values (including units) for a_{th} and a_{bias}, and reason why (3 points)*

A-9-5 Pedestrians at a traffic light

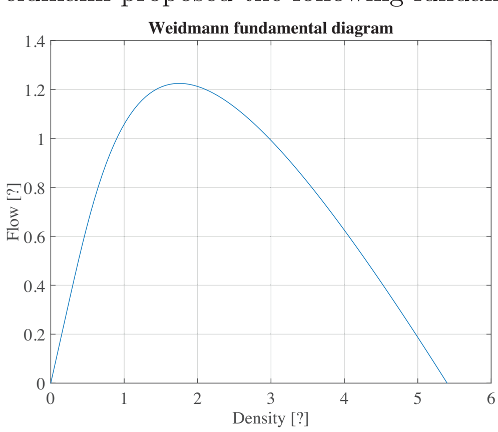
Consider pedestrians approaching the traffic light at a busy intersection. The pedestrians arrive independently at the traffic light at a flow of 1200 peds/h. The cycle time is 120 seconds. Assume all pedestrians cross the road simultaneously (next to each other) in negligible time and then the traffic light turns red again. (So the pedestrian traffic light is red for 120 seconds, then all pedestrians cross instantaneously and the traffic light turns red again instantaneously). A sketch of the situation is found below:



Exercise 231. *Give the name of the distribution of the number of pedestrians arriving per red time. (1 point)*

Exercise 232. *How many pedestrians cross during a green phase. (2 points)*

Weidmann proposed the following fundamental diagram for pedestrians.



Exercise 233. *Give the units for the values at the axes. (2 points)*

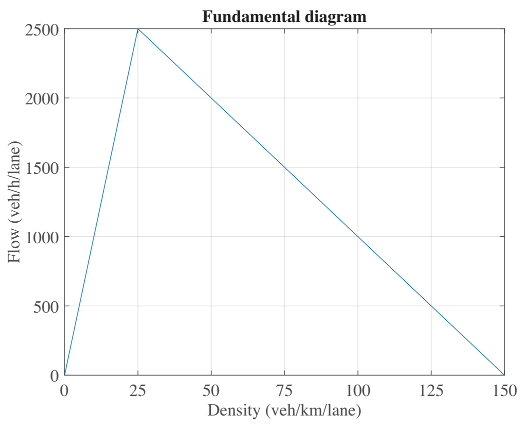
The pedestrians cross (simultaneously) during the green phase and arrive at the other side. After a short (negligible) distance, the sidewalk has a bottleneck of 60 cm wide. Assume that the Weidmann fundamental diagram holds. You may assume this bottleneck has no length (e.g., a door on the sidewalk).

Exercise 234. *Calculate how long would it take the pedestrians to pass the bottleneck. (3 points)*

Exercise 235. *Sketch the moved cumulative curves for this situation (inflow at the blue dashed line "A", outflow at the blue dashed line "B") for a period of one cycle. Start your sketch just before the moment the traffic light is green, i.e. the pedestrians are about to cross the road. Hint: at that moment, how many pedestrians have then crossed line A and not line B? (3 points)*

A-9-6 Hysteresis

On a road the following fundamental diagram holds under equilibrium conditions.



However, due to the delayed reaction of drivers, hysteresis occurs. Copy the fundamental diagram to your answer sheet.

Exercise 236. *In the fundamental diagram, sketch which path the traffic states make in the transition from congestion to capacity and back. Argue why. (2 points)*

Exercise 237. *Draw (not sketch) the fundamental diagram in the speed-spacing plane. (3 points)*

Exercise 238. *Sketch the hysteresis of question a in the fundamental diagram in the speed-spacing plane. (1 point)*

A-10 Problem set 10

A-10-1 Short questions

Consider a road where vehicles are arriving independently.

Exercise 239. *Give the name for the distribution of the number of vehicles arriving within a time interval on this road. (1 point)*

The pce value of a truck is claimed to be 2 for a road without queue warning systems.

Exercise 240. *Explain in your own words what a pce value of two means. (1 point)*

A study shows that by installing queue tail warning systems the capacity of the road increases. The factor by which it increases is 1.05 for passenger cars and 1.09 for trucks, meaning it can handle 5% more or and 9% more trucks.

Exercise 241. *Based on the information above, calculate the new pce value for trucks for the road with queue warning systems (3 points)*

Traffic simulation in Lagrangian coordinates has the advantage that there is less numerical diffusion.

Exercise 242. *Explain in your own words what numerical diffusion in traffic simulation is. (1 point)*

Exercise 243. *Explain what is meant by local instability in traffic. (1 point)*

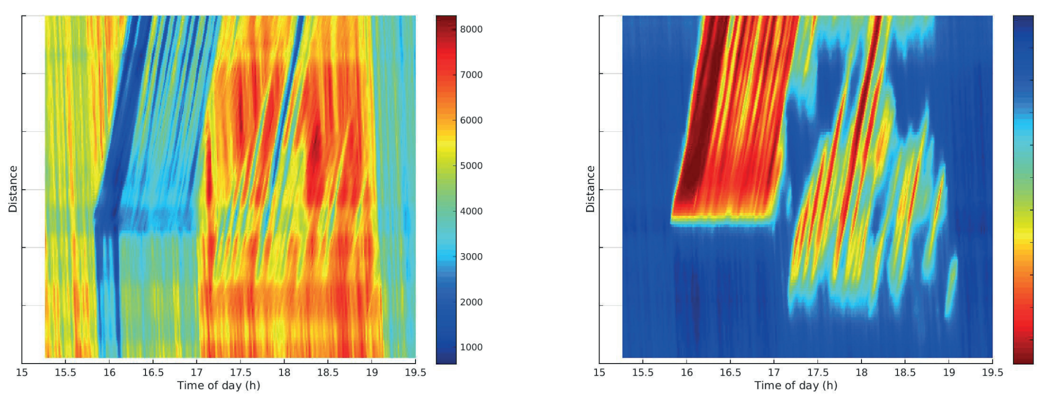
Exercise 244. *Can a stationary tail of the queue occur* due to *a moving bottleneck? If so, explain how. If not, explain why not. (2 points)*

Someone claims traffic on a road is behaving according toa Greenshields fundamental diagram with free flow capacity=2000 veh/h, critical density = 20 veh/km and kjam=150 veh/km.

Exercise 245. *Is this possible? If so, sketch the fundamental diagram in the speed-density plane and comment on the construction; if not possible, argue why. (2 points)*

A-10-2 Traffic flow properties

Consider the following flow contour plot (left) and speed contour plots (right):



Exercise 246. *Derive in which direction traffic is flowing (2 points)*

The spatial scale is lacking.

Exercise 247. *Reason what is the length of the road shown in the figure. Base your answer on traffic properties. (2 points)*

Exercise 248. *Does traffic at this day exhibit a capacity drop? If so, estimate the capacity drop, and explain how you find it. If not, explain how a capacity drop would show.* For clarity, hand in the graphs in which you indicate the various observation points. *(2 points)*

A-10-3 Speeds

Consider a road with slugs and rabbits

Exercise 249. *What are the behavioural assumption(s) with regard to lane choice* for rabbits *in Daganzo's theory of slugs and rabbits. (1 point)*

Traffic is in a two-pipe regime, and the speed of rabbits is 100 km/h and the speed of slugs is 80 km/h. Instantaneously on the the road, the amount of slugs and rabbits is equal.

Exercise 250. *Compute which fraction of flow is rabbits. (2 points)*

Exercise 251. *Compute the time mean speed. (2 points)*

Exercise 252. *Explain how the mean speed can be computed using Edie's generalised definitions.* Equations only do not suffice; explanation of the working of the equations is needed. *(2 points)*

A-10-4 Analyse a car-following model

For car-following, the following model is being proposed (Helly model):

$$a(t) = \alpha\left(v_i(t-\tau) - v_{i-1}(t-\tau)\right) + \beta\left(x_{i-1}(t-\tau) - x_i(t-\tau) - s_0 - Tv_i(t-\tau)\right)$$

In this model, x denotes the position, v denotes the speed and a denotes the acceleration, all evaluated at a time t or $t-\tau$ (τ >0 representing a reaction time); subscripts i and $i-1$ indicate the current vehicle or its leader, respectively. Further parameters of the model are s_0 and T, as well as sensitivity parameters α and β.

Exercise 253. *Explain the working of this model in words.(2 points)*

Exercise 254. *Sketch the congested branch of the fundamental diagram in the speed-spacing plane. To do so, first derive this relationship in mathematical expressions.(5 points)*

Due to hysteresis not all traffic states are on the fundamental diagram.

Exercise 255. *Sketch in the speed-spacing plane how traffic states move in the fundamental diagram, i.e. indicate hysteresis. Motivate your answer.(2 points)*

The fundamental diagram can be transferred to the flow-density plane.

Exercise 256. *Derive this relationship mathematically.(2 points)*

A-10-5 Traffic light

Consider a road at an intersection. The traffic can be described with a triangular fundamental diagram with a capacity of 1500 veh/h, a free flow speed of 50 km/h and a wave speed of -15 km/h. The inflow is 1000 veh/h.

Exercise 257. *Sketch the fundamental diagram in the flow-density plane.(1 point)*

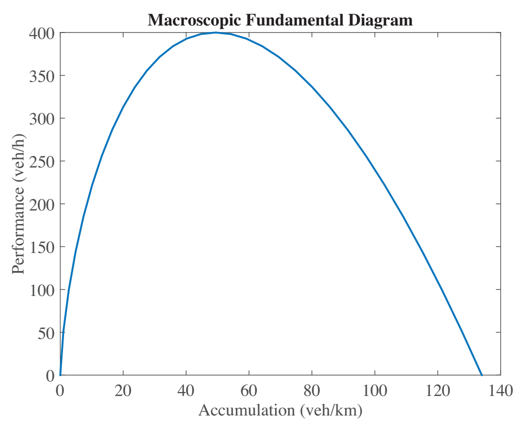
Exercise 258. *Compute the jam density. (2 points)*

At t=0, the traffic light turns red for 2 minutes. Then, the traffic light turns green for 0.5 minute, after which the traffic light turns red again for 1 minute. Afterwards, the traffic light stays green.

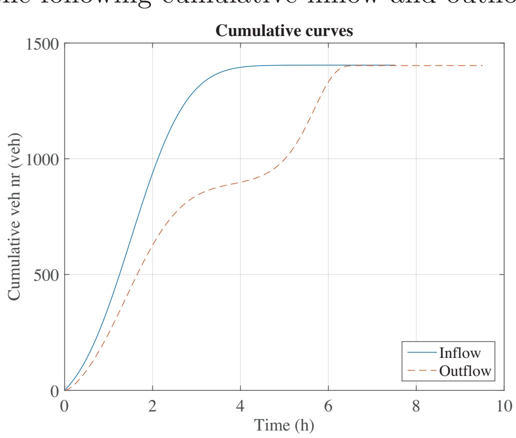
Exercise 259. *Using shockwave theory, sketch the traffic states in the space-time diagram and show the states and waves in the flow density diagram. Comment on how you find the lines. (9 points)*

A-10-6 Macroscopic Fundamental Diagram

For a zone the following Macroscopic Fundamental Diagram holds:



The performance on the vertical axis indicates the outflow out of the zone; the total road length in the network is 3.8 km. Assuming homogeneous and stationary traffic conditions, the following cumulative inflow and outflow curves are obtained:



The inflow curve is a result of (exogenously given) demand.

Exercise 260. *Sketch the inflow pattern as flow over time; estimate the maximum flow into the road. (2 points)*

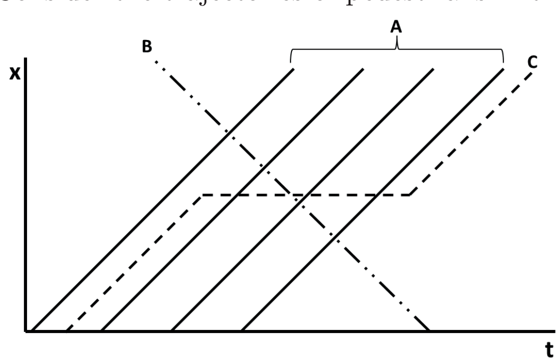
Exercise 261. *Explain the shape of the outflow curve; in particular, include a description of the process underlying the shape from t=2h to t=5h. (3 points)*

Exercise 262. *Explain how perimeter control can reduce delay. (2 points)*

Exercise 263. *Sketch in the graph the amount of total delay that can be reduced; comment on your construction. (2 points)*

A-10-7 Non-vehicular traffic

Consider the trajectories of pedestrians in this figure:



The trajectory of pedestrian C is temporarily horizontal.

Exercise 264. *Does this imply he is standing still? Argue why. (1 point)*

Pedestrian traffic can be described with the social force model

Exercise 265. *Give a conceptual description of the model (equations are not needed); in your answer, include whether the model is microscopic or macroscopic. (3 points)*

A-11 Problem set 11

A-11-1 Short questions

Exercise 266. *Describe what characterizes the difference between a cellular automaton model and another car-following model. (1 point)*

Exercise 267. *Give a reasonable value (with unit) for the capacity of pedestrians. (1 point)*

Exercise 268. *Explain why the MFD does not have the same "pointy" top as a fundamental diagram for a road. (1 point)*

Exercise 269. *Explain why in the method of characteristics, characteristics cannot cross each other in the xt=plane. Comment on an intersection of characteristics of the same traffic state and of different traffic states. (2 point)*

In a traffic stream with independent arrivals, both the exponential probability distribution function as well as the Poisson distribution function play a role.

Exercise 270. *Give the name of the variable of which the distribution is quantified by a Poisson distribution (be precise). (1 point)*

A-11-2 Biking queues

A group of 50 people approaches a traffic light with a flow of 1 cyclist per 2 seconds. A traffic light is red when they approach, and turns green 10 seconds after the last person has arrived at the traffic light. The capacity is 1 cyclist per second.

Exercise 271. *Sketch the cumulative curves upstream and downstream of the traffic light (assuming vertical queuing). (2 points)*

Exercise 272. *Indicate how the delay can be computed with cumulative curves. (1 point)*

One can approximate the traffic operations with a triangular fundamental diagram. Then, the same questions could be computed with shockwave theory.

Exercise 273. *Argue whether the delay computations would be different using shockwave theory, and if so, whether they would be smaller or larger. (2 points)*

A-11-3 Speed averaging

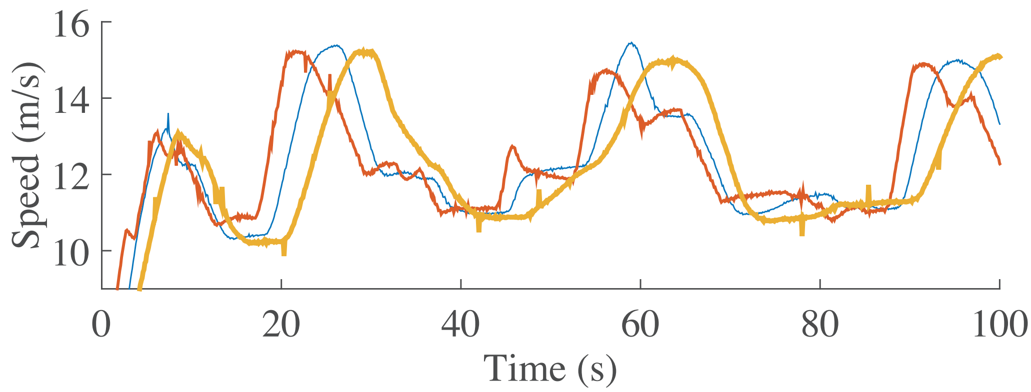
Consider a group of cyclists all following the same circular route: 80% rides a normal bike, and 20% a bike with electrical support motor. Cyclists at a normal bike move at 15 km/h and those with an electrical support motor move at 25 km/h

Exercise 274. *Argue whether the time mean speed is higher or lower than the space mean speed.(2 points)*

Exercise 275. *Compute the average speed on the track using Edie's generalised definitions of traffic. Show the steps you take. (4 points)*

A-11-4 Experiment of ACC equipped cars

An experiment was carried out with vehicles driving an adaptive cruise control, i.e., the vehicle adapts the speed to its predecessor, which is measured by radar. The graph below shows the resulting speeds.



The platoon consists of these three vehicles

Exercise 276. *Give the convention of numbering vehicles: upstream to downstream or downstream to upstream. (1 point)*

Exercise 277. *Indicate which line relate to each of the vehicles 1 to 3. Argue why. (2 points)*

Exercise 278. *Comment on the local stability for these vehicles. (1 point)*

Exercise 279. *Comment on the platoon stability for these vehicles. (2 points)*

A-11-5 Accident downstream of a bridge

Consider a road with a bridge. Drivers do not exceed the speed limit. In congestion, they act according to the Helly car following model. The acceleration a at moment t is given by:

$$a(t) = \alpha\left(\Delta v(t-\tau)\right) + \gamma\left(s(t-\tau) - s^*\right) \tag{A-15}$$

t is the time, v is speed, and Δv is the speed difference with the leader. s is the spacing, s^* is given by:

$$s^* = s_0 + Tv(t-\tau) \tag{A-16}$$

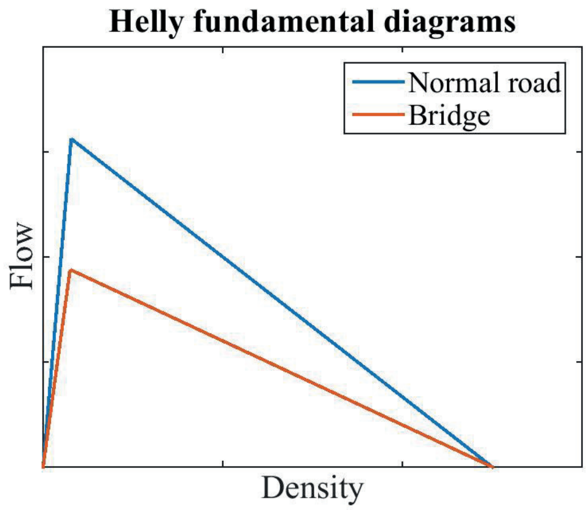
τ, s_0, and T, as well as α and γ are parameters on driving behavior.

Exercise 280. *Explain the working of this model. (2 points)*

Exercise 281. *Give a realistic value for τ (include units in your answer where relevant). (1 point)*

Exercise 282. *Reason which parameters determine the fundamental diagram. (2 points)*

On the road, there is a bridge At the bridge, driver behavior changes due to narrower lanes. The fundamental diagram hence changes. Both fundamental diagrams are shown.



Note both branches change.

Exercise 283. *Reason which model parameter(s) change(s). Only comment on parameters which can be derived from the fundamental diagram.2*

On a particular day, there are free flow traffic operations until an accident happens downstream of the bridge which temporarily blocks the road completely. The tail has spilled back further upstream than the bridge when the incident is cleared. Consider constant inflow.

Exercise 284. *Sketch the traffic operations in the time-space diagram, using a construction in the density-flow diagram. Show the construction steps. For simplicity, you may assume the bridge has a length of 0 (but the capacity constrain remains). (8 points)*

A-11-6 Moving bottleneck principles

Consider a two-lane road with a Greenshields fundamental diagram. The free flow speed is 60 km/h and the jam density is 100 veh/km/lane, hence 200 veh/km for the roadway.

Exercise 285. *Argue whether this information is sufficient to specify the fundamental diagram. If so, determine the roadway capacity. If not, indicate which information is missing, and state a reasonable value for this. (3 points)*

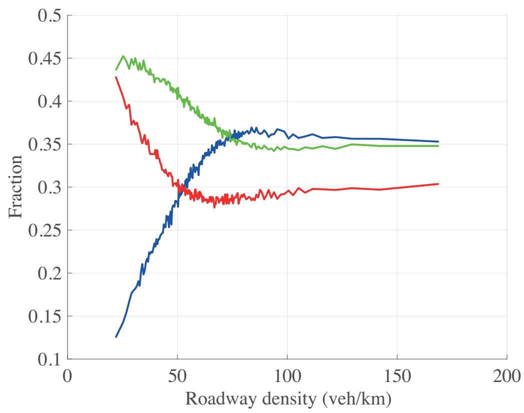
A truck is driving slowly at 15 km/h. Traffic can overtake the truck. The capacity point at the (moving) bottleneck is characterised by a density of 50 veh/km, and a flow of 1000 veh/h. This will yield congestion.

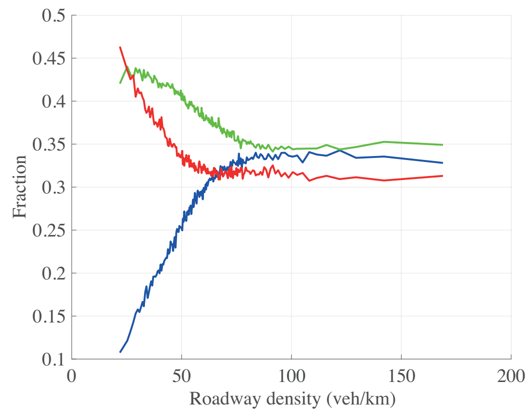
Exercise 286. *Sketch the fundamental diagram in the flow-density plane (if needed with the additional estimates).(1 points)*

Exercise 287. *Sketch how to find the states upstream and downstream of the moving bottleneck in the fundamental diagram. (2 points)*

A-11-7 Multi-lane traffic

Consider the following graphs, representing the fraction of flow and the fraction of density at each of the three lanes on the three lane part of the A13 motorway near Delft in the evening peak.



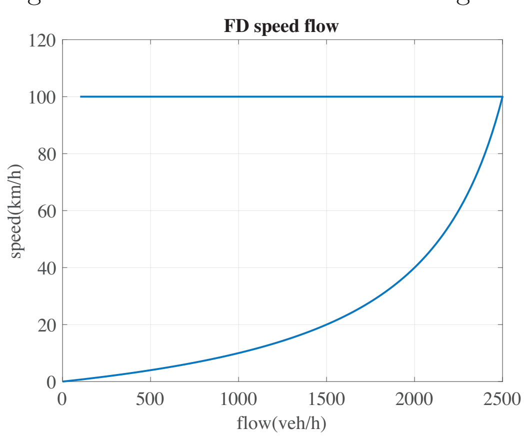


Exercise 288. *Argue which line denotes the left lane, which the center lane and which one the right lane.(2 points)*

Exercise 289. *Argue which graph denotes the fraction of flow and which the fraction of density. (3 points)*

A-11-8 Macroscopic traffic simulation

A road is split into cells of length Δx. For each of the cells, the same triangular fundamental diagram holds. This fundamental diagram is depicted in the speed-flow plane.



Exercise 290. *Draw (precisely, no sketch; use graph paper for this question) this fundamental diagram in the flow-density plane. Comment on the construction.3*

At a certain moment, the cells have the following densities:

Direction	→				
Cell number	1	2	3	4	5
Density (veh/km)	5	20	100	20	125

Exercise 291. *Calculate the flow (in veh/h) from cell 2 to cell 3 according to the cell transmission model.2*

Exercise 292. *Calculate the flow from (in veh/h) cell 3 to cell 4 according to the cell transmission model.3*

A-12 Problem set Urban Intersection control

A-12-1 Urban Intersection questions

An intersection has a layout as given in figure A-2. In this figure the movements are indicated with their approach and direction, for instance, E-Sb is East approach, South bound. For this intersection, in table A-1 the total flow, the saturation flow and the signal timing are given, in table A-2 the clearance time.

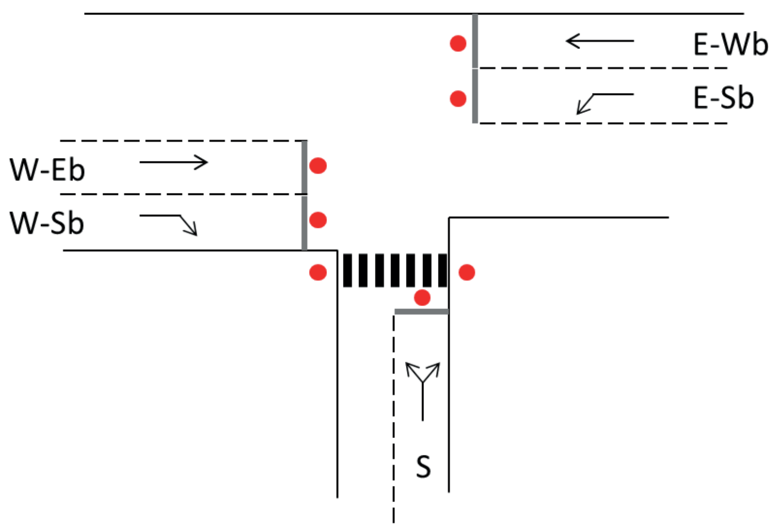


Figure A-2: Intersection layout

Table A-1: Traffic flow and timing at the intersection, ped: pedestrian crossing

	E-Wb	E-Sb	S	W-Sb	W-Eb	ped.
total flow (pcu/h)	800	300	500	300	400	100
saturation flow (pcu/h)	1800	1700	1600	1500	1800	∞
minimum green (s)	6.0	6.0	6.0	6.0	6.0	8.0
yellow time (s)	3.0	3.0	3.0	3.0	3.0	4.0
start loss time (s)	1.0	1.0	1.0	1.0	1.0	0.0
yellow use time (s)	1.0	1.0	1.0	1.0	1.0	0.0

Table A-2: Clearance times in seconds

	E-Wb	E-Sb	S	W-Sb	W-Eb	ped.
E-Wb		-	-	0.0	-	-
E-Sb	-		0.0	2.0	2.0	4.0
S	3.0	2.0		-	0.0	0.0
W-Sb	-	0.0	-		-	0.0
W-Eb	-	0.0	1.0	-		-
ped.	-	0.0	3.0	2.0	-	

Exercise 293. *Code the signals of the intersection in figure A-2 according to the Dutch standard coding.*

Exercise 294. *Explain that at least two different control structures are possible given the conflict matrix.*

Exercise 295. *Determine all control structures.*

Exercise 296. *Determine the Webster cycle time and the minimum cycle time for all control structures.*

A-13 Problem set 12

A-13-1 Short question

Exercise 297. *What are the assumptions for rabbits in Daganzo's theory on slugs and rabbits2*

A-13-2 Recognizing traffic states

Consider the following speed contour plots (left) and flow contour plots (right). There are two different situations (two rows of figures). There are no onramps on the road.

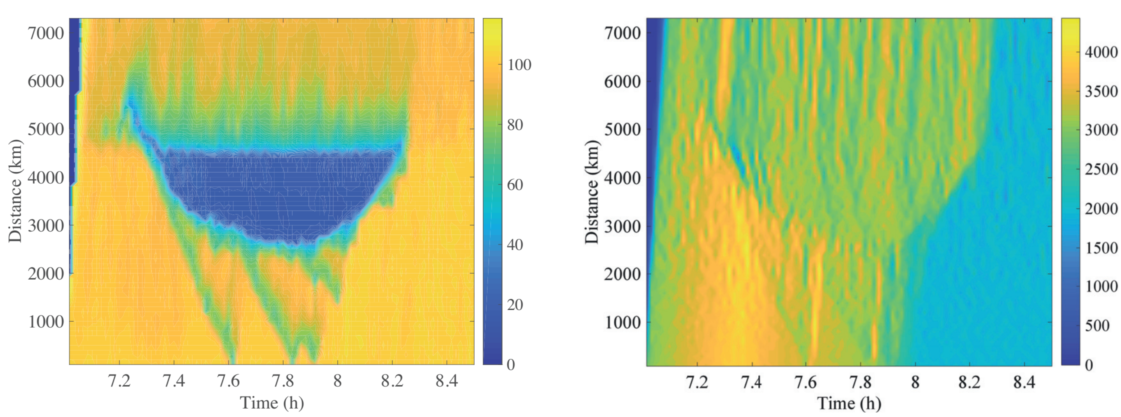


Figure: situation 1 (above this text)

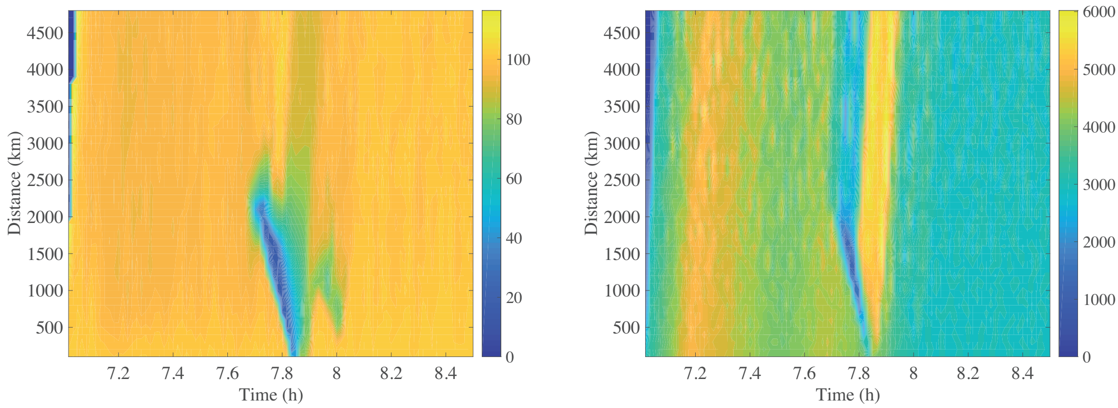


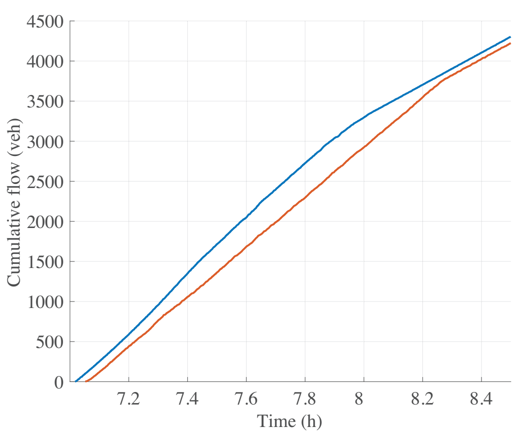
Figure: situation 2 (above this text)

Exercise 298. *Explain how it can be seen that no traffic leaves the road. (1 point)*

Exercise 299. *Describe the traffic pattern and give a likely cause for this traffic pattern for situation 1. (2 points)*

Exercise 300. *Describe the traffic pattern and give a likely cause for this traffic pattern for situation 2. (2 points)*

The following graph shows cumulative curves for one of the situations. The curves are created for locations which are useful for the situation at hand (and it is on purpose that the location specified in the question).



Exercise 301. *Argue for which of the two situation this set of curves is made. (1 point)*

Exercise 302. *Estimate the delay for this situation: give your reasoning and computation steps. (You can hand in the last page with the graph to show steps) (3 points)*

A-13-3 Systems and control, Ramp metering

Exercise 303. *Draw the block diagram of a feedbak control system. Give the names of the arrows and blocks. (3 points)*

Suppose we use the Alinea ramp control algorithm.

Exercise 304. *Specify for this case:*

- *the control signal,*
- *the controller,*
- *the controlled process,*
- *the measurement.*

. *(2 points)*

Exercise 305. *Explain whether ALINEA has a feedback or feed-forward control structure.1*

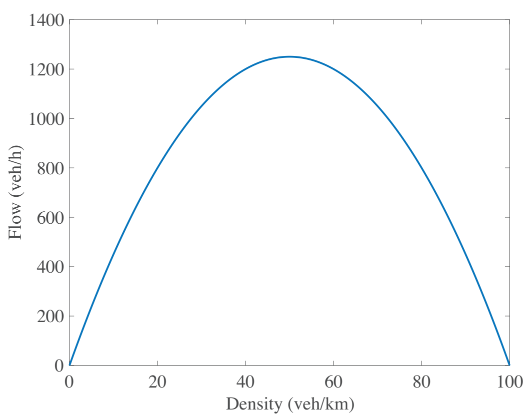
Given a multiple-car-per-green type of ramp metering system, with one metered lane, an effective green time of 20 s, a cycle time of 40 s, and a saturation flow of 1800 veh/h. The ramp metering system is replaced by a one-car-per-green type of ramp metering. The new system serves the same demand, but spreads the vehicles more evenly.

Exercise 306. *Compute be the average red time of the new system, if the yellow time is* 0.5 *s and the average green time is* 1 *s? Explain your answer. (2 points)*

Exercise 307. *Explain the advantage of a one-car-per-green type of ramp metering over the multiple-cars-per-green type. (2 points)*

A-13-4 Shock wave theory

Consider a road for which the Greenshields fundamental diagram holds for each of the lanes. The diagram in flow density is as follows



Exercise 308. *Draw (not sketch) this fundamental diagram in speed density plane. (2 points)*

The road consist of 2 lanes. Consider a temporal blocking of 1 of the two lanes. The inflow is constant at 2000 veh/h.

Exercise 309. *Sketch the traffic operations in the space time diagram. Show the steps of your sketch and their relations in the fundamental diagram (5 points)*

A-13-5 Performance measures and control objectives

Suppose we have a traffic network where the following control measures are applied to optimize the performance: route guidance, variable speed limits, and ramp metering. We need to choose an objective function that the controller optimizes over a pre-defined time horizon.

Exercise 310. *Explain whether* maximizing the total travelled distance *would be a suitable objective for this problem. (1 point)*

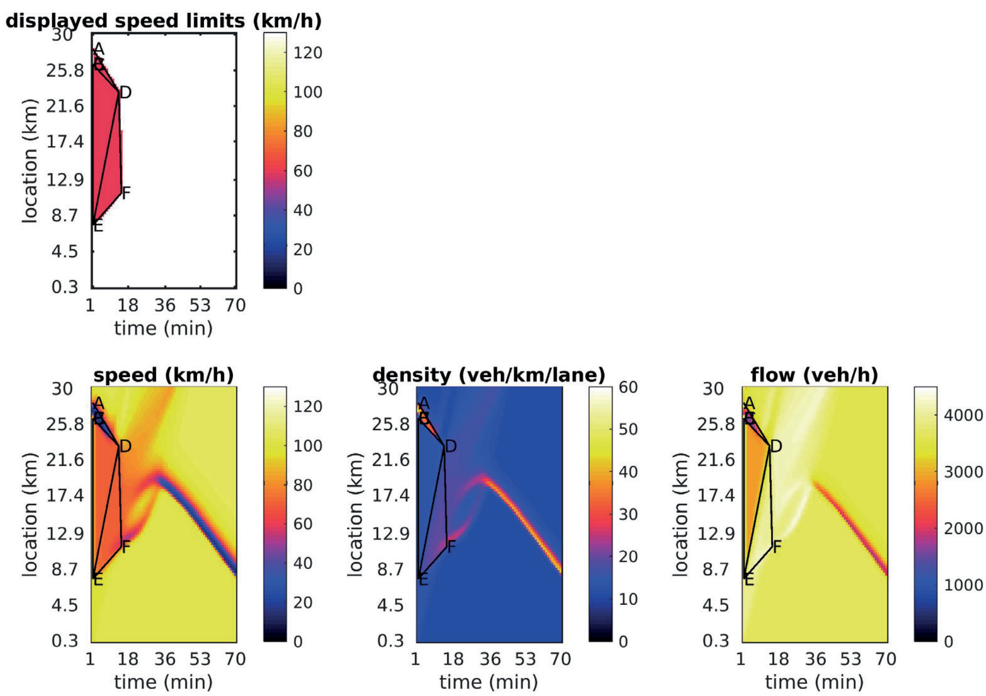
Exercise 311. *Explain whether* minimizing the total travelled distance *would be a suitable objective for this problem. (1 point)*

Exercise 312. *Explain whether* maximizing the total time spent *would be a suitable objective for this problem. (1 point)*

Exercise 313. *Explain whether* minimizing the total time spent *would be a suitable objective for this problem. (1 point)*

A-13-6 SPECIALIST

Given a SPECIALIST tuning as shown in the figure below:

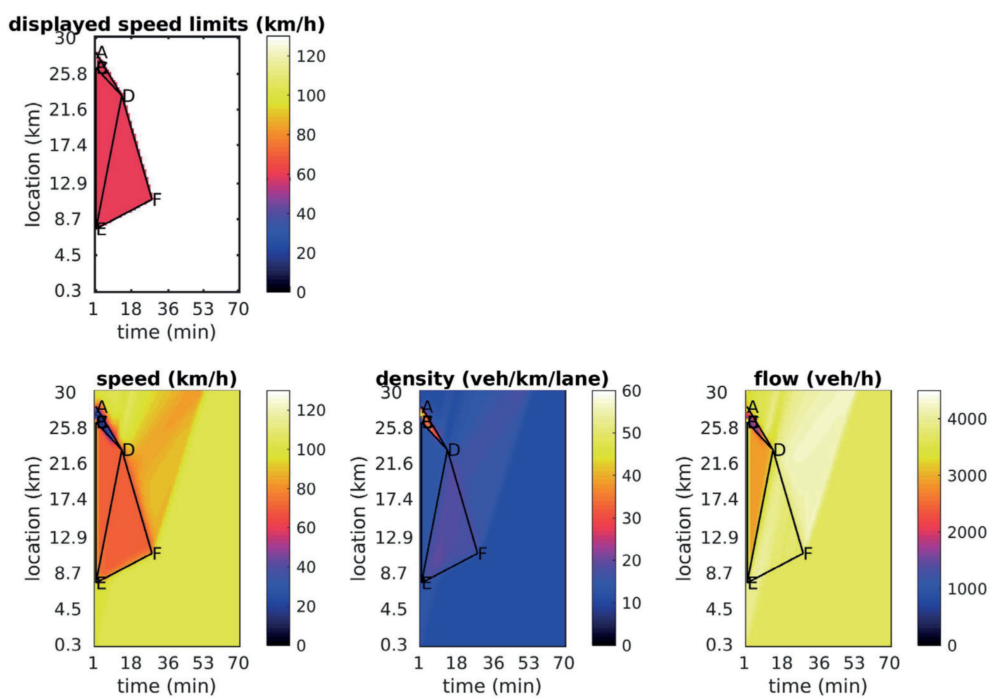


Exercise 314. *Explain which parameter you would change and in which direction to make SPECIALIST work correctly. (2 points)*

Given a SPECIALIST scheme as indicated in the figure below with the following coordinates and states:

A:	0.0167h	28.300 km
B:	0.0167h	26.500 km
C:	0.0183h	26.454 km
D:	0.2269h	23.198 km
E:	0.0183h	7.5588 km
F:	0.4705h	11.0163km

state 1:	v_1= 56.12 km/h,	ρ_1 = 58.3 veh/km,	q_1 = 3272 veh/h
state 2:	v_2 = 16.76 km/h,	ρ_2 = 114.2 veh/km,	q_2 = 1914 veh/h
state 3:	v_3 = 75 km/h,	ρ_3 = 40.81 veh/km,	q_3 = 3060 veh/h
state 4:	v_4 = 75 km/h,	ρ_4 = 54 veh/km,	q_4 = 4050 veh/h
state 6:	v_6 = 96.78 km/h,	ρ_6 = 40.81 veh/km,	q_6 = 3949 veh/h



Exercise 315. *Given a vehicle at t_0 = 0.25h, x_0 = 13 km. Calculate when this vehicle will leave the stabilization area. (4 points)*

A-13-7 Car-following and fundamental diagram

Consider the Intelligent Driver car-following model, given by the equations:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = a_0 \left(1 - \left(\frac{v}{v_0}\right)^4 - \left(\frac{s^*(v, \Delta v)}{s}\right)^2\right) \tag{A-17}$$

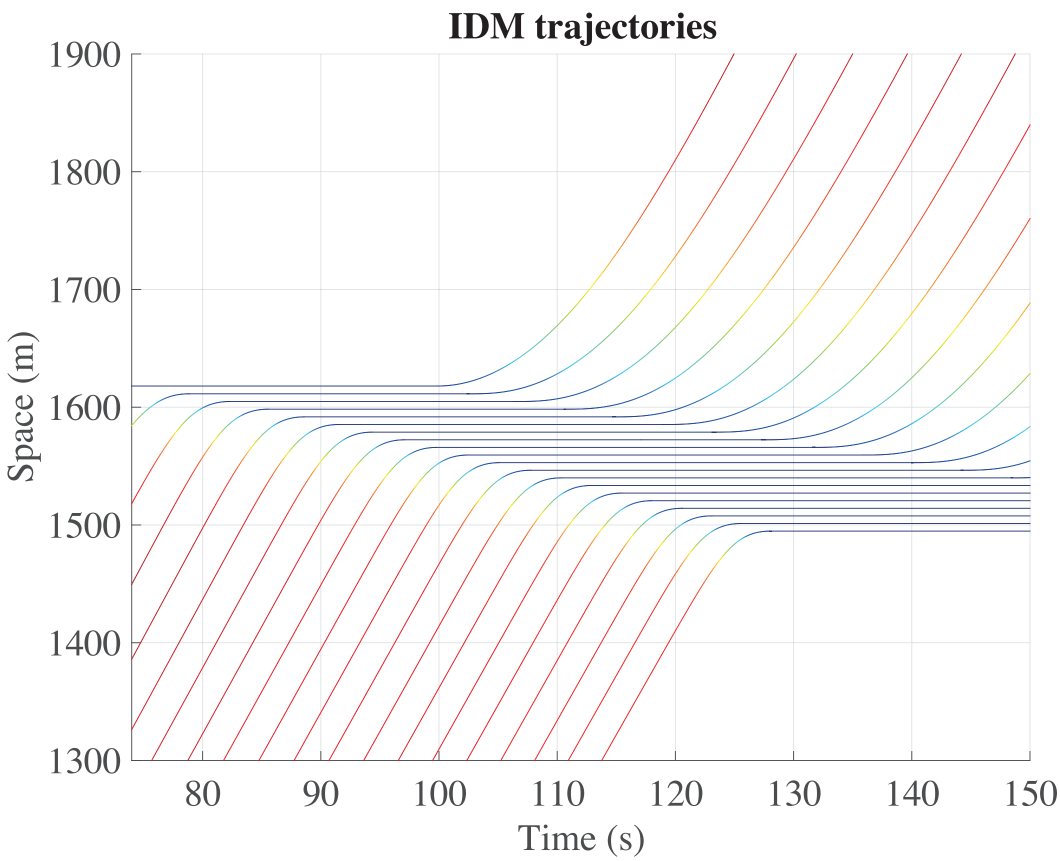
with a desired spacing s^* as function of speed (v) and speed difference between leader and follower (Δv):

$$s^*(v, \Delta v) = s_0 + vT + \frac{v\Delta v}{2\sqrt{a_0\, b}} \tag{A-18}$$

Model parameters are a_0, a reference acceleration, and b, a comfortable deceleration, and T the speed-changing part of the distance headway in equilibrium conditions. For parameter s_0 an interpretation is intentionally not given.

s^* and v_0 are often referred to as desired spacing and desired speed respectively.

The following trajectories are created using this model. The leader drives freely, and is then (abruptly) stopped and is kept stationary for a while, before it sets off following the same model equations as the other vehicles. All vehicles have the same parameters.



Exercise 316. *Give the term used in traffic engineering for estimating parameters of a model. (1 point)*

Exercise 317. *Estimate the desired speed v_0 from the trajectories. Explain your steps. (2 points)*

Exercise 318. *Estimate s_0 from the trajectories. Explain your steps. (2 points)*

Exercise 319. *Derive from the trajectories whether this model in these conditions is platoon-stable or not. Explain your reasoning. (2 points)*

Exercise 320. *Determine the pace of the last vehicle at t=120 s. (1 point)*

A-13-8 Urban Intersection questions

A micro-simulation is made of an intersection that is controlled by traffic signals. With this intersection simulation, trajectories are determined for all vehicles that approach and pass the stop line. There are two types of control that can be used for the traffic signals, vehicle-actuated control and fixed time control. To investigate the difference of these types of control, the simulation is run with the same arrival patterns, once for fixed-time control and once for vehicle-actuated control. In Figure A-3a and A-3b trajectory plots are displayed for one of the directions of this intersection, one figure when fixed-time control is used, and figure when vehicle-actuated control is used.

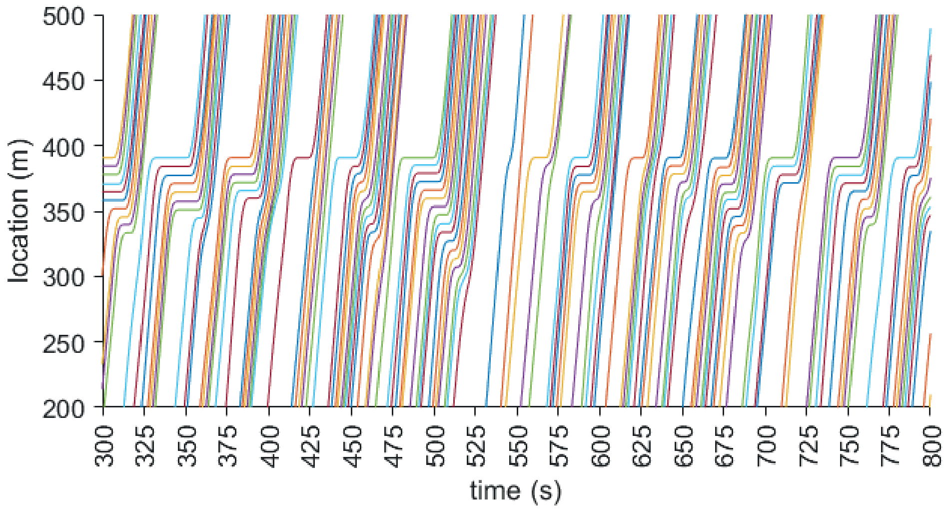
Exercise 321. *Explain, by giving two reasons, which figure shows the trajectories for the vehicle-actuated control.2*

Exercise 322. *Determine the Webster cycle time for the fixed-time control using the trajectory plot, explain how you determined it.3*

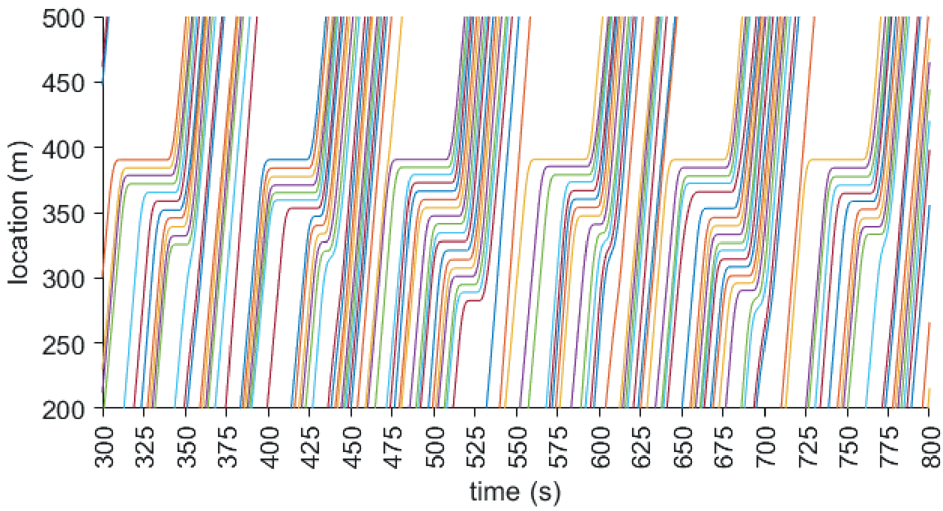
In Figure A-4 two intersections are displayed, which are controlled with traffic signals, using fixed time control. Also information is given to to calculate the Webster cycle time. For both intersections applies: the yellow time is 3 s, the start loss equals the yellow use, the minimum green for all directions is 6 seconds.

Exercise 323. *Calculate the Webster cycle time for both intersections.3*

Exercise 324. *Argue which of the intersections of Figure A-4 has a Webster cycle time that suits the trajectory data of the simulation, given that the simulation is run with the optimum Webster cycle time. 1*



(a) With control 1

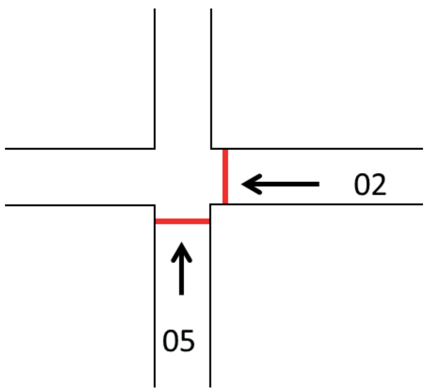


(b) With control 2

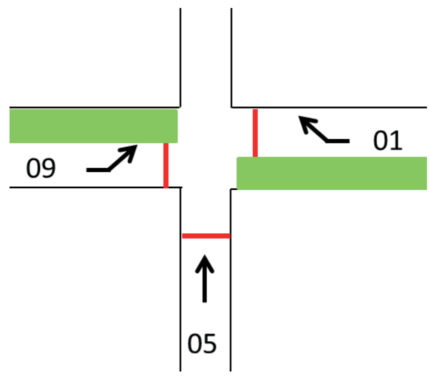
Figure A-3: Trajectories at an intersection, obtained from a micro-simulation

Layout of the intersections

Intesection 1



Intersection 2



Conflict matrix (clearance in s)

Intersection 1

	02	05
02		0.0
05	0.0	

Intersection 2

	01	05	09
01		0.0	0.0
05	0.0		0.0
09	0.0	0.0	

Demand/saturation flow (pcu/h)

Intersection 1

	02	05
demand	600	900
saturation flow	1800	1800

Intersection 2

	01	05	09
demand	600	750	54
saturation flow	1800	1800	1800

Figure A-4: Left: intersection 1, right: intersection 2.

Appendix B

Solutions to test questions

1 Note that there is a capacity drop here, so the fundamental diagram should either contain a capacity drop (e.g., inverse lambda) or be curved.

2 Slugs stay always in the right lane (1) Rabbits choose the fastest lane (1) Free speed of rabbits is higher than the free speed of slugs (1)

3 A count of the number of vehicles over time at one location (1).

4 A model describing the inflow and outflow of the vehicles, with a restriction on the maximum outflow rate. The queues will not occupy any horizontal space (1). This is basically two cumulative curves, inflow and outflow, at one location, in which the angle of the outflow is maximized at capacity (1).

5 Time mean speed, since the faster cars are weighted higher (1)

6 $u = 1/\sum_i (1/v_i)$

7

1. Free flow => free driving + speed
2. Synchonized flow => equal speeds in both lanes, speeds <70 km/h
3. Wide Moving Jams => backwards travelling waves, v very low.

(1 point for names, 1 for definitions)

8 Down-top => stop-and-go wave travelling backwards (0.5) Stop -and-go wave (0.5) and a local bottleneck at 482 (0.5) causing an area of synchronised flow (congestion) (0.5)

9 down-top. Finer structure within the congested part. (0.5) Moving bottleneck (0.5) => structure moves upstream (1) Data error at 8am (1)

10 down-top, from the shockwave speeds (0.5) Incident (1) Synchronized flows (and stop-and-go waves) (1) Moving bottleneck after resolving (1)

11 Trajectory of vehicle is its leaders' trajectory delayed in time (1) and moved backwards in space (1).

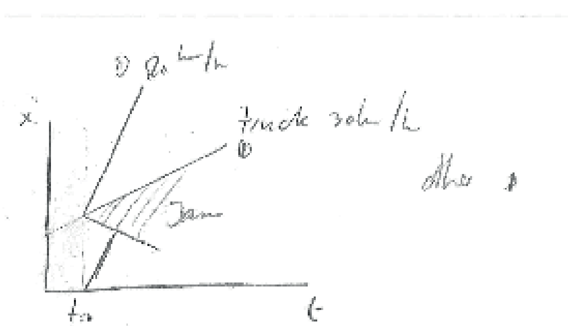
12
1. Follower-leader (local) => over reaction on disturbance in one leader-follower pair
2. Platoon => disturbance grows in a platoon
3. Traffic flow => space between platoons is insufficient, so a disturbance in a platoon causes a delay in the next platoon

13 Yes. Smoothing (1) of disturbances improves platoon stability (1). Thereby, the amplitudes of a disturbance are reduced (1).

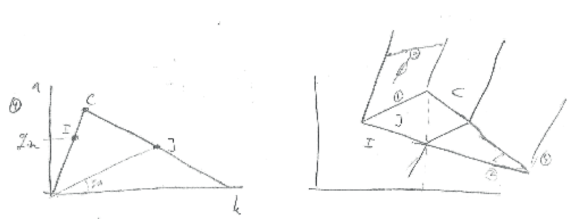
14 A slower vehicle can merge into another lane, which causes a disturbance, growing to a breakdown (1)

15 Now the lane distribution is uneven (1). If vehicles change towards lanes with lower densities, a jam might be prevented (1).

16 Use the triangular fundamental diagram. For the congested branch, the equation is q=2000-2000/100*(k-25). From the speed we q=kv=30k. Solving this set of equations (1), we find k=50, and thus the flow of 30*50 = 1500 veh/h (1)



17
Starting with shockwave theory can be rewarded with 1 pt.



18

19 in I: flow 5000 veh/h, and density 62.5 veh/km, using the equation for the congested branch used in e (1 point). In J: q=4500 veh/h, k=150 veh/km. Wij=-500/87.5=-5.7. So it moves with 4.7 km/h upstream (1).

20 Already in drawing

21 Speed is equal to the truck speed, 30 km/h. (1)

22 Width of a runner 75 cm (1 point for 50-100 cm), headway 1.5 sec (1 point for 0.9 to 2 seconds). 25 meters / 0.75 = 33,3 runners per roadway (both this, and the rounded values are correct). The capacity is 33,3/0,75 x 3600 = 79,200 pedestrians per hour.

23 30,000/79,200 * 60 min/h = 23 min. (1)

24 Halfway: 1 hour duration of the passing of all runners (1) 30,000 runners in 1 hour => 30,000/60 = 500 runners per min. (1) Assumed earlier: 75 com width, 1.5 second headway, so capacity is 1/(0.75 * 1.5) = 0,89 ped/m/s = 53 ped/m/min (1) The required width: 500/53 = 9,4 meters. (1)

25 The congested branch is not a line, but an area (1)

26 An MFD relates the *average* density in an area, the accumulation, to the *average* flow in the area

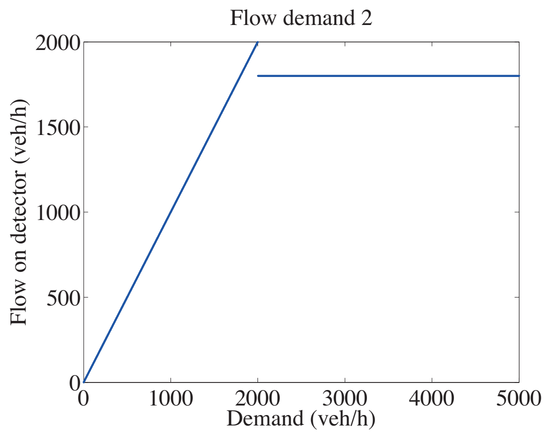
27 The MFD is flattened at the top compared to the regular FD (1). This is due to averaging effects: the top can only be reached if *all* links are operating at capacity. If some have a over-critical density and some an under-critical, the average density can be critical, but the flow is less than at capacity (1).

28 The CTM is a numerical scheme to calculate traffic flows in a macroscopic way (0.5), based on demand of the upstream cell and supply of the downstream cell(0.5). The flow is the minimum of both (0.5); demand is increasing with density up to the critical density and remains constant afterwards, supply starts at capacity and starts decreasing from capacity (0.5)

29 Lagrangian coordinates are used in macroscopic traffic flow description. Instead of fixing coordinates at moments in space, the coordinates go with the traffic (1). Variables used in this system are hence time, *vehicle number* (1) and speed. This is an advantage since the representation of traffic flow equations is more accurate (no numerical diffusion) (1)

30 The maximum flow on the road before the onset of congestion is higher than the flow of vehicles driving out of congestion. (0.5) This phenomenon of drivers keeping a larger headway after a standstill is called the capacity drop. (0.5)

31 Approximately a few to sometimes 30% is claimed

32 

(1 for shape + 1 for capacity drop):

33 the flow equals the demand (0.5), but is never higher than the capacity, and it will not decrease (no over-critical densities) (0.5). Capacity drop explained if present.

34 Capacity is approximately 2000 veh/h/lane (1500-2500: 0.5 pt, possibly lower if urban roads are assumed and explained), so 4000 veh/h for the roadway (0.5)

35 Effective green: $(g-3)$ seconds per direction (0.5). The cylce time is $c = 2(g+2)$ (0.5). Thus,$g = (c/2) - 2$ (0.5). Relative green time: $2\frac{g-3}{c} = 2\frac{c/2-5}{c}$ (0.5) $= 1 - \frac{10}{c}$.

36 Relative green time per direction is half that of the total relative green time: $\frac{1}{2} - \frac{5}{c}$ (1). The flow when effective green is 3600(sec/h)/2(s/veh)=1800 vehicles per hour, or 1/2 vehicle per second (either flow value: 1 point). The flow per direction is thus $1/2(\frac{1}{2} - \frac{5}{c})$ with flow in veh/s and c in seconds or $1800(\frac{1}{2} - \frac{5}{c})$ with flow in veh/h and c in seconds. (1 point)

37 Both methods will yield the same answer, since not the queue length, but the flow is asked. (1) Most vehicles come from direction 2, so the green time (still equal) should be based on direction 2 (1). Using the equation from the last question, we have $1800(\frac{1}{2} - \frac{5}{c}) = 800$ (0.5). Solving this for c yields $\frac{1}{2} - \frac{5}{c} = \frac{8}{18} = \frac{4}{9}$. So $\frac{5}{c} = \frac{1}{10}$, so $c = 50$ seconds(0.5).

38 Poisson distribution

39 The cycle time is 120 seconds. The expected number of vehicles in a cycle is $120/60 * 800/60 = 160/6 = 26.7$ (0.5) so $\lambda = 26.7$ (0.5) 120 seconds cycle time, so 60 seconds per direction. 5 seconds are lost (2 s clearance time and 3 s startup loss), so 55 seconds leading to floor(55/2)=27 vehicles at maximum through a green phase (0.5). The probability of an overflow queue is the probability that the number of cars arriving is 28 or larger, $P(X >= 28)$ (0.5); this can be calculated by $1 - P(X <= 27)$ (0.5). This is calculated as $p = \sum_{k=0}^{27} \frac{26.7^k e^{-26.7}}{k!}$ (0.5)

40 The spread of a Poisson arrival process is larger(0.5), so the probability of having overflow queues is larger(0.5; only with correct reasoning)

41 Vehicles accelerate in principle with acceleration a, but this reduces when they approach their desired speed (1) or if they approach their desired distance (1). The desired distance increases with speed.

42 It is a translation of the leaders' trajectory (0.5), translated forward in time by a fixed value τ and back in space by Δx (0.5).

43 Choose two:

- Drivers are considering multiple vehicles ahead
- Drivers are unable to judge perfectly speed and gap
- Drivers are not changing acceleration continuously
- There is also interaction with lane-changing

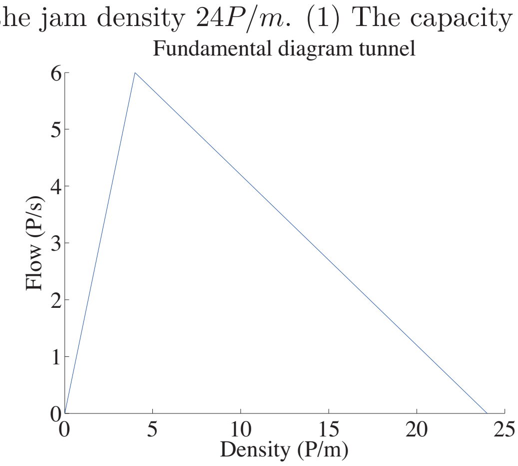
44 Choose 2:

- It better captures the non-equilibrium conditions.
- There is a possibility that stop-and-go waves form (instabilities).
- Drivers have a finite acceleration

45 Choose 2 (1 point per good answer):

- In the IDM, drivers will, unrealistically, never reach their desired speed, not even if the distance is larger than the desired distance
- It has less parameters to calibrate – more realistic
- It is better understandable

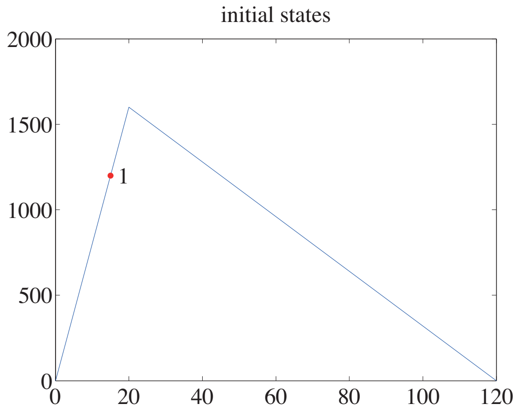
46 Convert the densities to tunnel-wide properties. The critical density then is 4P/m and the jam density $24P/m$. (1) The capacity is vk=6 P/s. (0.5)



(0.5)

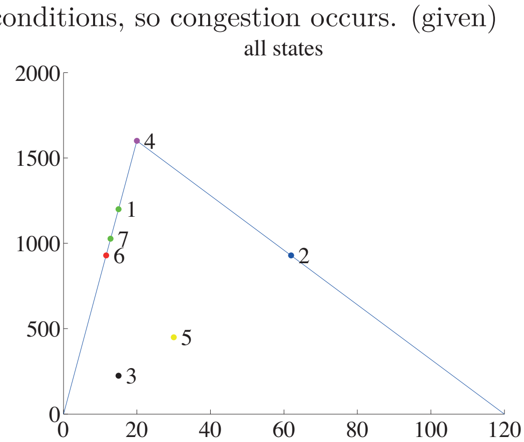
47 180P/min = 3 P/s (unit conversion: 0.5). Undercritical (of course, the tunnel is the bottleneck!), so v=vf. $k = q/v$=3/1.5=2 P/m (0.5). That is 2 pedestrians per meter length, so 0.5 pedestrians per m^2 (1)

48 30 vehicles in a 2 km section, so the density is 15 veh/km. (0.5)



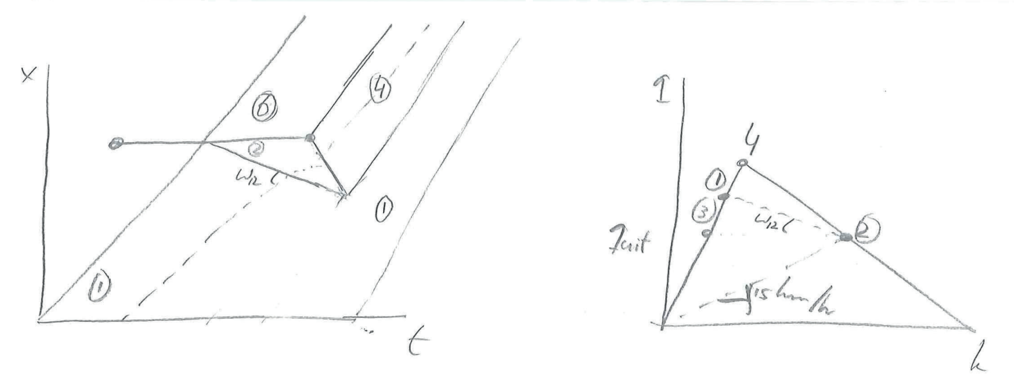
(figure: 0.5);

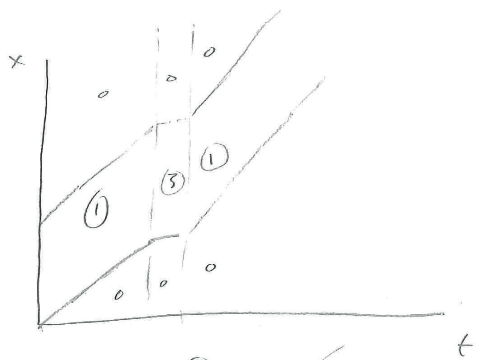
49 Draw a line with an slope of 15 km/h up in the fundamental diagram. You find the intersection with the congested branch at point 2. (0.5 point) The flow is lower than the initial conditions, so congestion occurs. (given)

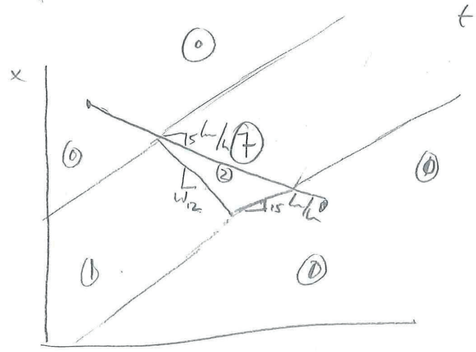


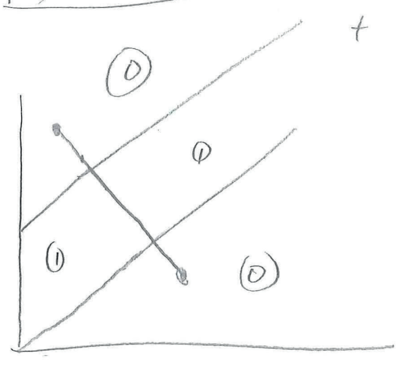
(0.5 points)

For the xt-plot we start with the bottleneck at a constant x from t1 to t2. This causes traffic state 2 to occur. The boundary moves backwards with $w_{1,2}$, which is the same as the angle in the fundamental diagram between state 1 and 2. (0.5). Downstream, the flow is the same as the capacity in 2, but in free flow (state 6) (0.5). After removal of the bottleneck, there are capacity conditions (0.5) and the shock waves propagate with speeds equal to the angle 2-4 and 1-4 (0.5 combined). From $t1$ there also is a shock wave separating state 1 and 6, moving with a speed equalling the slope of line 1-6 in the FD (0.5). Trajectories: 0.5.
All xt-plots are found in the next figure


x
t
w12 l
I
4
2
k
15 km/h




x
w12
15 km/h



50 Since it is temporal, the density remains constant (1 point). This leads to the construction of point 3 a the intersection of the line with a slope of 15 km/h and the density of 15 veh/km (0.5, including showing this point in the graph). Trajectories 0.5.

51 It is given that the tractor moves back very slowly, so we start at the situation comparable to b/ That means there will be a congestion state at the fundamental diagram where people drive 15 km/h (0.5), state 2 in the graph. The moving vehicle forms a boundary in shockwave theory (1), separating state 2 from another state where traffic moves out of the queue. By using the speed of the tractor, we find state 7 (0.5) as outflow state. After the tractor leaves the road, traffic is at capacity (0.5) and restores just like in question c (0.5).
fundamental diagram (0.5) & trajectories: (0.5)

52 The tractor forms a moving bottleneck, and the boundary between two traffic states (0.5). It separates the initial traffic state 1 from another traffic state where vehicles travel at 50 km/h (1). Constructing this in the fundamental diagram is done by drawing a line with a slope of 50 km/h down from point 1. We find the traffic state next to the tractor at this point (0.5), drawing point 5 in fundamental diagram: 0.5. The end of the tractor has a similar effect of a boundary (0.5) moving with 50 km/h. Constructing this in the fundamental diagram gives back point 1 (0.5). So at both ends, we have traffic state 1 (1), and next to the truck state 5 (no points for that). There are no other shock waves. Trajectories move either straight on, or are delayed a bit next to the tractor. (0.5)

53 The left one is Eulerian. Is shows numerical diffusion. Moreover, the grid structure in time and space is clearly visible, whereas in Lagrangian coordinates the grids follow a fixed time and the vehicles (trajectories are visible in the right figure) (1 point for one of the explanations)

54 The representation of traffic flow equations is more accurate (no numerical diffusion) (1); also, the calculations are more efficient since it is not needed to consider both the upstream and downstream cell as with Eulerian coordinates (e.g., CTM) (1)

55 The relative amount of space (in time! -0.5 if not mentioned) that a non-passenger car vehicle occupies on the road

56 Trucks are longer, which influences the pce value. If the pce is the ratio of the length + net headway, assuming the same net headway the pce value increases with deceasing speed.

57 Trucks are longer, which influences the pce value. If the pce is the ratio of the length + net headway, assuming the same net headway the pce value increases with deceasing speed.

58 Vehicles accelerate in principle with acceleration a, but this reduces when they approach their desired speed (1) or if they approach their desired distance (1). The desired distance increases with speed.

59 In regular cf-models, the drivers react on actions by their leader. In multiple-leader models, drivers take several leaders into account in determining their action

60 Calculate for both leaders the desired distance:

$$s_1^*(v, \Delta v) = s_0 + v\,T + \frac{v\,\Delta v}{2\sqrt{a\,b}} \tag{B-1}$$

and

$$s_2^*(v, \Delta v) = s_0 + 2 * v\,T + \frac{v\,\Delta v}{2\sqrt{a\,b}} \tag{B-2}$$

Note that the desired distance to the second leader has twice the speed component.(1) Each of these will lead to an acceleration:

$$a_1 = a_0 \left(1 - \left(\frac{v}{v_0}\right)^4 - \left(\frac{s_1^*(v, \Delta v)}{s}\right)^2\right) \tag{B-3}$$

$$a_2 = a_0 \left(1 - \left(\frac{v}{v_0}\right)^4 - \left(\frac{s_2^*(v, \Delta v)}{x_2 - x_0}\right)^2\right) \tag{B-4}$$

(1) (Note the desired distance to the second leader is compared with the distance between the second leader and current vehicle) These have to be combined into one acceleration (0.5). For safety, we choose the minimum acceleration rather than for instance the mean:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = min\{a_1, a_2\} \tag{B-5}$$

(0.5)

61 This basically is the same as the 2-leader model. First formulate the desired distance to the n-th leader

$$s_n^*(v, \Delta v) = s_0 + nv\,T + \frac{v\,\Delta v}{2\sqrt{a\,b}} \tag{B-6}$$

(1) Then calculate the resulting acceleration for each of them:

$$a_n = a_0 \left(1 - \left(\frac{v}{v_0}\right)^4 - \left(\frac{s_n^*(v, \Delta v)}{x_n - x_0}\right)^2\right) \tag{B-7}$$

These have to be combined into one acceleration (0.5). For safety, we choose the minimum acceleration rather than for instance the mean:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \min\{a_n\} \tag{B-8}$$

(0.5)

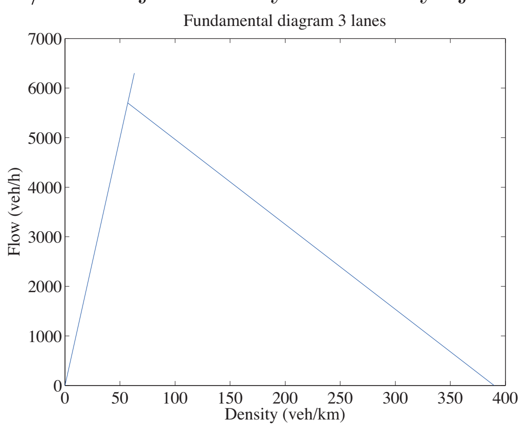
62 The edgy line is the flow, since at the start the demand will exceed the flow before this is reversed. (Also: the flow will be capped at capacity (queue discharge rate) and is expected to be constant for a while). Alternatively: the integral these curves are cumulative curves, and the cumulative curve of the demand cannot exceed the one of the flow. Therefore, the one reducing earlier is the flow)

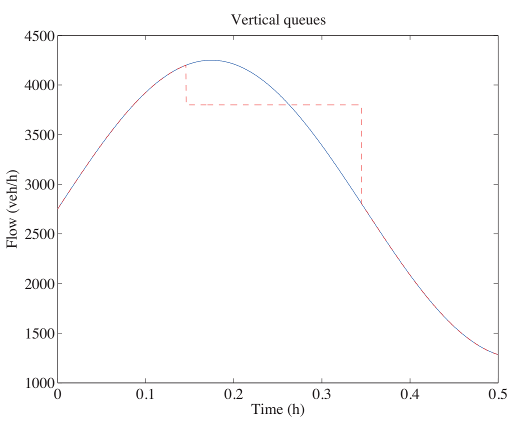
63 The free flow capacity can be found by the flow at the upstream detector before congestion sets in (1), so 4200 veh/h (read from the graph at the time congestion sets in)(1)

64 The queue discharge rate can be found by the flow after congestion has set in (1), so 3800 veh/h (1)

65 The free flow capacity is higher than the queue discharge rate, i.e., there is a capacity drop (1 for either explanation)

66 The FD should have a capacity drop (see previous questions), so the simplest is to assume an inverse lambda shaped FD (0.5 pt either by naming or in drawing) The free flow capacity is found in question a: 3x2100 veh/h. Also the free flow speed is known (from the speed figure): 100 km/h (0.5 for both combined). This gives the free flow branch (0.5 point for drawing). The density matching the queue outflow rate can be found on the free flow branch, by looking up the density for the queue outflow rate (0.5). Kc=qc/vfree=1900/100=19 veh/km. (multiply by 3 for the 3-lane section (1)). Furthermore, it is known that with a flow of 2 lanes x 1900 veh/h/lane (the flow through the bottleneck), the speed is 22,6 km/h (read from graph) (1 pt), which gives a density of 3800/22.6=168 veh/km (0.5 pt). Now, the congested branch can be constructed. We find the wave speed w=$\Delta q/\Delta k$=((3*1900)-(2*1900))/(3*19 -168)=-17 km/h. The jam density is found by kj=Kc-qc/w=3*(19-1900/-17.1)=390 veh/km (0.5 pt).

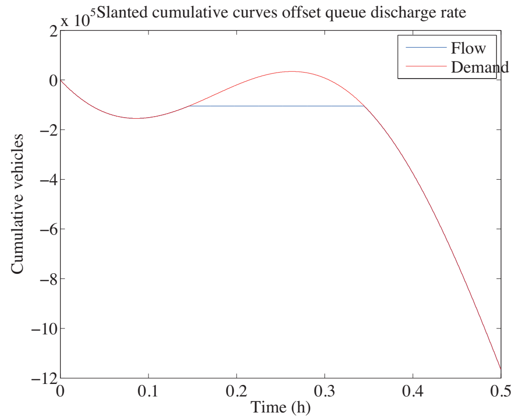




67
Points for the general shape (1) and for the correct times of flow change (1)

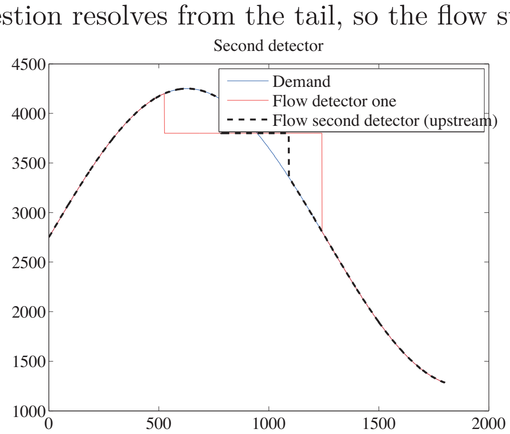
68 The most useful offset is the maximum flow of the road which is maintained for a while, i.e. the queue discharge rate (1 pt). (Alternative: the free flow capacity: 1/2)

69

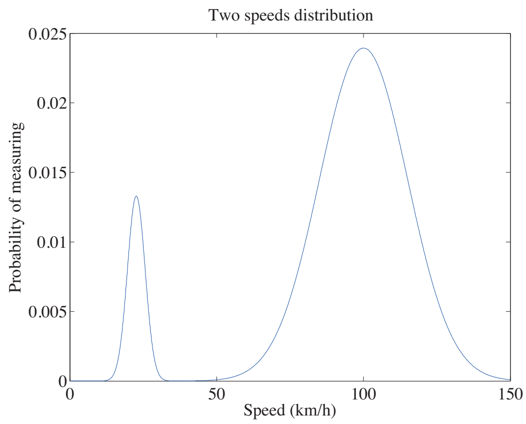


70 This is the area between the curves

71 It takes a while before the congestion spills back to the second detector. Similarly, congestion resolves from the tail, so the flow starts following the demand already earlier in time



72 If we do not relax the assumption of the fundamental diagram, there are only two speeds: the free flow speed and the speed for the state in the congested branch where the flow matches the queue discharge rate for the 2 lane stretch. (1) The frequency at which they are measured depends on the relative time of congestion (typically, 2 peaks at each 1 hour of congestion, so approximately 10% of the time congestion) (0.5). Relaxing the assumption of the fundamental diagram now, we do not find two exact values of the speed, but a distribution centered around two values:



(0.5)

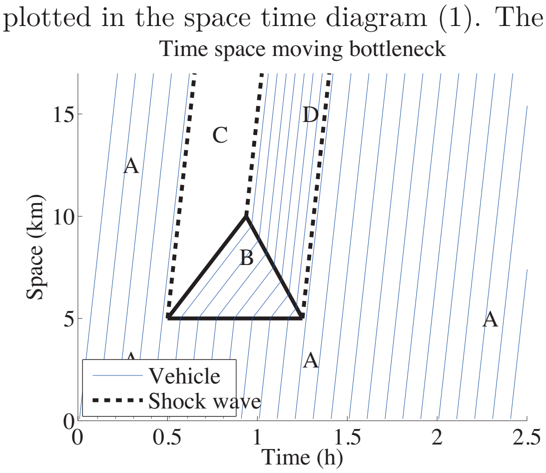
73 Capacity is maximum flow, q=ku=25*80=2000 veh/h/lane. For a two lane section, this is 4000 veh/h. (1 for points) plus 1 for graph

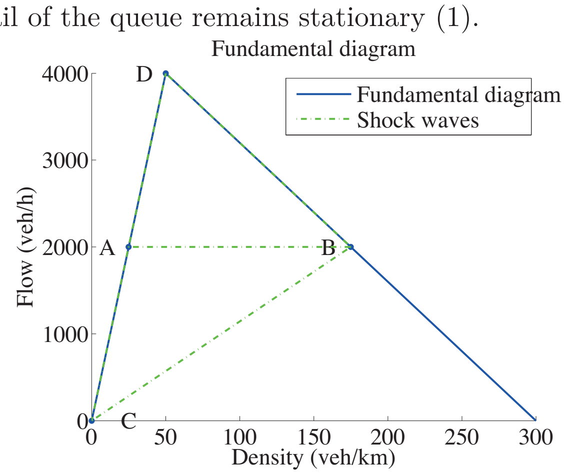
74 Since the tail is at the same place, the flow in the queue should be equal to the inflow (1), so 2000 veh/h.

75 The fundamental diagram is described by q=C(1-((k-kc)/(kj-kc))); use kj=300;C=4000;kc=50; (1) Use the fundamental diagram to find the density at q=2000 veh/h. (0.5) Solving q=1000 veh/h gives k=175 veh/km (0.5)

76 The speed of the vehicle is equal to the speed in the congested state (1). Using q=ku (0.5), this gives v=2000/175=11,4 km/h (0.5)

77 The truck drives at 11,4 km/h, and forms the moving bottleneck for 5 km. This can be plotted in the space time diagram (1). The tail of the queue remains stationary (1).





For the solution of the traffic jam, construct the line in the FD between the jam point and the capacity point (1). Finishing the figure: 1 point.

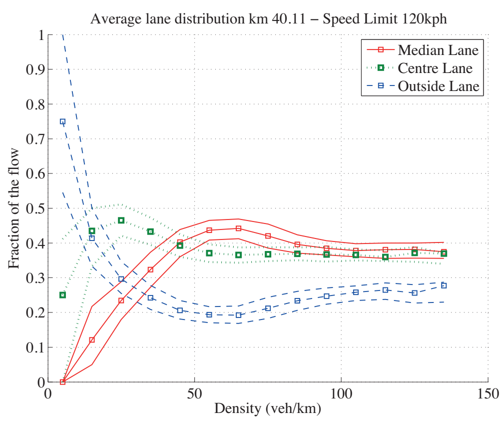
78 The queue has at maximum a length of 5 km, and the density is 175 veh/km – so 5*175= 875 veh.

79 Two-pipe regime and one-pipe regime

80 Two-pipe regime (1). Trucks drive in one lane, and passenger cars in the other, at higher speeds (based on the density) (1).

81 The speed of the passenger cars is higher, so they will pass the stationary detectors more "often" (1). Hence, the fraction of the flow will be higher (1).

82 Requested is the fraction of the flow of passenger cars, (class 1): $\rho = q_1/q_{tot} = q_1/(q_1 + q_2) = (k_1v_1)/(k_1v_1 + k_2v_2)$. (1 point for q=kv per class).
From the image we know k2=4/7*k1 (0.5 point), so $\rho = (k_1v_1)/(k_1v_1 + 4/7k_1v_2) = v_1/(v_1 + 4/7v_2)$ (0.5 point) in which v indicates the speeds. Assuming v_1=120 km/h and v_2=80 km/h (1 point – assuming equal speeds maximizes the nr of points for this question to 1), we find: $\rho = 120/(120 + 4/780) = ...$ (1)



83 If it is really quiet on the road, people keep right (hence blue is right). When it gets busier, more people go towards the middle lane, and then towards the left lane. (0.5 for the right reasoning, 0.5 for drawing the right conclusions; no points for only conclusions; 0.5 if middle and median lane exchanged.)

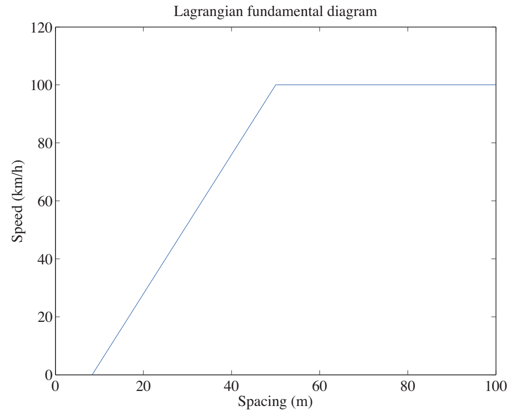
84 Slugs stay always in the right lane (1) Rabbits choose the fastest lane (1) Free speed of rabbits is higher than the free speed of slugs (1)

85 Near capacity, the left lanes travel still faster then the right lane (1). All rabbits are hence in the left lane, leaving gaps in the right lane (0.5) which hence has a lower flow (0.5)

86 It averages the flow and densities (0.5) *for an area* (by which it differs from the regular FD) (1) and relates them to each other (0.5)

87 The average flow decreases for the same average density if the traffic is distributed less homogeneously over the network.

88 The speed is increasing with increasing spacing (0.5). There is a minimum spacing (0.5) and a maximum speed (0.5). Plot: 0.5



89 Vehicles have a desired speed v^* which is determined by the distance to their leader (1). The acceleration is proportional to the difference between their speed and the desired speed.

90 Traffic must be stationary and homogeneous

91 If traffic is stationary, the acceleration is 0 (1), so $v = v^*$ (1). Realising that $k = 1/\Delta x$ (0.5), or correcting for the units (density is in veh/km!) we find $k = 1000/\Delta x$ (0.5), leading to:

$$v = 16.8(tanh(0.086((1/k) - 25) + 0.913)) \tag{B-9}$$

(0.5) This is the speed in m/s, which should be translated into km/h, so:

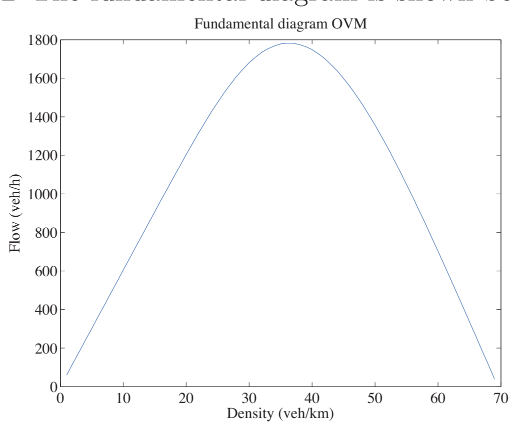
$$v = 3.616.8(tanh(0.086((1/k) - 25) + 0.913)) \tag{B-10}$$

(0.5) Now applying q=ku=kv (0.5), we find

$$q = k3.616.8(tanh(0.086((1/k) - 25) + 0.913)) \tag{B-11}$$

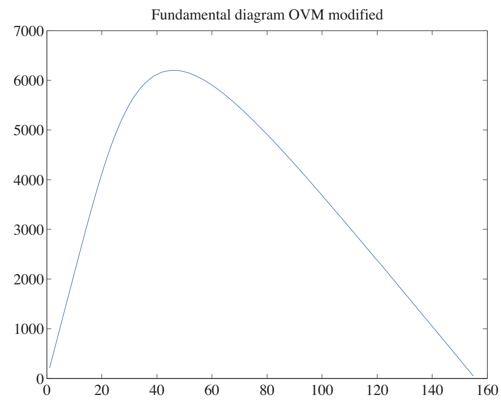
(0.5)

92 The fundamental diagram is shown below:



From this, we read a capacity of approximately 1750 veh/h, the free flow speed is approximately 60 km/h and a critical density of approximately 70 veh/km(1).

93 A free speed of 60 km/h could match a provincial road. However, the capacity for a provincial road is usually not so high. For sure, the jam density would be more in line with the jam density of motorways, i.e. well over 100 veh/km. Therefore, the fundamental diagram is not very accurate. It would be better to have a higher jam density, so scale the graph on the horizontal axis , meaning the number 0.086 should be lower (by about a factor 125/75=1.75 – increasing the jam density to 140). Then, the speed drops even further, so that should be corrected (increase a factor 1.75), and then once more increased by a factor of almost 2 to get a motorway speed . This means increasing 16.8 to 1.75 x 2 x 16.8=... While this would give a proper free speed and jam density, the capacity increases to extreme values (1), so a proper parameter set is not possible



94 Drivers are expected to change lanes whenever it is beneficial for them (measured in terms of acceleration) (1). They do take the other drivers' benefit into account, but less then their own (at factor p) (1).

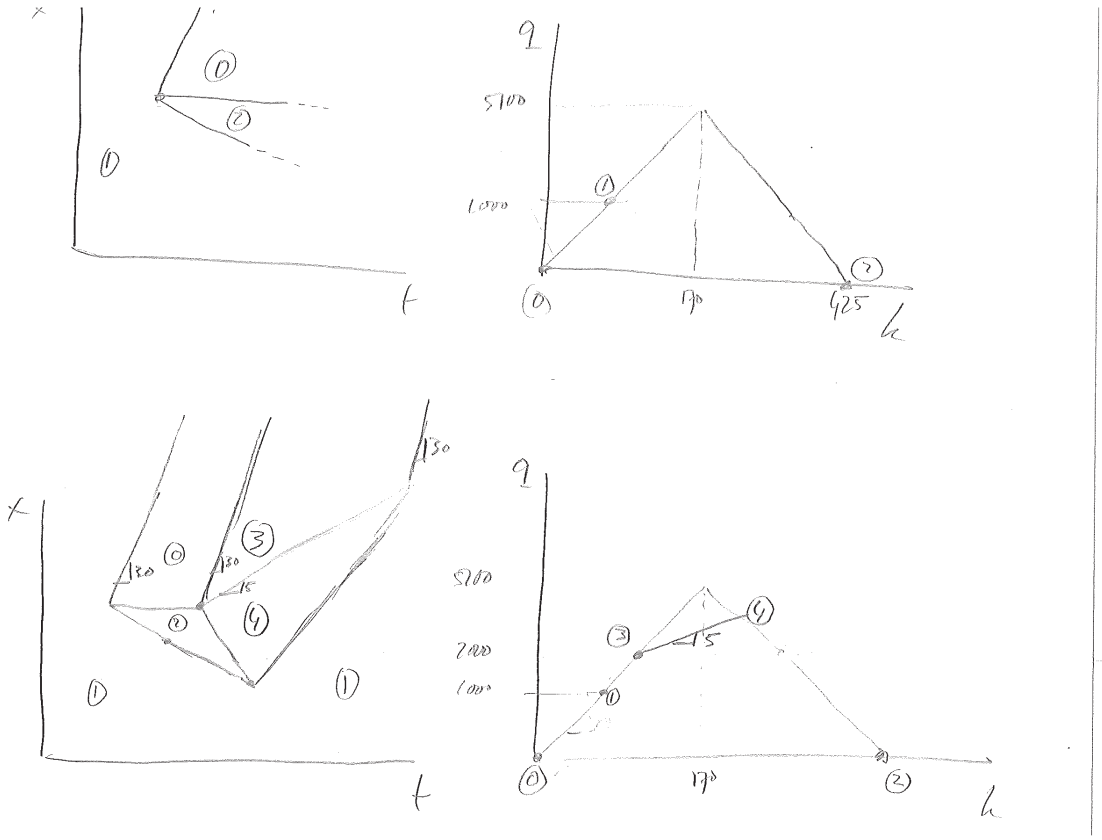
95 It can be expected that drivers take other drivers into account (p>0), but value their benefits less then their own (p<1).

96 No, it does not. The drivers are all at the left lanes. There is no more acceleration in the left lane, so the model does not give higher utility to the left lane.

97 Once the trucks come to a complete stop, upstream an congested area at jams density will be created (0.5); downstream, there will be an empty road (0.5). There are three shock waves. One with the free speed (30 km/h) separating the empty road with from the initial state (1). One at the location of the stopped trucks, separating the empty road from the jam state, at 0 km/h (1). Finally, there is a wave at the tail of the queue, propagating at a speed of $\Delta q/\Delta k$=1000/((1000/30)-425)=-2,5 km/h (upstream). (1 point)

98 The start is the same as described in b. Then, at the head of the queue, vehicles are freed. The fastest vehicles, passing the moving bottleneck, will drive at a free flow speed of 30 km/h (1, including drawing in xt). The snow plows will form a moving bottleneck driving at 5 km/h (1, including drawing in xt). Downstream of this moving bottleneck there will be a free flow

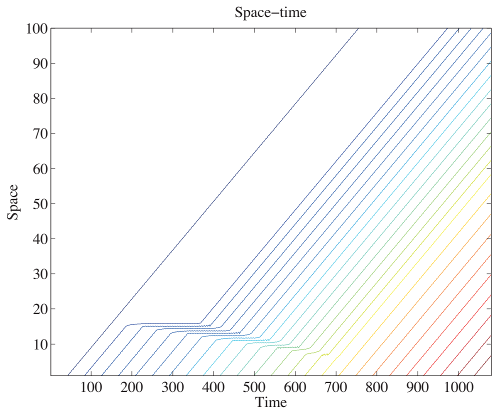
(0.5) traffic state with a flow of 2000 veh/h (given, 0.5) – this is state 3. The plows will form a moving bottleneck, and hence the separation of two traffic states, a congested upstream (state 4, unknown yet) and an uncongested downstream (state 3). (0.5 point). Plotting: 0.5 point. State 4 can be found by plotting the shock wave in the fundamental diagram, from state 3 upwards with a slope of 5 km/h (1). The shock wave speeds between 4 and 1, as well as between 3 and 1 can be found by the slope of the lines connecting these states in the FD (1).



99 The inflow is constant (constant raise of N at x=4 - 1 point), but the outflow is temporarily zero (no extra vehicles temporarily in the graph (1). Hence, there is a temporal blocking (0.5) completely blocking the flow. (0.5)

100 The total delay can be derived from the area between the cumulative curves (1). In this case, we compare the cumulative curves where there is no delay (at x=4km) and that where the delay is maximum (near x=25 km) (1). The surface between the lines is (calculate: 1 point).

101 There is a temporal blocking, so the general pattern is as follows:



(1)

The head of the queue, and thus the location of the blocking is between location 16 and 20 (1). The tail is between km 8 and 12, since there is no delay (1)

102 The capacity is the derivative of cumulative curve during the outflow (0.5), here 2000 veh/h (0.5). The jam density can be derived from the number of vehicles between the cumulative curves in standstill: here 120 veh/km (480 veh in 4 km) (1). The slope of the congested branch van be determined by the speed at which the head of the queue (0.5) moves backward, e.g. at 16 km at t=0.6 and 12 km at t=0.8, i.e. 4 km in 0.2h = 20 km/h (0.5). The critical density then is the jam density minus the capacity divided by the shock wave speed (120-2000/20=120-100=20veh/km - 0.5 pt). The free speed then is the capacity divided by the critical density, i.e. 2000/20=100 km/h.

103 The limiting point for the flow is not the amount of space in the room, but the visitors watching the piece of art (1). Because then there are more visitors that can have a look at the art at the same time, because they do not stand crowded around a single object but have a large circumference to stand around (1).

104 I assume every visitor wants to be in the front row (1 point) for the assumed watching time of 30 seconds per person (10 s - 2 minutes: 1 point). Finally, I assume the width of a pedestrian to be 75 cm (50-110 cm: 1 point). The circumference of the show case is 2x3+2x1=8 meter, which means 8/0.75=10 2/3 visitors can stand at the circumference. (Both rounding to an integer number of visitors, as continuing with the real number is OK – arguing that the spots are not predefined, but dynamically filled). The capacity is 10 visitors per 30 seconds = 20 visitors per minute.(1)

105 When stepping on the moving walkway, visitors keep a gross distance headway of approximately 75 cm (50-110 cm: 1 point; net=gross-20cm); assume a single lane of visitors (for the best view) per walkway. The walkway moves with an assumed 0.5 m/s (0.2-1.5: 1 point). The flow is then 1/0.75*.5=2/3 visitor per second = 40 visitors per minute. (0.5) Since there are walkways at both sides, the capacity is increased to 80 visitors per minute. (0.5)

106 The figure shows the size of the square approximately 500x50/2=12,500 square meters(0.5) The maximum density for pedestrians is approximately 5 ped/square meter (3-10: 0.5 pt). So, at times square at maximum 75,00 people can stand (0.5). The claim cannot be strictly right.(0.5)

107 Speeds in both lanes are equal(1). Alternatives with speeds etc: 0.5 points

108 Sketch showing the areas in more or less the right place (0.5), with appropriate names (0.5). Note that synchronized flow is an area (1)

109 There is no capacity(0.5), since there is a probability of breakdown from free flow to synchronised flow(0.5). This is an increasing function with flow. (Remarks on the area in synchronize flow: no points)

110 The coordinates move with the traffic, rather than that they are fixed at a certain location

111 No numerical diffusion, so more accurate results

112 Traffic is flowing from bottom because the congested patterns propagate in the opposite direction.

113 Bottleneck at km 14 (0.5), creating congestion and instability (1 point for this causality) leads to wide moving jams (0.5 for name)

114 Local: a change of speed of the vehicle cause oscillations with increasing amplitude in speeds with the direct follower (or not)
String: the amplitude of speed disturbances grows (or not) over a platoon,
Traffic: the disturbances within a platoon are propagated (or not) to the next platoon
(one point for the names, one point for the explanation)

115 The stability of traffic is one of the levels of above – we only need to consider that level. We see wide moving jams (or stop and go waves) emerging from an area of congestion, so traffic is unstable in the congested branch. (1) That means traffic is also platoon unstable. (1). Traffic is in general locally stable (only instable in simulations)

116 In the two pipe traffic state (0.5) traffic has different speeds in each lane / slugs and rabbits are separated (0.5); in one pipe traffic state (0.5) slugs and rabbits mix and the speeds are the same (0.5)

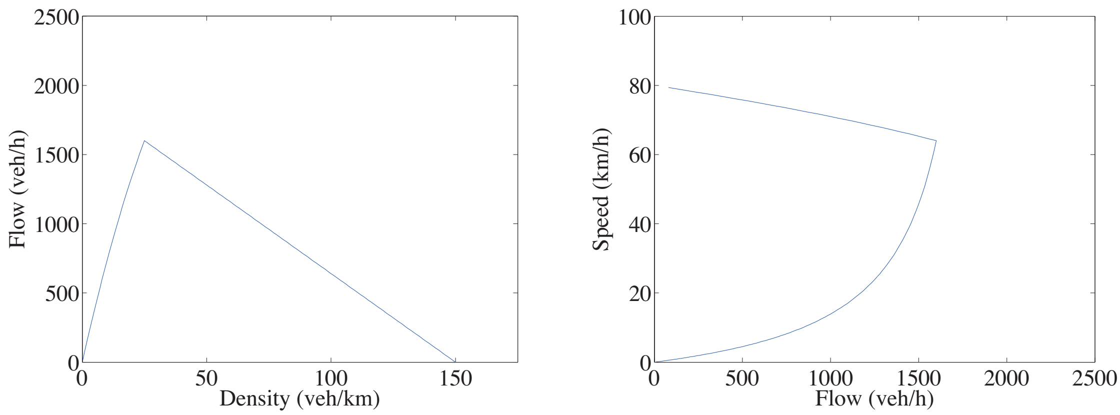
117 According to slugs and rabbits, the traffic is more present in the left lane (1), whereas for the Mobil lane change model there is a stronger tendency for drivers to take the right lane if there is no other vehicle.(1)

118 Stationary means that the traffic stream does not change over time (1). Homogeneous means that the traffic state does not change with space (the same at different locations) (1).

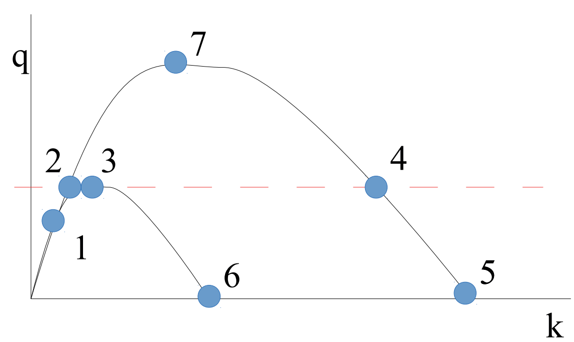
119 It cannot, since the FD gives the equilibrium conditions (1), and drivers have different ways to reach these equilibrium conditions. (1)

120 The right one is the triangular, since there is no influence of density in the speed in the free flow part

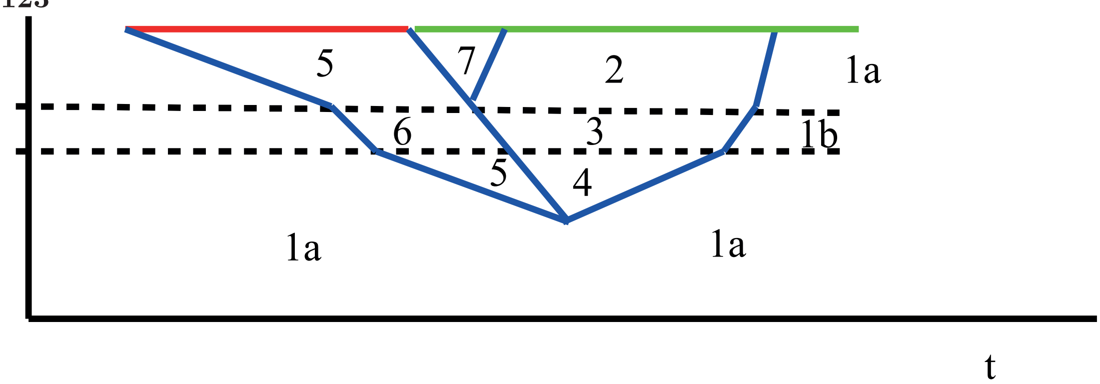
121 2 plots - for each one point. 1 point for the values



122 This is an urban area, so the free flow speed is around 50 km/h. (0.5). The jam density is hardly dependent on the conditions, so around 125 veh/km/lane (0.5) The capacity is a bit lower than in motorway conditions, at around 1200 veh/h/lane (1000-2000 veh/h/lane: 0.5 point). Draw this for a 4 and 1 lane part (0.5 point)

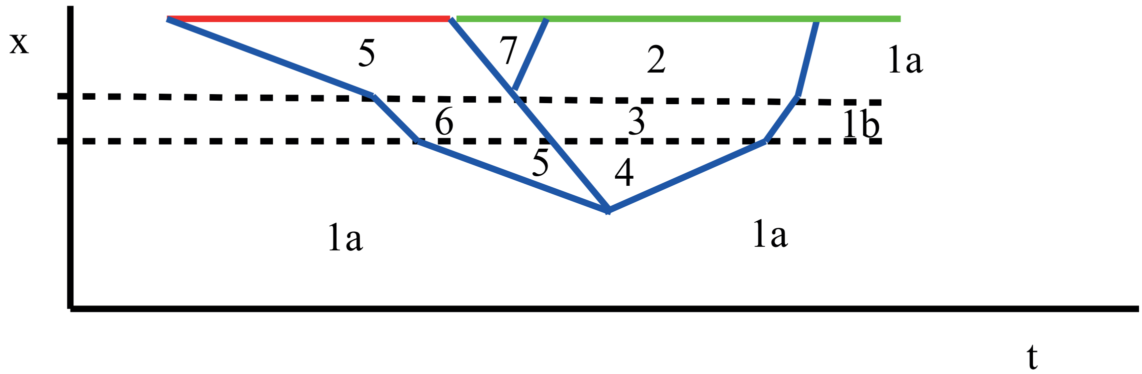


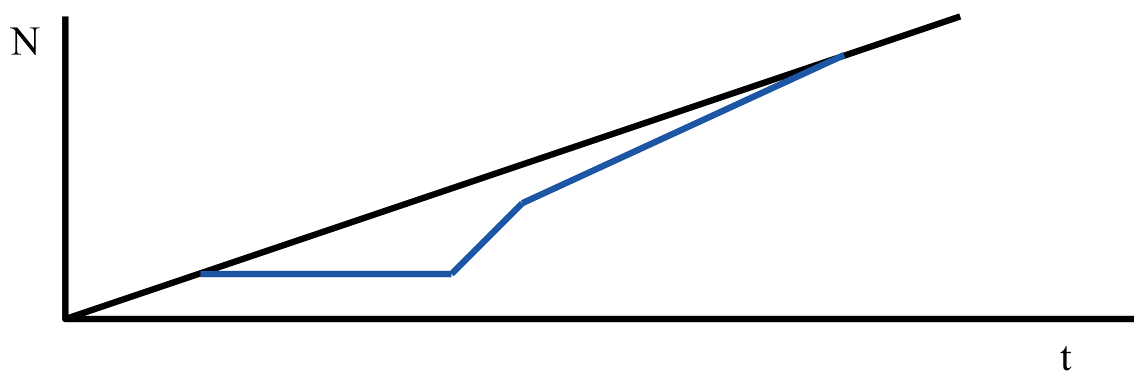
123



If during the green phase the queue completely dissipates, the inflow should be lower than the capacity for the one lane part (1). Starting with the traffic light in green/red: 1 point. During the red phase, the flow is zero, so we get jam density in the 3 lane part (0.5) and the 1 lane part (0.5). Once the traffic light turns green, there is capacity outflow (state 7 - 0.5 point). Once the backward shockwave reaches the part where the crane is, the flow in the part where the crate is, is now limited to the of one lane (0.5). In the downstream part, we now have the same flow in the free flow branch (0.5). An congested traffic state with the same flow occurs upstream of the crane area (1 point) (5.5 points for the traffic states). Now, connect all traffic states (2.5 points).

124 Upstream, the inflow is equal to the flow in 1 (0.5). Downstream, the flow is zero (during the red phase)(0.5), the flow of 7 (0.5), the flow of 2 and the flow of 1 once the queue has been dissolved. (0.5)

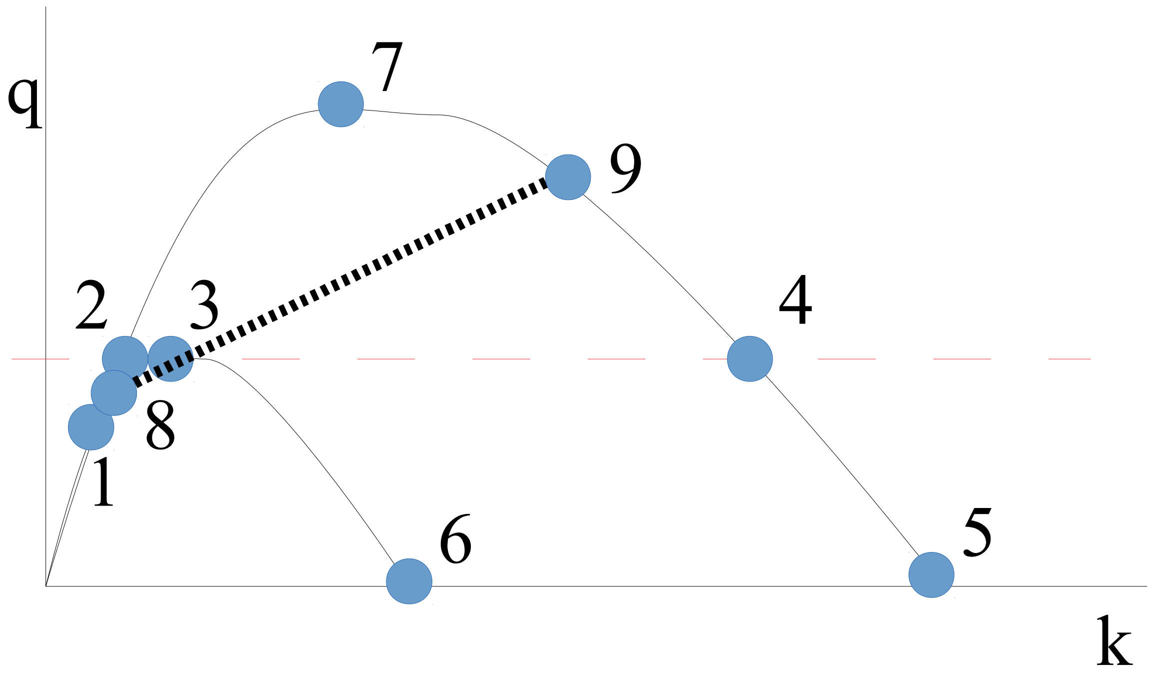


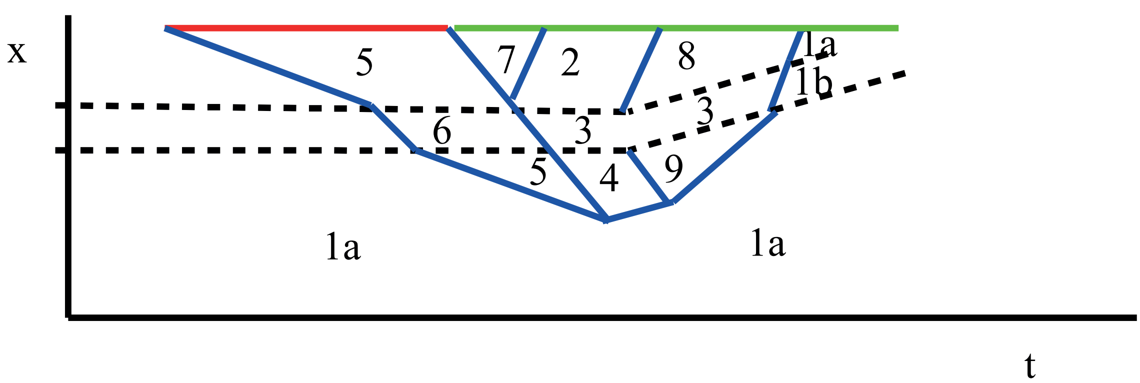


125 The delay is the area between the blue and the black line

126 If the crane is further from the intersection, state 7 lasts longer (1). Hence, the farther upstream the crane is, the longer the steep increase of the blue line (0.5) and the smaller the delay (0.5).

127 Still, the traffic operations next to the moving crane are at capacity conditions as long as there is a queue waiting. (0.5) Hence, we have point 3 (0.5). The crane forms a moving bottleneck and these are shocks which connect to the free flow and congested states on the fundamental diagram downstream and upstream (1 point for mentioning, 1 point for drawing correctly in FD). The remainder of the shocks can now be drawn (1 point for finishing)





128 The Network Fundamental Diagram describes the production (0.5) average flow (0.5) as function of the accumulation (0.5) , i.e. the number of vehicles in a network (0.5) Also possible: the performance (0.5), i.e. the outflow (0.5), as function of the accumulation. Wrong combination of terms and meaning costs 1 point.

129 NFD plot with flattened top (top: 0,5 point).

130 The NFD goes up and down, just like the regular FD, but the top is flattened (1 pt, including graph). The flat top is caused by the fact that not all roads can be operating at capacity (1)

131 Once the roads get more congested, the blocking back causes the traffic to slow down.

132 See the course slides. If the inflow is limited, the production can be higher, and hence the outflow can be higher (1). Without control, the inflow into the protected network can be higher than the outflow out of the protected network(1), thereby steadily increasing the accumulation(1). The higher the accumulation, the lower the outflow, until finally a gridlock situation occurs.

133 A car-following model describes how the position, speed or acceleration of a follower depends on the movement of the leader

134 The relative amount of space (in time! -0.5 if not mentioned) that a non-passenger car vehicle occupies on the road

135 The road could handle 1 passenger car per unit of time (without loss of generality, choose time unit appropriately such that this is correct) or 0.5 truck (0.5 for this reasoning, or any other leading anywhere). Now, this would be 1.05 passenger car of 1.09*0.5=0.545 truck (1). To get the pce value, one has to divide the maximum truck flow over the passenger flow (1): 1.05/0.545 = 1.92.

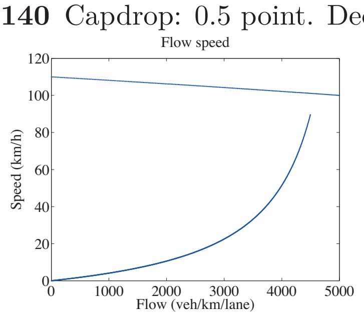
136 This means that the the solution is more exact; a platoon or stop-and-go wave will for instance spread in a Eulerian coordinate system, whereas this will stay together (resp the wave will remain sharp) in the Lagrangian coordinates.

137 Comfortable accelerations lie in the order of 1 m/s2, strong braking in the order of 6 m/s2. A threshold value to accelerate would hence be around 1 m/s2 (0.3 to 2: 1 point)

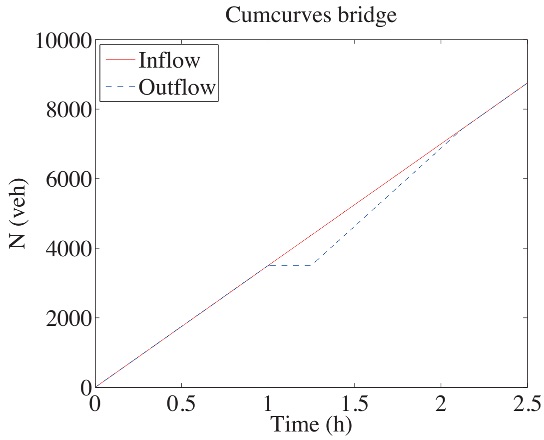
138 It is the politeness – how people value the acceleration of others. This can be assumed to be between irrelevant (value: 0) and just as important as themselves (value: 1)

139 Flow in ped/s/m (max: 0.6-2), density in ped/m2 (max: 4-12 1 point) Base for instance the flow on a 1.5s headway per lane (similar to car traffic), leading to 1/1.5=0.67 ped/s/lane. In 1 meter, at 60cm width per ped, 1,67 fit next to each other (leading to 1 or two lanes). Being partly next to each other, the lane capacity is multiplied by 1.67: cap=.97*1.67=1.1 ped/m/s

140 Capdrop: 0.5 point. Decreasing speed): 0.5 point. General shape: 0.5



141 Free flow speed (100-120 km/h), critical speed(75-95 km/h), critical density(17-30 veh/km/lane), jam density (100-200 veh/km/lane). 1 point for the densities, 1 for the speeds. 1 point for matching numbers (q=ku) at capacity (free flow / critical)

142 

143 In 15 minutes 3500/4=875 vehicles arrive, which are in the queue. After the bridge opens, the queue length reduces (3500 veh/h in 4500 veh/h out)

144 See the cumulative curves. The total delay is the total area between the lines.(1) In this case, the congestion reduction phase lasts 875/(4500-3500)*60 = 52,5 min, (1) so in total congestion lasts 67,5 minutes. (0.5). This is a triangle, which surface is being determined by the height times the width times divided by two. The surface then is 67,5 min * 875 veh/ 2 /60 (min/hr) = 492 veh h. (0.5)

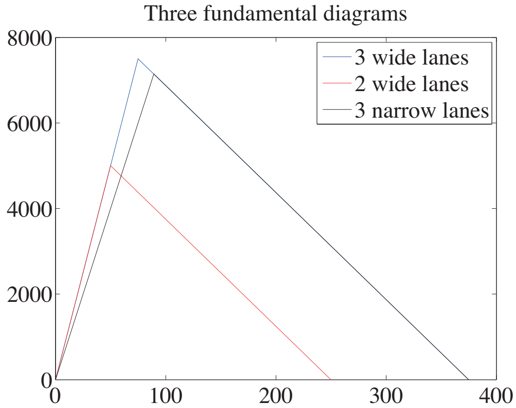
145 From one point different waves occur: a fast one to high speeds, and a backward moving one to lower speeds.

146 At the beginning there is no flow (the first vehicles need some time to arrive - 0.5) and it then gradually increases (0.5) to the capacity (1) which it will asymptotically reach (so no touching!) - 1.

147 This is downstream. The fastest wave travels at free flow speed (0.5). This is derived from the fundamental diagram (0.5). The Greenshields fundamental diagram is a parabola (0.5). To find the parameter (1 pt): the speed decreases linearly with the density. The capacity is found halfway the density range, so at 60 veh/km. The speed there can be determined by v=q/k=5000/60=83 km/h. Since the speed-density relation is linear, the speed at k=0 is twice as high: v0=83*2=166 km/h Other possibility: for the sake of calculation, flip the parabola to a form upside down, and move it to such that the top at (0,0) and increasing flow. The parameters are found by $a(60)^2 = 5000$ (0.5), so $a = 5000/3600 = 1.39 veh/h/(veh/km)^2 = 1.39 km^2/veh/h$ (0.5). The speed is the derivative (0.5), i.e. 2ax at x=60. That means v=2*1.39*60 = 167 km/h. (0.5) Note that it is unlikely that a wave will travel that fast in Dutch motorways. So the wave arrives at 200 m downstream 500/(167/3.6)=11 seconds after the bridge closes.

Because this is a hard question, the final points are divided by 2.

148 For each of the FDs one point. Note the congested branch is uninfluenced for the congested part, so is overlapping with the three lane FD

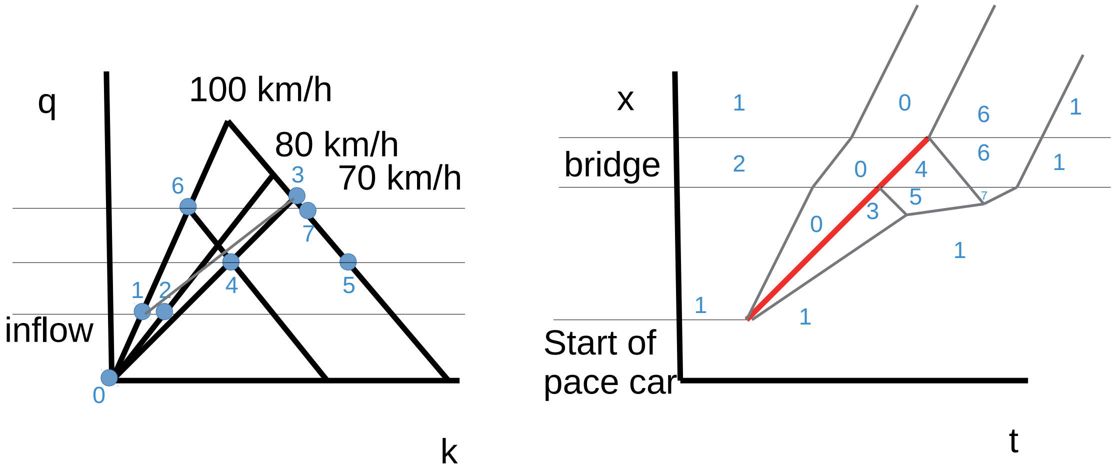


149 The question relates to: where do the free flow line for 80 kms/h and the congested line cross. (0.5) The free flow line is: q=80k (0.5). The congested line is q=100*25-25*(k-25). (0.5) Setting q equal gives kc=29.8 veh/km/lane and q=2381 veh/h/l/lane. (0.5) For three lanes this gives 7129 veh/h. (0.5)

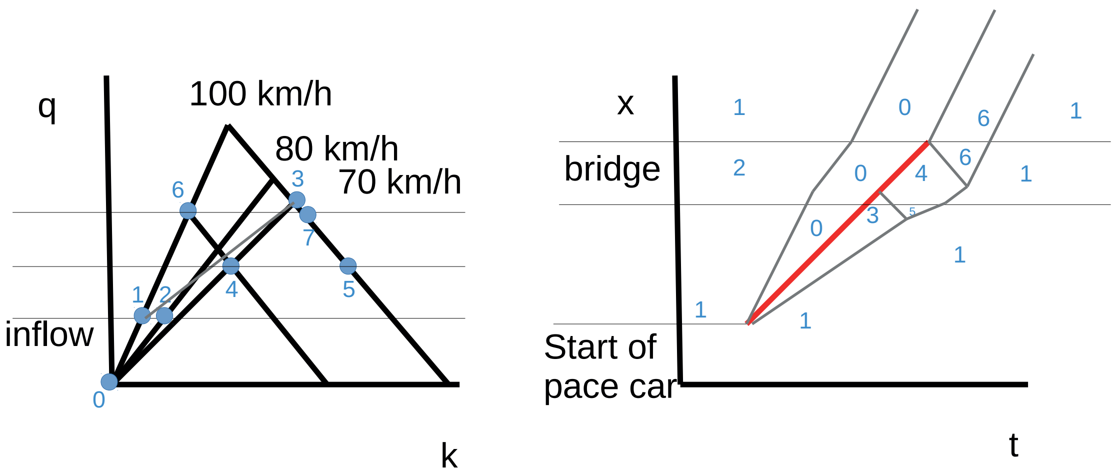
150 In busy conditions, the increased capacity causes less queuing, and hence a lower travel time (0.5). In free flow conditions, the higher speed causes lower travel times. (0.5)

151 The last vehicle downstream of the pace car needs to travel 1000 meters, where the pace car travels 700 meters. Upstream, the traffic has a speed of 100 km/h (0.5), so the time is 1000/(100/3.6)=10*3.6=36 seconds (0.5). The pace car hence has to drive 700 meter in 36 seconds (0.5), hence the speed is 700/36 * 3.6 = 70 km/h. (0.5) (the factor 3.6 is for unit conversion to km/h.

152 Traffic is in free flow conditions (0.5), at a flow of 3000 veh/h (0.5), giving a state in high speeds (state 1) (0.5) and at 80 km/h in the narrow lane section (state 2) (0.5 – lower speeds). At a certain moment, the pace car enters, driving at 70 km/h leaving a gap in front (state 0) (0.5) Upstream there will be a congested state at the fundamental diagram for three wide lanes (0.5) at 70 km/h (0.5) (state 3). The capacity for the two lane part at 70 km/h (the pace car still determines the speed) is point 4 (0.5). This is lower than the local demand, state 3. Hence, there will a congested state upstream (state 5 - 1 pt) with the same flow as the congested state in the two lane section. Once the pace car leaves the road, a flow equal to the capacity of the two lane section exits the congested state 4. The head of this tail is moving backward (just like a head of a queue once the traffic light turns green). Downstream in the 3 lane section, the same flow (capacity flow of the two lane section) is found. Then, two possibilities arise (based on your sketch – no need for exact computation): (1) backward wave gets to the end of the bridge before the forward bridge reaches it. Then, in the three lane section a congested state with the same flow as the capacity in the two lane section is arises. Possibility (2) is that the forward wave is earlier, and in that case the congestion solves from the tail. Drawing the space time diagram: 2 points: (cumulative: 9.5, divide points by 1.5 because it is a hard question)



Solution 1: backward wave gets to the end of the bridge before the forward bridge reaches it

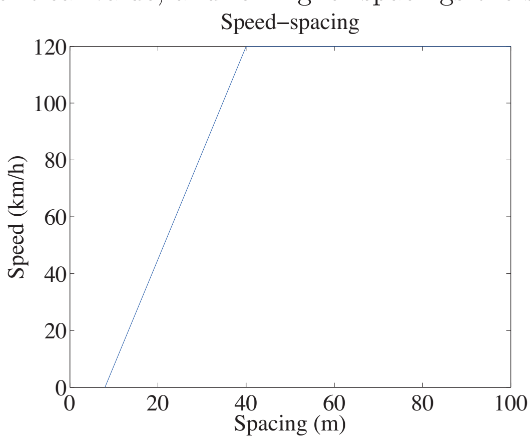


Solution 2: backward wave gets to the end of the bridge after the forward bridge reaches it

153 Synchronized flow

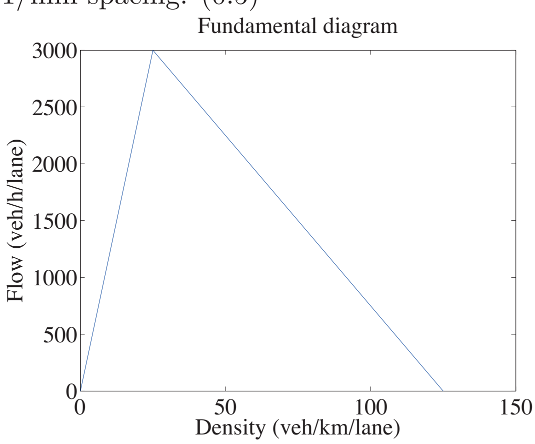
154 Wiedeman states that drivers will not react to minor speed differences if they are below a certain threshold (1). That is similar to the synchronize flow area (0.5) which states that drivers are willing to accept various time headways at different speeds (0.5)

155 Speed increases with increasing spacing (0.5). No values for smaller spacings than a critical value, and for higher spacings the speed is constant (0.5) (at the desired speed).



(1 for graph)

156 Free flow speed is slope (0.5), critical density is 1/critical spacing (0.5), jam density = 1/min spacing. (0.5)

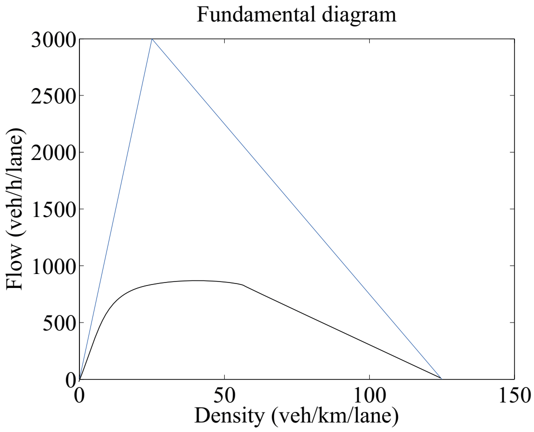


A piecewise linear speed-spacing diagram leads to a triangular fundamental diagram (0.5 – not asked in wording, but a wrong shape will get a reduction in points).

157 Production

158 Accumulation

159 (it is lower and it has a flattened top. (1) Plotting: (1). The right part is mainly theoretical, since one cannot reach that state (not needed for points).



Good reasoning as well: in a network, the speeds are lower, so the overall NFD is lower.

160 The controlled intersection causes headways to be more spread (1). In this case, the negative binomial (1) function is a good way to describe arrivals.

161 If the platoon is stable, it means traffic disturbances damp out within the platoon (1). That means that they cannot grow from platoon to platoon, which is the definition for traffic instability (1)

162 N(x,t) indicates how many vehicles have passed location x at time t

163 With calibration, one finds the best parameter settings for a model (1), whereas in validation, one checks how good this (optimized) model works for another situation (1)

164 An MFD indicates the relation between the outflow out of (or internal flows in) *an area* to the number of vehicles in that area

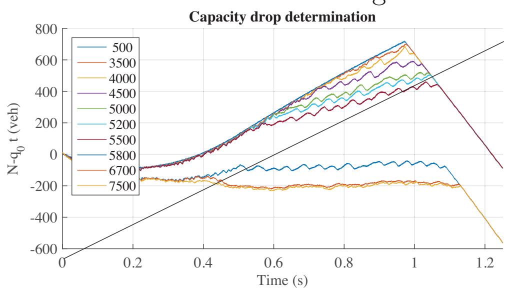
165 Capacity is given in ped/m/s

166 The capacity indicates the number of pedestrians per unit of time per unit of width.

167 (1) One has to estimate the number of vehicles in the road section at the start, (2) there is a increasing error from miscounts (3 – less of an issue) one does not have observations from the vehicles leaving or entering the road section (one has not here in this case either – therefore 0.5 point)

168 The queue starts at the bottleneck, since there is delay between x=6700 and x=5800 (there is a distance between the moved slanted cumulative curves of 6700 and 5800) (1). The delay spills back to approximately 3500m (there is still a little delay between 3500 and 4000m) (1). The congestion solves by a lower demand, seen from the demand curve (only that: 0.5 pt) and the fact that the delay at the more upstream locations solves earlier (add that: 0.5).

169 The most accurate reading is from the slanted curves.



Before congestion it has a slope of approximately (650- -400)/(1.2 - 0.2)=1050 veh/h (1). This should be added to the offset to get the flow (0.5): 1050+3900 = 5050 veh/h (0.5) – reading from the non slanted curves 1 pt. Reading the maximum flow from the highest value on the three lane section: 0.5 point (it is a flow, but not related to the capacity)

170 The difference between the flow before congestion sets in and after it has set in is the capacity drop (1). After congestion has set in, the slope is approximately 0 (0.5) (or equalling a flow of 3900 veh/h). The flow 1050 veh/h lower than before congestion, which is the value for the capacity drop (0.5)

171 Yes, the delay would differ since the delay of the vehicles leaving at the offramp would not be taken into account (1). Hence, the delay computed with a vertical queuing model would be less (1)

172 The partial densities can be calculated by k=q/v per class (1), leading to k1=500/120=4.2 veh/km and k2=500/90=5.6 veh/km (1). The total density is the sum of these two (0.5), 9.7 veh/km (0.5)

173 For the space mean speed, one can (a) average the paces or (b) Put weight on the speed values the speeds using the densities (either method: 1 point). (a) The result is 1/((1/90+1/120)/2) (0.5) = 102.8 km/h (0.5) (b) The result is 4.2 * 120+ 5.6 * 90 /(4.2+5.6) (0.5 point) = 102.8 km/h

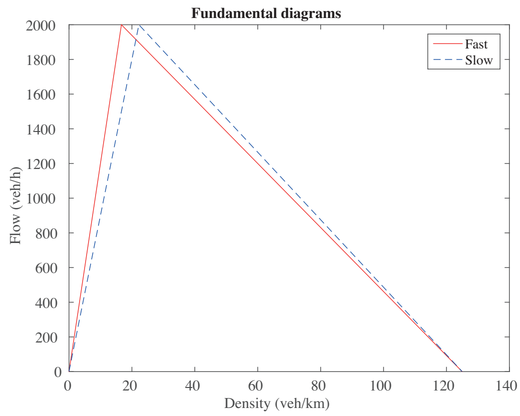
174 (i) For the time mean speed, one observes as many slow as fast vehicles (equal flows). (1) Hence, the mean speed is (90+120)/2 (0.5 pt) =105 km/h (0.5 pt). If b is the correct answer to c and c the correct answer to b: 1 point total.

175 $k_{\text{Edie}} = \frac{\text{Total time spent}}{\text{Area in space-time}}$ (1 point)

176 Use the equation qrel=kfast*vrel (1). kfast=4.2 veh/km (0.5) and vrel=120-60=60 km/h (1). Therefore,qrel=4.2*60=250 veh/h

177 Calculate – in the same way as above – the flow of the slow moving vehicles: qslow=5.6*20=113 veh/h (2). Now, average the speeds: vmean=(qslow*80+qfast*120)/(qfast+qslow) (1 point) = 107.6 (0.5)

178 The flows passing a fixed observer are equal, and relative speed difference with the fast vehicles is larger. (1) Hence, using qrel=k*vrel, the flow of the fast vehicles becomes relative larger (1).



179
1 point for each line

180 1 since the capacities of both vehicle classes are equal

181 Newell's car-following model is described by the equation

$$x_{\text{i}}(t) = x_{\text{i-1}}(t - \tau) - s_{\text{jam}} \tag{B-12}$$

(1 point) τ can be found as 1/capacity (1) 1/2000 (/h), or 3600/2000=1.8 s (0.5). s_{jam} can be found as 1/jam density (1), or 1/125 (/km), or 1000/125=8 m (0.5)

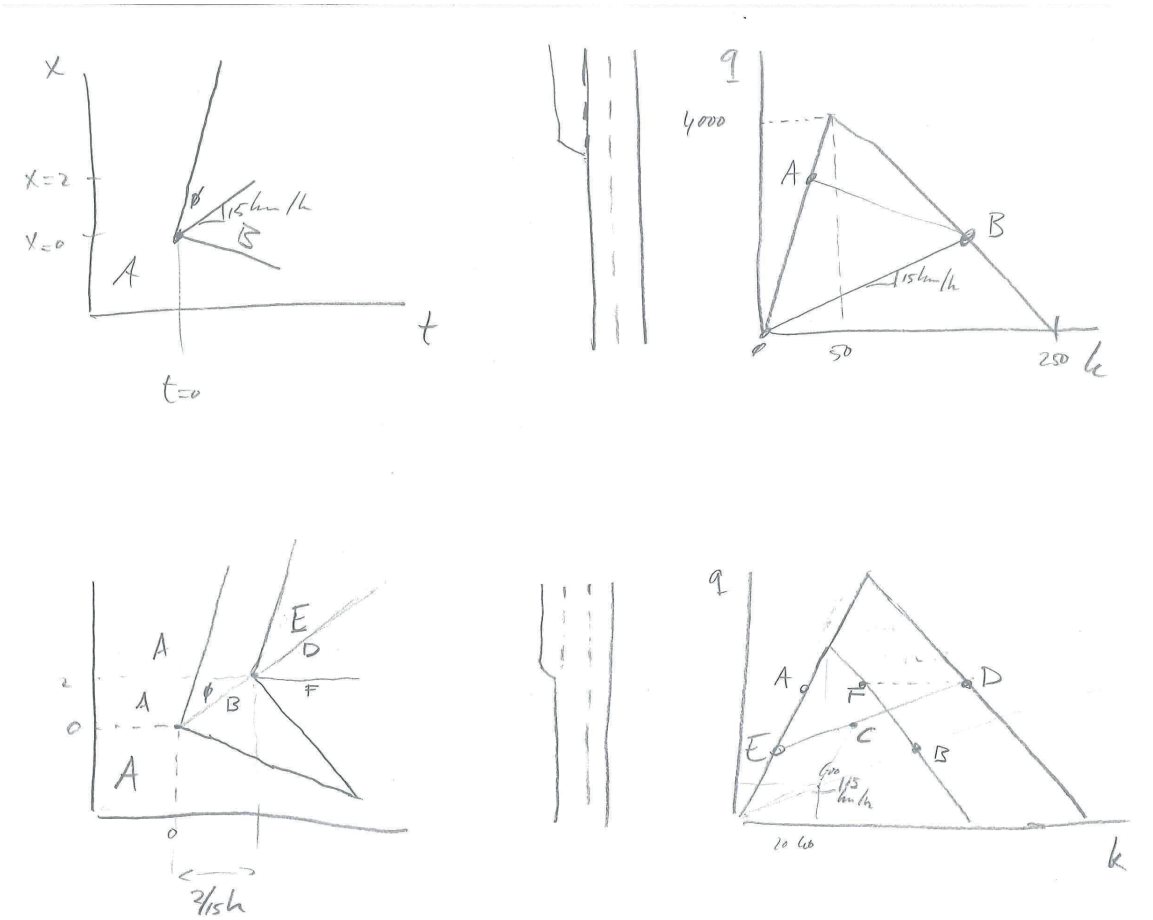
182 Slugs and Rabbits

183 Synchronised flow

184 Only the two lane part is relevant here, so draw that fundamental diagram (1 point for the FD with the correct axes (per roadway, rather than lane). The initial conditions (A) can be indicated in the FD (0.5). Now, the traffic upstream of the truck is driving at 15 km/h (0.5), hence we get a congested state at the line q=15*k, leading to B (0.5 point). The trajectory of the truck can be put in xt, driving at 15 km/h (1 point). The congested state B spills backwards with a speed of which the slope equals the slope of A-B in the fundamental diagram (1 point). Downstream of the truck, there is an empty road (0.5) and the matching shockwave between that and state A (0.5)

185 Draw the fundamental diagrams (1 point), and the capacity point C (0.5). The truck forms a bottleneck, and hence moving at 15 km/h, (0.5) through point C, there is a line with a slope of 15 km/h (0.5), connecting downstream point E (0.5) and upstream point D (0.5). Shockwaves 0-E (0.5) and ED (0.5) can now be drawn. In the two lane section, state is at the two-lane fundamental diagram: a congested state at the same flow of D – point F (1 point).

Shock wave BF at the speed of the line BF in the FD (0.5 point).


x
x=2
x=0
A
B
t
t=0
q
4000
A
B
50
250
k
E
D
A
B
F
A
A
2
0
0
2/15h
q
A
D
F
C
E
B
k

186 Exponential

187 Binomial distribution

188

$$f_i = \frac{q_i}{q} = \frac{q_i}{\sum_{\text{all lanes}} q_i} = \frac{1/\langle h_i \rangle}{\sum_{\text{all lanes}} 1/\langle h_i \rangle} \qquad \text{(B-13)}$$

(The step at each equal sign is 1 point)

189 Passenger car equivalent

190 The demands are equal, hence the flows of slugs and rabbits are equal (0.5). They are separated in free flow conditions (1), so the flow in both lanes is equal (0.5)

191 If a platoon is unstable (0.5) a small disturbance can grow over the platoon (0.5). If the disturbance also grows beyond the platoon into the next platoon (0.5) this is called platoon instability (0.5). A small disturbance hence can grow until vehicles come to a complete stop (0.5).

192 Breakdowns occur stochastically, and with a higher flow, the probability of a breakdown is higher. Then, there is not a single value for capacity.

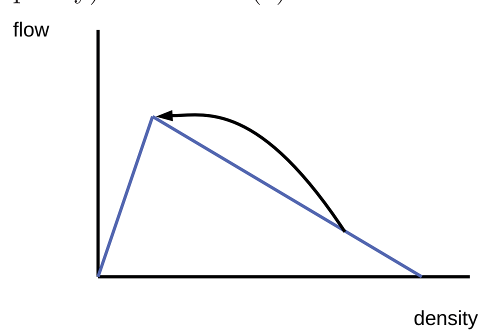
193 It means that the predicted traffic states are not exact; shock waves get spread out over space if time progresses (1)

194 Traffic must be stationary and homogeneous

195 Traffic characteristics move with the slope of the fundamental diagram. (1 pt, word characteristics not needed) Simplifying to a triangular fundamental diagram, they move either forward with the free flow speed or backward with the wave speed. A parallelogram with slopes free speed and wave speeds (1) aggregates hence a more homogeneous traffic states.

196 Edie's (generalised) definitions

197 Due to this multi-anticipation, drivers will accelerate earlier out of the congested state (1) (even before the density will go down as much as usual – i.e., the FD). Hence, the speed and thus the flow go up faster than normal (1). The end situation (capacity) is the same(1). We



hence get a transition line above the congested branch (1: sketch).

198 The flow through the bottleneck is equal to the flow in the jam (1). This is found by q=ku (1). k=1000/11*3=273 (0.5 + 0.5 for the *3 part) and u=8 (0.5). Hence, q= 273*8=2182 veh/h (0.5)

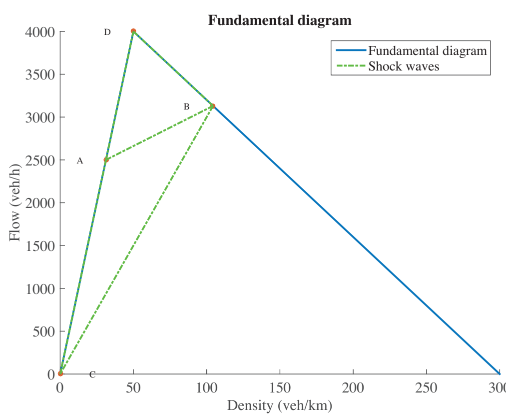
199 The capacity point for the three lane part is three times the capacity of the one lane part (0.5), which is the flow through the bottleneck (0.5). So the capacity is 3*2128=6545 veh/h (0.5). The critical density is hence 6545/cfree=65 veh/km (1). The wave speed can be found by realising the point q=2182 and k=272 lied on the FD (0.5). Now, the jam density can be found in various ways. For instance: the wave speed is dq/dk, with dq=-(6545-2182)=-4364 and dk=272-65=207 w=-21km/h. kj=kc+C/w=67+3*2182/21=375 veh/km (1)

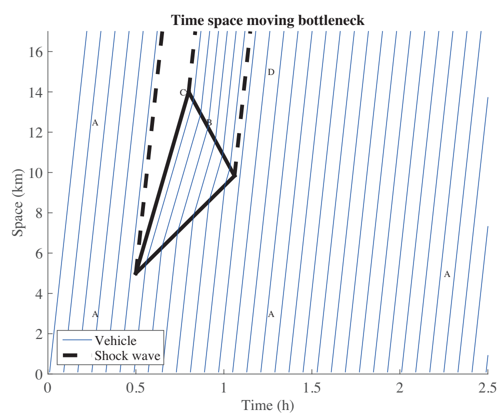
200 This question relates to the flow in the two lane section in a congested state (0.5) for a flow of 2182 (0.5). The density is exactly between the critical density (2/3*65=44 veh/km) and the jam density (2/3*375 = 250 veh/km), so at 294/2=150-3=147 veh/km. (1) The speed then is u=q/k=2182/147=15 km/h (1).

201 Since the traffic state is stationary, the shock waves must have slope zero, so the flows must be equal. (1)

202 For all stationary states, the flow is the same for all locations. Hence, the average flow is the same as the outflow. (1) This increases as in the FD up to the critical density, and then remains at that maximum (1).

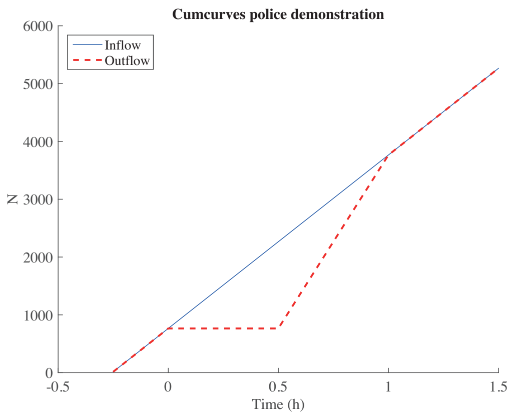
203 Sketch the FD (1). Draw a line with slope - speed - 30 km/h from the origin (0.5). This is the line where the congested state must lie, i.e. at the intersection of this line and the FD – indicate in FD (1). The original traffic situation is at a demand level higher than this flow (given) at the free flow branch (0.5). After the police cars speed up, the state returns to capacity (1). Indicate the connections in the FD (1) and in space-time (1)
The following graphs are sketches, and the units are not matching the question at hand.





204 Calculate the intersection point of the line with slope v0 and the FD (1) Description of the congested FD: q=C-C*(k-kc)/(kj-kc) (1). Fill in values, and solve:
$C - C * (k - kc)/(kj - kc) = v0k$
k=104 veh/km (1) q=v0 k (0.5)= 30*104 = 3130 veh/h (0.5)

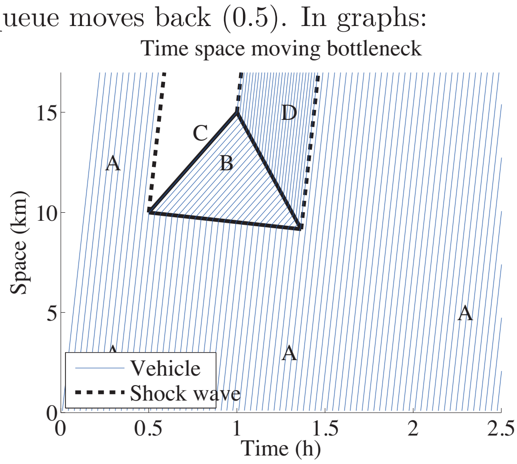
205 Inflow is constant (0.5), so a straight line with slope of demand (0.5). Outflow is equal to the inflow until the moment that the police cars enter (0.5). From that moment in time, the outflow hence slope is 0 (1). After they leave, the flow is the capacity (not asked), so steeper outflow (0.5), up to the moment they are at the same vehicle level (0.5), after which they are at the same line (0.5)

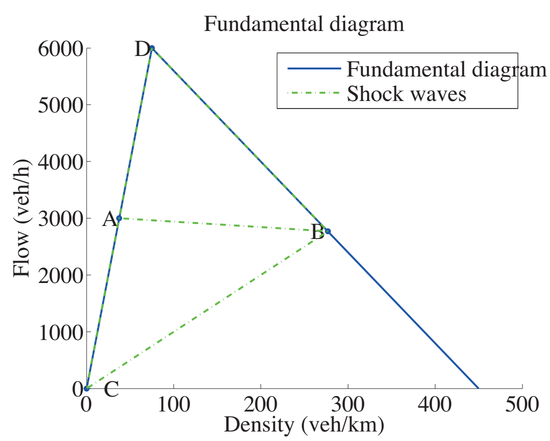


206 The delay is given by the area between the two lines in the graph above (1). The max nr of vehicles in the queue ($N_{\max}$) is $d * q$in. The time to solve is then $t_{\text{solve}} = N_{\max}/(C - q\text{in}) = d * q\text{in}/(C - q\text{in})$ (1). The area is $1/2 * N_{\max} * (d + t_{\text{solve}}$ (0.5). Replacing the terms gives$D = d * q\text{in} * (d + d * q\text{in}/(C - q\text{in})$. Further simplification (proof) leads to

207 The delay approaches infinity if the demand approaches the capacity. (1). That is because the jam will never solve if the demand is equal to the outflow (1), and all following vehicles will get a delay (1).

208 This is basically the same as the question b, with the difference that the demand is above the congested flow on the free flow branch of the FD (0.5). The segment connecting this point to the congested traffic state hence goes down (0.5), and this means the tail of the queue moves back (0.5). In graphs:





209 It does – it is based on the flows and the gaps, and they remain the same.

210 See the cumulative curves: the area is the delay (no points - awarded earlier). This increases proportional with both the number of vehicles in the queue, as well as with the delay for the first vehicle (1), which are both proportional to the duration of the demonstration (1), so in total that means a quadratic relationship (1)

211 Expected headway: 5 seconds. (1). Flow per lane: 3600/5 peds/h. (1) Flow total: 3600/5*2 = 1440 ped/h. (1)

212 In total the number of cyclists is the area under the graph (1), calculated by $\int_0^(2points)q(t)dt$, by geometry, leading to: 2 (min) * q0 *1/2 = q0 [cyclist/min] (1). This equals 300 cyclists (0.5). Therefore, q0=300 cyclist/min

213 The queue grows as long as q is larger than C (0.5), hence we need to find the number of cyclists that have passed at the moment q reduces to a value lower than C (0.5). It is (1-(C/q0)) minutes (0.5) *after the peak of the flow* (0.5) at one minute, so $t_E = (1-(C/q0))+1 = 2-C/q0$. The amount of cyclist which then have passed is $N_E = \int_0^{t_E} q(t)dt$, (1) which equals (or from geometry)

$$
\begin{aligned}
N_E &= 1/2 * q0 + (t_E - 1) * C + 1/2 * (t_E - 1) * (q0 - C) && \text{(B-14)}\\
&= 1/2 * q0 + (1 - C/q0) * C + 1/2 * (1 - C/q0) * (q0 - C) && \text{(B-15)}\\
&= 1/2 * q0 + (1 - C/q0)(C + 1/2(q0 - C)) && \text{(B-16)}
\end{aligned}
$$

(1)

214 Production

215 Yes, we can. The critical density for the Greenshields fundamental diagram is halfway the density range (0.5), hence at 50 veh/km (0.5). The speed at that density is half the free flow speed (linear relationship, 0.5 pt), being 40 km/h (0.5 pt). The capacity hence is q=ku (0.5 pt) = 40*50 = 2000 veh/h (0.5 pt)

216 An acceleration fan occurs at the head of a jam. Since characteristics of the lowest density move upstream, and characteristics of highest density move downstream, the sharp shock at the head of the queue (1) will spread out (1).

217 The density should be derived from the quotient of the flow and the space mean speed (1). The time mean speed overestimates the space mean speed (1), hence the density is underestimated (1)

218 Locally stable traffic means a vehicle will reduce the amplitude of it speed variation over time; from vehicle to vehicle, the amplitude can still grow, so platoons can be unstable.

219 In a stochastic simulation package (1 point), the traffic situation is depending on a random draw. You should do multiple simulations to check to which extent the results depend on the random draw.

220 FD holds in equilibrium, hence Δv equals 0, and a equals zero. This means (first equation) that $s = s^*$. (1). The second equation gives a relation between speed and density. We can rewrite this to

$$v = (s - x_0)/T \tag{B-17}$$

(1 pt) Now substituting s=1/k gives:

$$v = (1/k - x_0)/T \tag{B-18}$$

(1 pt) Multiplying both sides with k gives the relation q=q(k)

$$vk = q = k(1/k - x_0)/T = 1/T - kx_0/T \tag{B-19}$$

(1 pt)

221 After validation, it is known to which extent the model can be applied in other conditions and what it's quality is

222 Since in the congested branch (equation above) q decreases with k, the capacity is found at the intersection of the congested branch and the free flow speed. We hence have to solve

$$v = q/k = k(1/k - x_0)/T/k = v_f = (1/k - x_0)/T \tag{B-20}$$

(1) Solving this to k gives

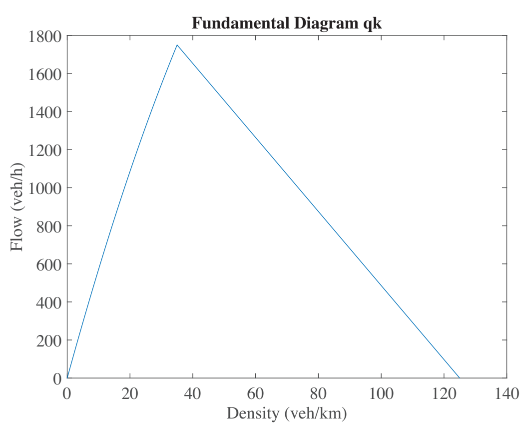
$$k = 1/(Tv_f + x_0) \tag{B-21}$$

(1) Using the parameters, we find k=1/(1*30+20)=1/50 veh/m. (0.5) Computing q from this gives q=kv=1/50 * 30 = 30/50 =3/5 veh/s (1) = 2160 veh/h (0.5)

223 Consider the FD in speed-density. Its form does not change, only the distance changes hence the values at the density axis are halved (0.5 pt). Moreover, the speed reduces to half, hence the values on the speed axis are also halved (0.5 pt). The effect on capacity is that the product of density and speed (1 point) is half times half = 1 quarter of the original capacity (1 point). The capacity is hence 2160/4=1080/2=540 veh/h

224 Smulders fundamental diagram

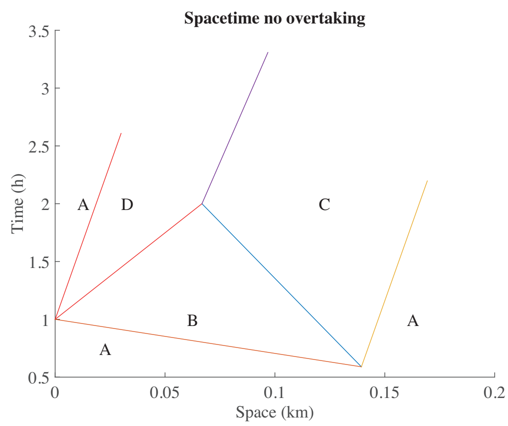
225 Capacity is at the point where the congested branch starts (0.5); reading out the graph gives 35 veh/km and 50 km/h (0.5), so a capacity of (q=ku, 0.5 point), 35*50 = 1750 veh/h

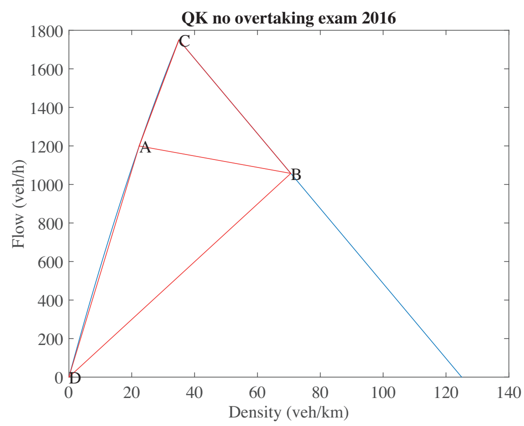


226
There is a slight curvature in the free flow branch, and a straight congested branch

227 This is the traffic state at 15 km/h (0.5), for which the FDs can be used (0.5). Hence, read out the FD at 15 km/h (0.5), yielding a density of approximately 70,5 veh/km (0.5). Using q=ku (0.5), this gives a flow of approximately 1060 veh/h

228 First, find the traffic state upstream of the moving bottleneck, which is found by constructing the intersection point of the line q=15k and the fundamental diagram (1 point, point B). Downstream of the MB is an empty road (state D, 1 point). Inflow point is at the free flow branch (0.5 point), and after the MB leaves the road, the state is at capacity (1 point).
Constructing the space time: start with the MB (1 point), construct AB, AC and CB (1 point), finishing 0.5 point.





229 A driver will change lanes whenever the utility for changing lanes is higher then staying in the current lane (1). The utility consists of the accelerations (no points, mentioned), for the driver itself and for the vehicles directly influenced (n and o) (1), where the latter two are discounted using a politeness factor $\mathcal{P}$ (1)

230 The values are accelerations so should be given in m/s^2 (1 point). a_{bias} should be higher than a_{th}, otherwise no lane change to the right. Accelerations in the order of 1 m/s^2 are sizeable (i.e., quite some throttle opening), so values should be in the order of one thenth of that. Values between 0.01 and 1 are given 1 point (Treiber and Kesting propose 0.3 and 0.1 m/s^2 for a_{bias} and a_{th} respectively.)

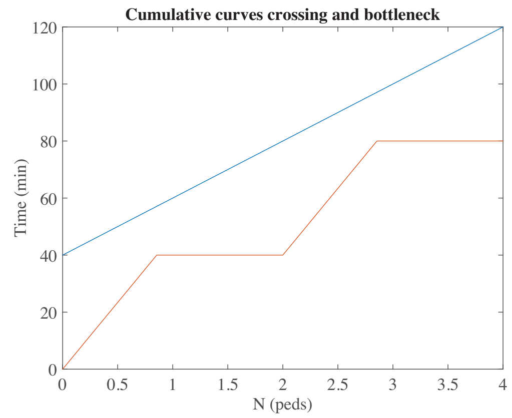
231 Poisson

232 1200 peds/h = 1200/60 peds/min = 20 peds/min. In 120 seconds = 2 mins 2*20 = 40 peds arrive (and cross) (1 point for flow times cycle time, 0.5 point for unit conversions, 0.5 point for end answer)

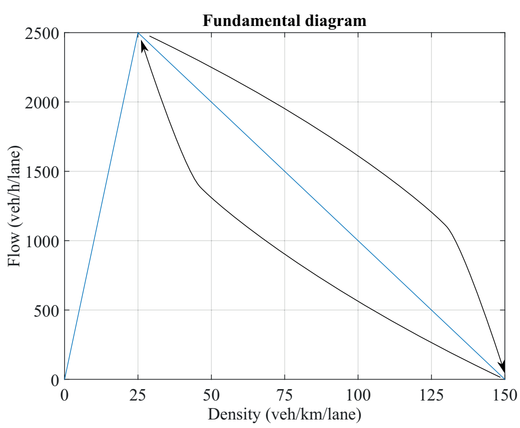
233 Flow in peds/m/s and density in peds/m^2 (1 point for each)

234 The capacity (read from graph) is 1.22 peds/m/s (1 point for using the capacity value), so the capacity for the bottleneck is 0.6*1.22 = 0.732 peds/s. (1). It will hence take 40 peds/ 0.732 peds/s (0.5) = 54 seconds (0.5)to pass the bottleneck.

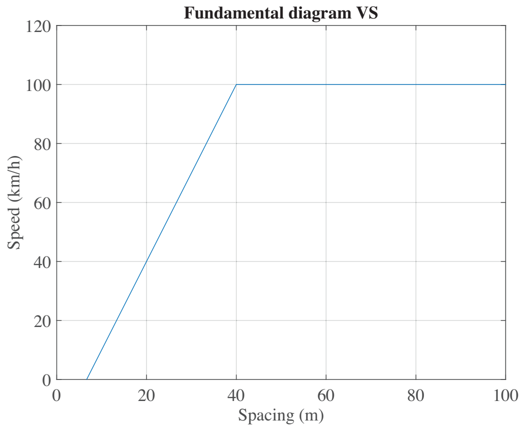
235 Starting with the hint: at the start, 40 pedestrians have crossed A and not B (see question b). (0.5) So start the curves with 40 pedestrians difference (0.5). Then, the inflow steadily increases with 1200 peds/h, or 40 peds/cycle (1). For the outflow, the queue discharges at 0.732 peds/s up to the moment that 40 pedestrians have exited, and remains constant then (1). (Indeed, by that time other pedestrians are waiting at the traffic light again, so the lines do not touch.



236 Consider for instance an acceleration from low speeds to high speeds. Traffic starts right bottom, and then the density decreases. However, since the speed of the vehicle is not adapted, it is still low, so it curves counter-clockwise to the top. (1) The other way round, under deceleration, consider a point near the top. Then, first the density increases (traffic states moves to the right), and the speed is still higher than for the newly increased density. Hence, the traffic state moves counter-clockwise down.(0.5 for a second reasoning)

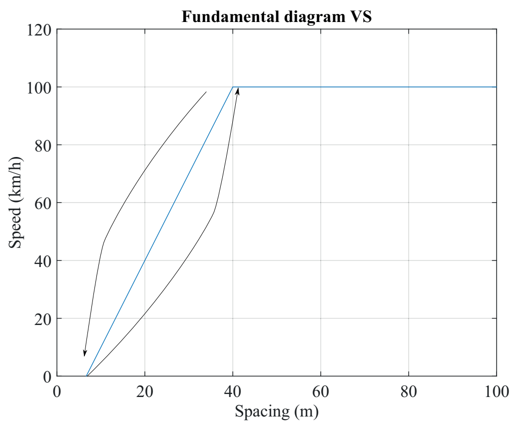


237 Triangular, so the free flow speed is fixed (at 2500/25=100 km/h). Straight line in congested branch means q and k are linear, say form q=1/r-wk. Divide both sides by k, and use substitution s=1/k: q/k=v=1/k 1/r-w = 1/r s - w, so also v and s are linear. From the fundamental diagram, find 1/r= 3000 veh/h (point where the congested line would intersect the y-axis) and w= 2500/125=20km/h. (Also accepted: triangular fundamental diagram in qk leads to a piecewise linear fundamental diagram in vs, with smin=1/kjam, scrit=1/kcrit and vmax=qcrit/kcrit). Plotting the minimum of congested speed and free flow speed gives:



(1pt)

238



239 Poisson distribution

240 Every truck counts for two vehicles in determining the capacity.

241 If we relate it to time headways (1) of a passenger car without queue warning systems, we get for cars a pce=1/1.05 and for trucks 2/1.09. (1). The new pce value is the ratio of the two (0.5): 2/1.09/(1/1.05)=1.05/1.09*2=1.92 (0.5)

242 It means that the edge of a traffic jam flattens (contrary to predicted with shockwave theory)(1)

243 It means that after a braking manouever of a leader, a follower will not find the same speed, but keeps continuing adjusting speed (0.5) in an increasing amplitude, i.e. going to higher and lower speed for every iteration (0.5).

244 Yes, if the inflow is equal to the flow in the queue behind the moving bottleneck, the tail remains stationary

245 This is not possible. For Greenshields, the critical density is half the jam density (1). Since this is not the case here, we cannot draw. (1)
A linear FD in VK is awarded 1/2 point.

246 Stop-and-go waves travel in the opposite direction of traffic, so traffic is moving top-down. No point rewarded if the tail of the queue of the accident is indicated.

247 Given one specific traffic jam and its slope (choose one, 1 point), we find the length of the shown road is around the distance travelled in 2/3 of an hour. Since stop-and-go waves travel with approximately 18 km/h, this equals or 12 km (1).

248 The figure shows an accident at the road at the second tick from the bottom. This is not the fixed bottleneck. The fixed bottleneck is located at around the first tick from the bottom. The capacity drop would be visible at a location just downstream of the fixed bottleneck (0.5). There, the flow would be higher before congestion has set than after congestion has set in (1). In the current graph, this is not visible. (0.5)

249 Rabbits take the lane where the speed is the highest.

250 The density is equal (from question,0.5), and the flow is proportional to density times speed (0.5), i.e. a normalized flow of 80 of slugs and 100 of rabbits(0.5). That means 100/180 = 5/9 of the flow is rabbits(0.5).

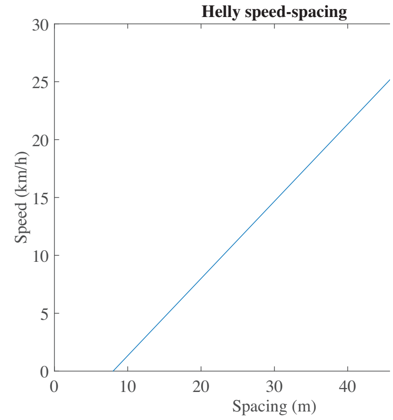
251 Weighted average speed proportional to the flows (1): 5/9*100+4/9*80=91 km/h.
Computing space mean speed, or assuming flows are equal gives no points.

252 Edie's generalised definitions suggest that the speed is the total distance covered divided by the total time spent. TTD is the sum of all distance covered by all vehicles within a predefined area in space-time; TTS is the sum of all time spent of all vehicles within a predefined area in space-time.
Only the equations do not give any points.

253 The model prescribes an acceleration based on a speed difference (first term, sensitivity α) (1) and a difference between the spacing $(x_{i-1}(t-\tau) - x_i(t-\tau)$ and a desired spacing $(s_0 + Tv_i(t-\tau))$ (second term, sensitivity β.) (1) The desired spacing is linearly increasing with the speed of the vehicle.

254 FD is homogeneous and stationary conditions (0.5), meaning $v_i = v_{i-1} = u$ (homogeneous, 0.5), and a=0 (stationary: 0.5). Hence, $x_{i-1}(t-\tau) - x_i(t-\tau) - s_0 - Tu$ (0.5 – 2 points up to now). Spacing can be obtained by $s(t-\tau) = x_{i-1}(t-\tau) - x_i(t-\tau)$ (0.5) and stationarity $s(t) = s(t-\tau) = s$ (0.5). Then, $a = 0 = 0 + \beta\,(s - s_0 - Tu)$, or $u = (s - s_0)/T$. (1) Plotting gives a linear increase of u as function of spacing

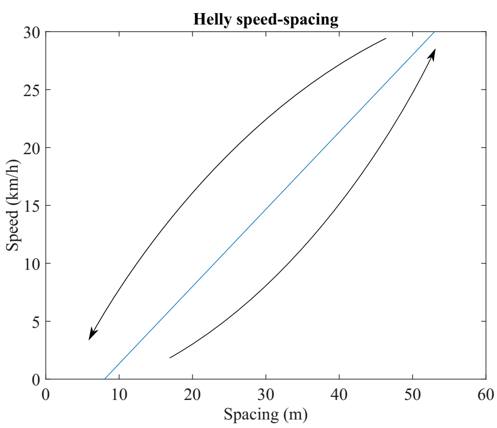
Explicitly making the assumptions is essential. A non-motivated sketch in vt gives no points;



starting from $u = (s - s_0)/T$ and plotting that will provide 1 point in total.

255 The delay τ causes a delayed reaction. If the leader accelerates (going right in the spacing), the follower goes later up in speed (1); similar reasoning for decelerating (0.5), so we obtain a anti-clockwise hysteresis loop (0.5).

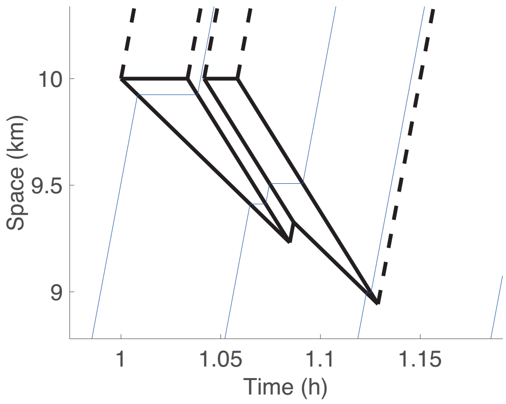
Plotting arrows based on an assumption delayed/anticipation, or providing both will give 1 point.

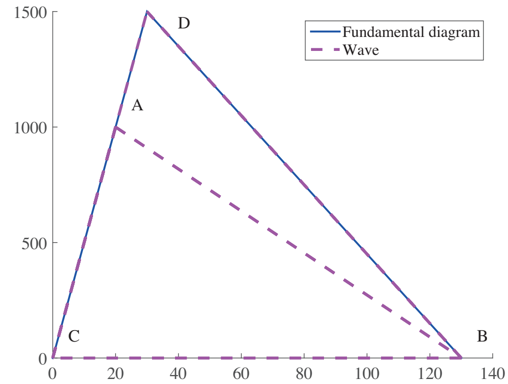


256 Starting from $s = s_0 + Tu$, we apply $s = 1/k$ (0.5) and $u = q/k$ (0.5). We obtain $1/k = s_0 + Tq/k$, rewritten to $1 = ks_0 + Tq$, or $q = (1 - ks_0)/T$ (1).

258 The critical density is 1500/50=30veh/km (0.5); The jam density is 1500/15=100 veh/km higher than the critical density (1), hence the jam density is 30+100=130 veh/km (0.5)

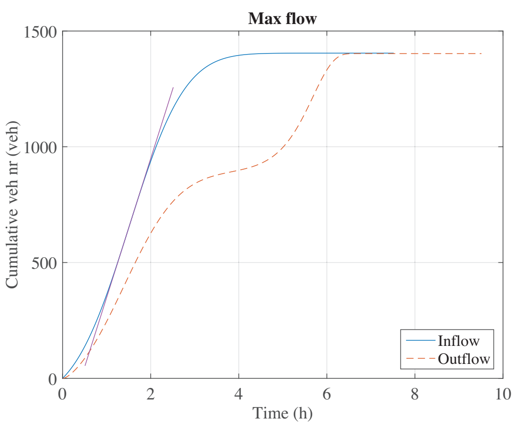
259 Indicate the inflow state in FD (A, 0.5). Then, the first light causes an empty congested state downstream (state C, 0.5) and a state at jam density upstream (B, 0.5). This creates three waves: AB, AC, BC (0.5 each; 1.5 total). When the traffic light turns green, a capacity state D (0.5) emerges, with shock waves CD and BD (0.5 point each – total so far: 4.5). The next red phase causes again states B (0.5) and C (0.5). However, this now gives waves BD (0.5) and CD (0.5). When waves AB and BD of the first traffic light join, wave AD emerges, up to the moment that this intersects with BD (0.5). Then, a wave AB continues (1). Finishing goes in the same way as finishing the first red phase: state D, waves CD, BD (0.5). When wave DB intersects with BD, the wave AD occurs and moves forward (0.5).

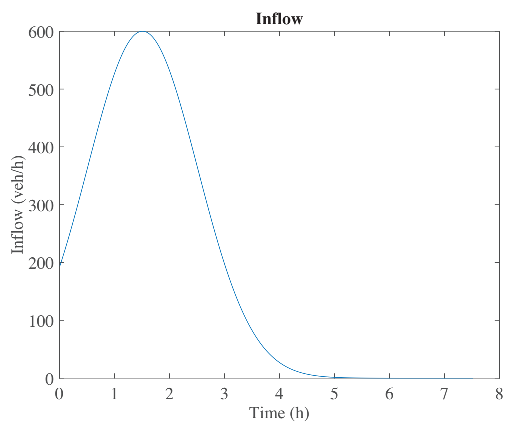




For this answer, the graphs have been constructed automatically. In your answer, provide the reasoning as above. Moreover, place letters in within the space-time diagram. The dotted shock waves should be drawn just as the others. Trajectories are not needed (not part of the question)

260 The flow is the slope of the cumulative inflow, which first increases (0.5) and then decreases, see graph above (0.5). The maximum flow is the highest slope, estimated at 600 veh/h (0.5).
Half a point reduction if the line does not increase at the beginning.

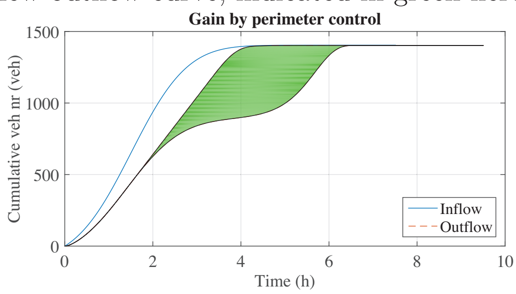




261 The outflow curve increases when traffic starts to accumulate (1). The slope of the line is the outflow, given by the accumulation. If the accumulation increases to values over capacity, the outflow decreases, yielding a flatter outflow curve (1). Only when the inflow reduces to values under the low outflow, the accumulation decreases and the outflow increases again (1).

262 By limiting the inflow and not exceeding the critical accumulation, an outflow of 600 veh can be maintained (1); the remaining queuing is outside the network (1).

263 Whenever there are too many vehicles in the network, we assume they can be queued outside the network (perimeter control), and the outflow will be 400 veh/h, i.e. the capacity outflow from the MFD (0.5). The cumulative outflow under perimeter control hence keeps increasing with 400 veh/h whenever there is an excess number of vehicles between cumulative inflow and outflow(0.5). The amount of travel time gained is the area between the old and new outflow curve, indicated in green here.



264 No, he can be moving in the y-direction

265 The SFM describes for individual pedestrians (hence, microscopic, 1) how they move. A "force" changes their direction: they are attracted by their destination (1), and repelled by obstacles (0.5) and other pedestrians (0.5).

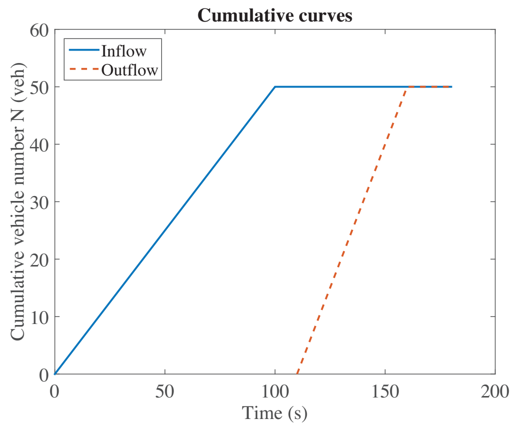
266 In a cellular automaton model, the space is discretized (1).

267 1.22 peds/m/s. Unit, 0.5 point. Value between 0.6 and 2: 0.5 point.

268 One can only see the top they all operate at capacity. (0.5) Whereas this is theoretically possible, in practice this will not happen (0.5). [Additional/alternative: it can only happen if all roads have the same FD, which is unlikely to be the case (1) – 1 point max]

269 A characteristic indicate a traffic state. At one moment in time, a location can only have one traffic state, hence characteristics of different traffic states cannot cross. (1) For the same traffic state: they have the same speed (because fully determined by the state), hence they do not get to each other; moreover, once they are at the same location, they will leave as one at one speed (because fully determined by the traffic state)(1).

270 The number of arriving vehicles in a period of time (1 point; no points for headways; no points for mentioning independent process, or "independent arrivals", since this was given in the intro).

271 

272 The delay is the area between the two lines (1).

273 Shockwave theory adds a spatial extent, but this is not relevant(1), hence the delay is the same(1). (In case of a non-triangular fundamental diagram, delays would increase since the speeds would remain low – this goes beyond the required answer.)

274 The faster cyclists are observed more frequently (0.5), and hence have a higher weight (0.5). Hence, the time mean speed is higher (1)

275 We would need to compute total travelled distance (TTD) in an area of space-time and total time spent (TTS); both should be divided by the area in space time, and then the quotient should be taken (1). We consider the full track (asked), and a time T (one could for simplicity take 1 hour). Assume n cyclists (or for simplicity 100 cyclists). (0.5) Since the quotient should be taken, the area in space-time drops out and is is not needed to compute the area and: v=TTD/TTS. (0.5)

$$TTS = n * T \tag{B-22}$$

(there are n cyclists each present during the full time T, 0.5 point)

$$TTD = 0.8 * n * 15 * T + 0.2 * n * 25 * T \tag{B-23}$$

(There are 0.8*n cyclist on a normal bike each travelling 15 T in time T, and 0.2*n cyclists each travelling 25 T in time T, 1 point). Now, we compute: $v = TTD/TTS = 0.8{*}15{+}0.2{*}25 = 17$ km/h (0.5) This is indeed the same as the space mean speed, and could be obtained directly. That was not the question, and hence would receive no points.

276 The most downstream vehicle has the lowest number

277 The vehicles *react* on a leader, hence the information comes later to vehicles further upstream in the platoon (1). Hence, red=1, blue=2, and yellow=3. (1 for correct numbers, 0.5 subtracted for a mistake)

278 Changes in speed dampen out quite nicely, so there is local stability

279 The speed reductions of the first vehicle are exaggerated by the following vehicles, e.g. at the first speed decrease, the first vehicle reduces speed to around 10.8 m/s and the last vehicle to 10.2 m/s. (1) If this is the case, this is called platoon unstable (1).

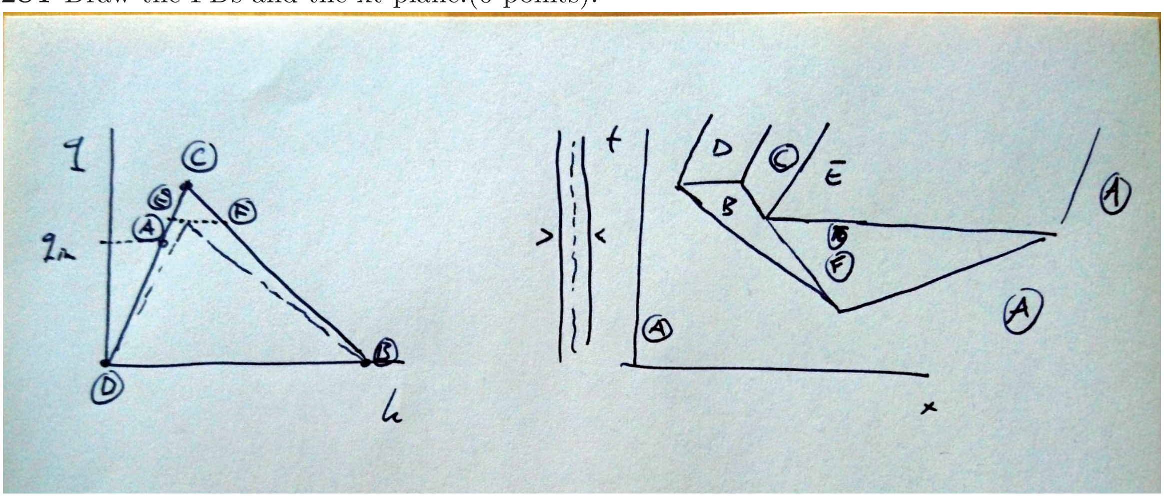
280 A desired distance is s^*, computed increasing (linearly) with speed v. (1) The acceleration is a linear combination of the speed difference of the leader and the difference of spacing and desired spacing, (1) each with its own sensitivity parameter

281 τ can be compared to a reaction time, of around 1 second (0-2 seconds: 1 point; one might also argue about anticipation which reduces reaction time).

282 The FD is based found in equilibrium, hence Δv equals 0, and acceleration equals 0, so γ and τ is irrelevant (1). The FD is hence determined by $s = s_0 + Tv$ (0.5); so with parameters s_0 and T. (0.5) (s and v are the variables, and no parameters)

283 The jam density stays the same, hence s_0 is the same (1). The congested branch is lower, hence T changes.(1)

284 Draw the FDs and the xt plane.(0 points).

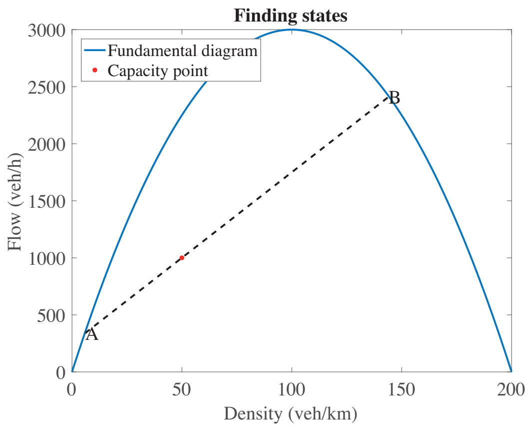


Consider an inflow lower than the capacity of road (no congestion, given), leading to state A on the free flow branch (1). A road block will lead to empty state D downstream (0.5) and jam density upstream, state B (0.5). Waves AD (0.5) and AB (0.5) follow the speeds of the connections in the FD. Once the incident is removed, there is capacity outflow (0.5), with wave DC (0.5). When DC reaches the bridge, no more outflow is possible then the capacity of the bridge, leading to state E downstream (1) and F upstream (1). Waves CE (0.5), EF (0.5) and AF (0.5) can now be constructed, followed by EA (0.5)

285 It suffices: Greenshields assumes a linear relation between speed and flow, so with two given points it is fully determined (1). The capacity is found halfway the density range (1 for knowing or deriving). The speed is half the free flow speed, so the capacity is 30*100=3000 veh/h. (60*1000, ignoring the speed reduction would not give the third point.)

286 Any convex FD will do... We can draw exactly, see next question. Not curved FD: -0.5 point.

287 The fundamental diagram is as follows:

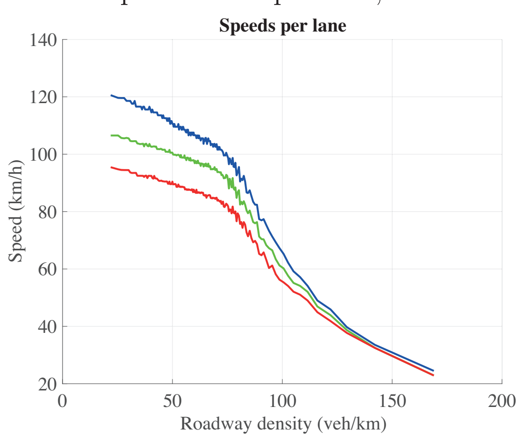


(1 point)

One can draw a line through the capacity point with a speed equal to the speed of the moving bottleneck (1 point). This connects the (free flow) state downstream of the moving bottleneck (state A) and the (congested) state B upstream of the moving bottleneck.

288 The blue line starts from 0 at the low densities, which is the left lane (keep right rule prevents people from using the left lane if not needed) (1). The red line is the right lane, since it starts (like the green one) high, but reduces more quickly and to lower flows than the green one (1).

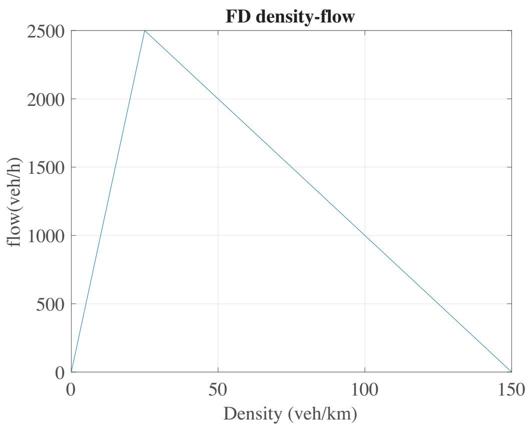
289 The speeds differ per lane, see the fundamental diagrams:



Speeds in the left lane are highest(1), so the graph with the highest fraction in the left lane is the flow (1), so the left figure is the flow (1). Similar reasoning (but the other way around) for right lane.

290 Given is that the fundamental diagram is triangular. The free flow branch is a straight line with a slope of the free speed, 100 km/h, (0.5) up to the maximum flow (2500 veh/h), yielding the critical density of 25 veh/km. (0.5). The congested branch is a straight line down

(0.5), hence one point on the congested branch suffices (0.5). We read a flow of 2000 veh/h and a speed of 40 km/h, yielding a density of 2000/40=50 veh/km (0.5). The fundamental diagram is finilized by drawing a line through the capacity point and the found point (0.5)



291 We take the minimum of demand of cell 2 and supply of cell 3. Cell 2 is undercritical, and has a demand read from the fundamental diagram at k=20 veh/km (0.5), being 2000 veh/h. Cell 3 is overcritical and has a supply read from the fundamental diagram at k=100 veh/km (0.5), being 1000 veh/h. The flow is determined by the minimum, here the supply at 1000 veh/h. (1)

292 We take the minimum of demand of cell 3 and supply of cell 4. Cell 3 is overcritical, hence demand equals capacity, 2500 veh/h (1). Cell 4 is undercritical, hence supply equals capacity, 2500 veh/h (1). The minimum of both is the flow (0.5), hence 2500 veh/h (0.5).

293 The signals coded according to the Dutch standard coding are given in figure B-1

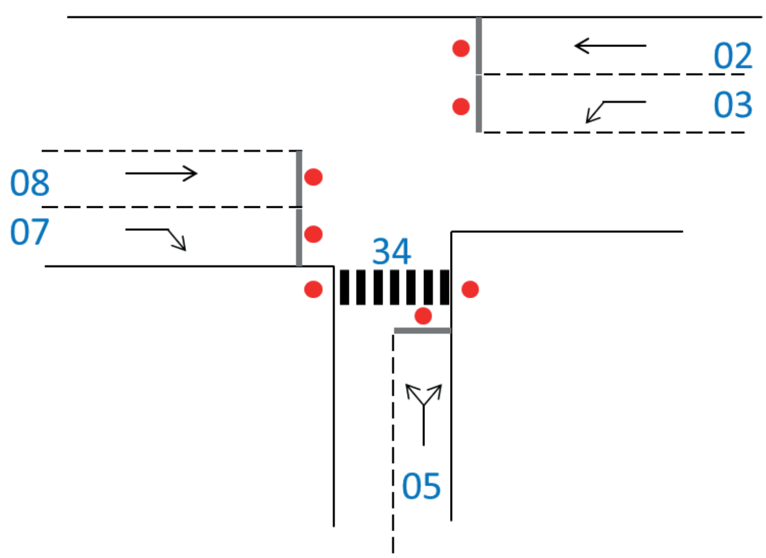


Figure B-1: Intersection layout with standard coding

294 In order to answer this question, the conflict groups have to be determined. The conflicts are: 02-05, 03-05, 03-07, 03-08, 03-34, 05-08, 05-34, 07-34. Since there is mutual conflict 03-05, 03-08, 05-08, this results in conflict group 03-05-08. In the same way 03-05-34 and 03-07-34 can be derived. The conflict group 02-05 cannot be extended. So the maximum conflict group has a size of 3. The order 03-05-08-03 and 03-08-05-03 are different, so there are at least two conflict groups.

295 Given the conflict groups, two control structures are possible:

Control structure 1:

conflict group	1	2	3	4
stage 1	03	03	03	02
stage 2	05	05	07	05
stage 3	08	34	34	02
stage 1	03	03	03	02

Since 02 has no conflict with either 03, 08 or 34, so 02 can be in both stage 1 and stage 3.

The second structure, stage 2 and 3 are swapped with respect to structure 1:

Control structure 2:

conflict group	1	2	3	4
stage 1	03	03	03	02
stage 2	08	34	34	02
stage 3	05	05	07	05
stage 1	03	03	03	02

296 The conflict group cycle time has to be determined for all conflict groups. The maximum conflict group cycle time, so the cycle time of the critical conflict group, is the cycle time of the structure.

Table B-1: Conflict group cycle time Structure 1

Conflict group i	03-05-08-03	03-05-34-03	03-07-34-03	02-05-02	
Internal lost time (s)	9.0	10.0	12.0	9.0	
TGmin (s)	0.0	8.0	8.0	0.0	
$\sum \frac{q_i}{s_i}$	0.722	0.5	0.367	0.778	
$TC_{webster,i}$ (s)	66.6	56.0	48.9	83.2	$TC_{webster}$= 83.2 s
$TC_{min,i}$ (s)	32.4	36.0	31.6	40.5	TC_{min} = 40.5 s

Table B-2: Conflict group cycle time Structure 2

Conflict group i	03-08-05-03	03-34-05-03	03-34-07-03	02-05-02	
Internal lost time (s)	14.0	19.0	16.0	9.0	
TGmin (s)	0.0	8.0	8.0	0.0	
$\sum \frac{q_i}{s_i}$	0.722	0.5	0.367	0.778	
$TC_{webster,i}$ (s)	93.6	83.0	58.4	83.2	$TC_{webster}$= 93.6 s
$TC_{min,i}$ (s)	50.4	54.0	37.9	40.5	TC_{min} = 54.0 s

Mind that for this structure, the Webster cycle time and the minimum cycle time lead to a different critical conflict group.

297 Rabbits have an higher free speed than slugs (1) and use the fastest lane (1).

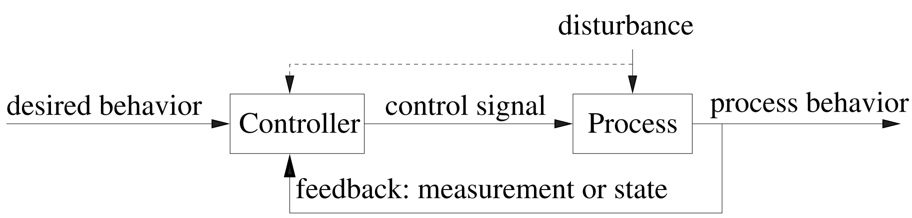
298 The flow contour plots show no decrease in flow anywhere, without congestion

299 situation 1: bottleneck (lane drop, no ramps since no flow change): fixed head, solves from the tail (required for 2 points). Caused by an increase in flow (from flow contour plot). Some stop-and-go waves are visible from the tail (not required).

300 Situation 2: accident/temporal blocking of 1 lane. Low (but not zero, so no full blocking, -1 if claimed otherwise) flow downstream, increase of flow to capacity after recovery. Head moves backwards during solving.

301 The delay is increasing gradually and decreasing gradually, already from an early time (7,4 h), so for situation 1

302 First move the cumulative curves by the free flow travel time (derived from separation at the end of the beginning (1). Then compute the area between the lines (1). Right computation/answer (13,400 veh mins) (1)



303
(process behaviour = process output. Process output is what can be measured, and only what can be measured can be fed back to the controller, so there is only one arrow leaving the process.)

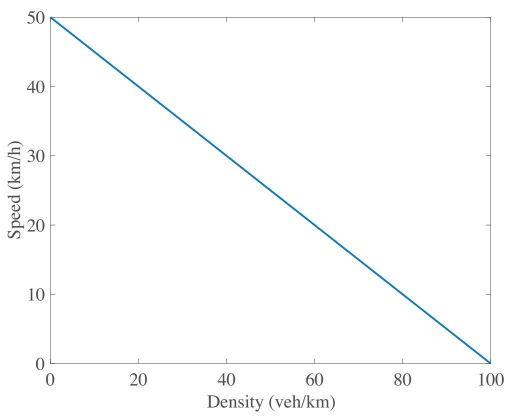
304
- the control signal: the ramp flow rate,
- the controller: the ramp metering system,
- the controlled process: the traffic of the mailine freeway and the on-ramp traffic,
- the measurement: the measured occupancy (or density)

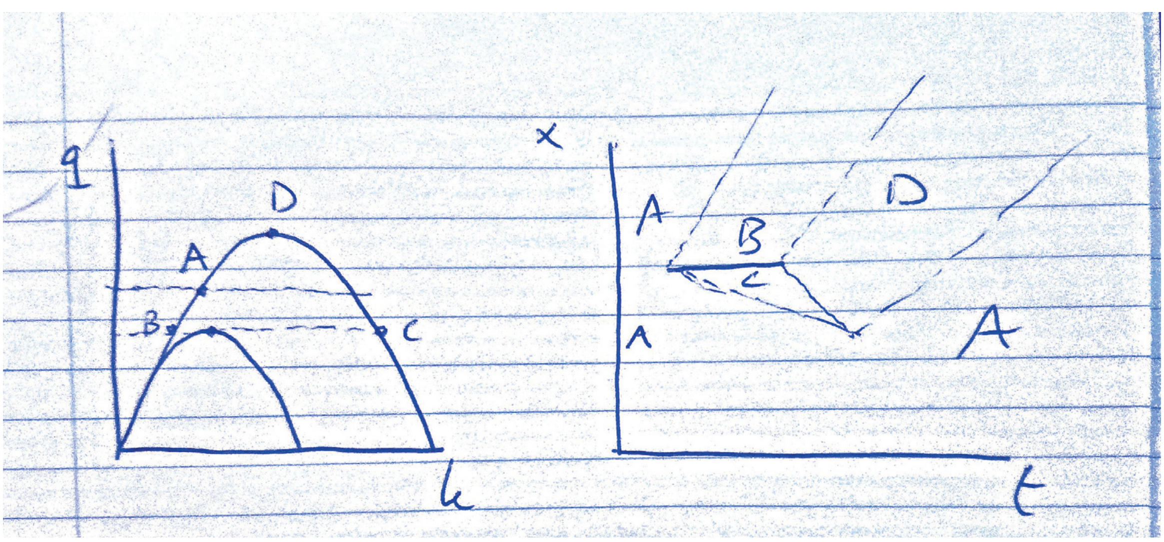
305 Alinea measures the occupancy (density) downstream of the on-ramp on the freeway. This measurement determines the ramp flow in the next time step, which on its turn influnces the traffic state on the freeway at the measurement location. So, the information travels in a loop, and the control structure is feedback.

306 The average flow of the multiple-cars-per-green system is 20s*1800(veh/h)/3600(s/h)/40s = 0.25 veh/s. This means that one car crosses the stop line every 4s on the average. Thus, for the one car per green the cycle time should be 4s, so 0.5s + 1s + t_{red} = 4s, and thus t_{red} = 2.5s.

307 The one-car-per-green spreads the vehicles more evenly when they enter the freeway. So, they cause smaller disturbances, and the chance of a breakdown is smaller. The multiple-car-per-green type sends platoons to the freeway, and the simulataneous merging of several vehicles may cause a breakdown.

308 The Greenshields fundamental diagram has a linear relationship in speed density (1). Reading out some points (at least two), and applying q=ku (0.5) gives the following:





309
Draw both FDs in one plane (0.5). Indicate the inflow point A on the FD and the xt plane (0.5). Blocking (indicated in xt as a horizontal line = state separation, 0.5) limits the outflow to the capacity of the 1 lane road, yielding state B downstream (0.5) and C upstream (0.5). This gives waves AB and AC. (1)Once the blocking has been removed, traffic operates at the capacity of the 2 lane road (0.5). This yields BD, DC, and DA (not parallel, 1)

310 maximizing the total travelled distance: this is not suitable for routing, because the controller will tend to send vehicles on longer routes, even if it increase the travel time. Note that for RM and VSL control only (in networks without route choice), maximizing the TTD may be a reasonable option. If during the time period the TTD is higher, that means that more vehicles have completed their route, or have completed a larger part of their route, so they are closer to exit. This is usually good, but not always. Vehicles may also be sent faster into some jam.

311 minimizing the total travelled distance: this is not suitable due to the variable speed limits and ramp metering: the controller will tend to hold back traffic in order to reduce the driven distance (by RM and VSL). For RG the controller will send the drivers to the shorterst route, even if it is congested, which is not desirable.

312 maximizing the total time spent: not suitable, longer travel times are never desirable.

313 minimizing the total time spent: this is suitable, because it expresses that vehicle should leave the network as fast as possible, and all control measures will serve this purpose. Vehicles will only be sent to longer routes/will only be delayed by VSL or ramp metering, if it can reduce the TTS, so if vehicles on the average arrive earlier at their destination.]

314 The breakdown/jam occurs in the traffic that has left area 4 (to area 5), which indicates that state 5 is unstable (too close to the top of the fundamental diagram). To lower the outflow from area 4, the slope of line DF (v_{45}) has to be made less negative. If in the fundamental diagram state 4 is fixed (in the density-flow plane), then the resulting state 5 (on the free-flow branch of the fundamental diagram) will be lower.

315 The equation of the trajectory of the vehicle in area 4 (stabilization area) is: $x = (t - t_0)v_4 + x_0$. The vehicle will cross line DF when it leaves area 4, which has a slope $v_{45} = (x_D - x_F)/(t_D - t_F) =$ -50 km/h. The equation of the line that the vehicle will cross (line DF) is : $x = (t - t_D) * v_{45} + x_D$. Solving the two equations for t and x gives $t = 0.32$h, $x = 18.43$ km.

316 Calibration

317 We check the speed of the first vehicle, which stops accelerating at when the desired is reached (1). The speed can be read from the slope of the line after the standstill (0.5), and is approximately 20 m/s.

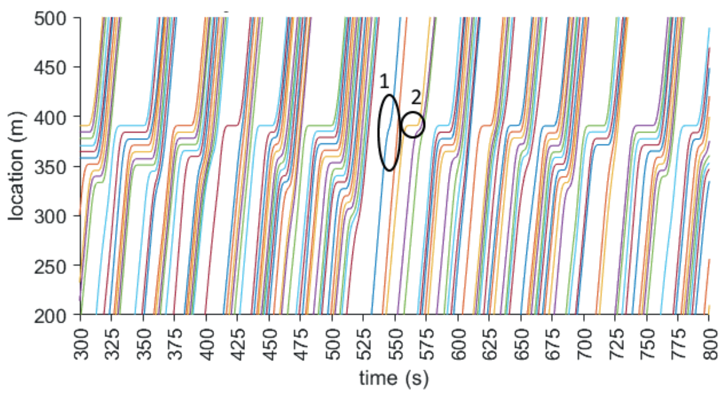
318 Consider the distance at standstill. There are 19 inter-vehicles s_0, which together take (readout) (1620-1490) m. Therefore s_0= 1620-1490)/19=6,8 m

319 The first vehicle brakes and the following vehicles do not reach lower speeds than the previous vehicles. (1) There are hence no indications that the model is platoon unstable (1).

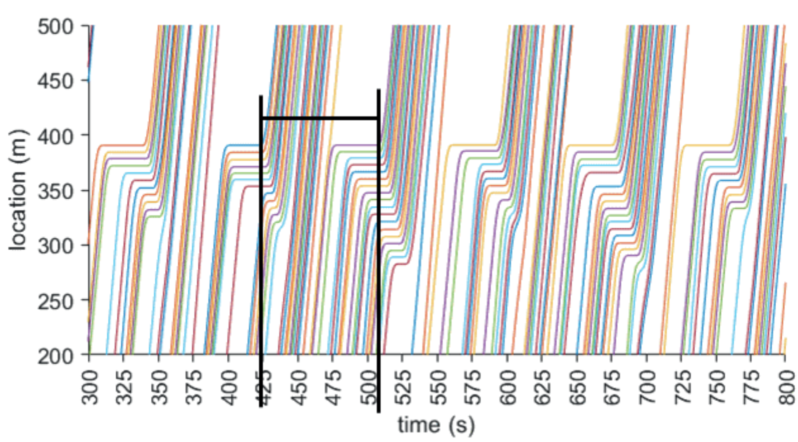
320 The speed is approximately 20 m/s, so the pace is 1/20 s/m.

321 1. Because the time difference between the trajectories of the first vehicle that passes the stop line at start green is the same for all cycles, figure A-3b gives the fixed time control, so Figure A-3a is the vehicle-actuated control.

2. In Figure A-3a, a trajectory near 550 s, shows a vehicle that arrives in red, because it has to break a bit, but apparently there is no conflicting traffic, so it can continue right away and accelerates again. So the green start is acting upon vehicle arrivals, as vehicle-actuated control does. Also the trajectory directly before 500 s has only to stop a short while.



322 The cycle time can be determined by the start green of a cycle to the start green of the next cycle, so about 507-425 s=82 s (everything between 80-88 s is ok) For full points a complete explanation has to be given, with the time stamps that mark the start of green. Better not use the end of green, since the last arrival not necessarily coincides with the end of green, the green time being fixed. If the next trajectory that shows a stop is close to the last-through-green, you can argue the end of green does coincide with the last arrival that passes green (no points are deducted if the end of green instead of the start of green is used). Since the cycle time is fixed, determining one cycle is sufficient to get full points, determining more and take the average is indeed more precise.



323 The Webster cycle time is:

$$T_{Cwebster} = \frac{1.5T_{LI+}5 + \sum_{i==i_{min}} T_{g,min}}{1 - \sum_{i \neq i_{min}} (q_i/s_i)} \tag{B-24}$$

Intersection 1
The (critical) conflict group is 02-05-02, so internal lost time T_{LI} is 6 seconds, sum of the load q_i/s_i is (600+900)/1800=0.833,
this makes the Webster cycle time: (5+1.5*6)/(1-0.833)=84 s.
The equation for the (effective) green time is:

$$T_{g,eff,i} = (y_i/Y)(T_C - T_{LI}), \tag{B-25}$$

so for direction 02: (600/1800)(84-6)/(1500/1800)=31.2 s,
for direction 05: (900/1800)(84-6)/(1500/1800)=46.8 s,
both larger than the minimum green time.

Intersection 2
The (critical) conflict group is 01-05-09-01, or 01-09-05-01, the lost time is: 9 s, the sum of the load is ((600+750+54)/1800)= 0.78,
this makes the Webster cycle time:
$T_{CWebster}$=(5+1.5(9))/(1-((600+750+54)/1800))=84.1 s.
The green time for 01 is: (600/1800)(84-9)/(1404/1800)=32.1 s,
for 05: (750/1800)(84-9)/(1404/1800)=40.1 s,
for 09: (54/1800)(84-9)/(1404/1800)=2.9 s, this is not sufficient green time for 09, since the minimum green time is 6 s.
The green time of 09 is set to 6 s, the cycle time has to be recalculated.
$T_{CWebster}$=(6+5+1.5(9))/(1-((600+750)/1800))=98 s,
direction 09 has a green time of 6.0 s,
01 has a green time of (600/1800)(84-9-6)/(1350/1800)=36.8 s,
05 has a green time of (750/1800)(84-9-6)/(1350/1800)=46.1 s.

324 The trajectories lead to a cycle time around 84 s (give or take a few seconds by determining the cycle time from the graph), the calculated cycle time of intersection 1 is 84 s, for intersection 2 is 98 s. The Webster cycle time is determined based on the idea that given Poisson distributed arrivals, the cycle time will be optimum. Since 84 is much lower than 98 s, a cycle time of 84 s will lead to queues for intersection2. So intersection 1 is used.

Appendix C

Matlab code for creating (slanted) cumulative curves

```
function cumcurves()
%This function will give the cumulative curves for a flow which
%has (in time) three values, and one bottleneck halfway
q0=[3600;5000;2000];%the three demands
Tchange=[60;90];%times in minutes at which the demands change
c=4000;%capacity
T=0:200;%minutes
dt=1/60;%time steps (in hours: time step is 1 min)
dem=q0(end)*ones(size(T));%pre-allocate demand function to the last demand ...
    value
for(i=numel(Tchange):-1:1)
    dem(T<Tchange(i))=q0(i);%adapt the demand function
end

figure;
plot(dem,'linewidth',2)
hold on
plot(repmat(c,size(dem)),'r--','linewidth',3)
ylim([0 6000])
legend('Demand','Capacity','location','Northeast')
ylabel('Flow (veh/h)')
xlabel('Time (min)')
exportfig('Demand')
%%

Nin=dt*cumsum(dem);
qout=zeros(size(Nin));
qout(1)=dem(1);%in veh/h
queued=zeros(size(dem));
for(t=2:numel(T))
    qout(t)=1/dt*min(dt*c, dt*dem(t)+queued(t-1));
    queued(t)=queued(t-1)+dt*dem(t)-dt*qout(t);
end
```

```
Nout=dt*cumsum(qout);
figure;plot(Nin,'linewidth',2,'color',[0.5 0.5 1]);hold ...
    on;plot(Nout,'r--','linewidth',3)
legend('N_{in}','N_{out}','location','Northwest')
ylabel('Cumulative flow')
xlabel('Time (min)')
exportfig('Cumulative curves')
%%

%compute total delay:
TotalD=dt*sum(queued)%then the total delay in hours
NrVeh=Nin(end);
AvgDelay=TotalD/NrVeh%then the total delay in hours
AvgDelayMin=60*AvgDelay;%then the total delay in hours

%%
Tin=interp1(Nin,T,1:NrVeh);%time to enter for each vehicle -- interpolation
Tout=interp1(Nout,T,1:NrVeh);%time to exit for each vehicle -- interpolation
DT=Tout-Tin;%additional travel time
figure;
plot(1:NrVeh,DT,'linewidth',2)
xlabel('Vehicle number')
ylabel('Delay (min)')
exportfig('Delay per vehicle')

figure;
plot(Tin,DT,'linewidth',2)
xlabel('Entry time (min)')
ylabel('Delay (min)')
exportfig('Delay as function of time')
```