An improved first-order macroscopic flow model for highway traffic simulation

T. Bellemans
ESAT/SISTA, K.U.Leuven
Kasteelpark Arenberg 10, B-3001 Leuven (Heverlee), Belgium
tom.bellemans@esat.kuleuven.ac.be

B. De Schutter
Control Lab, Fac. of Information Technology and Systems
Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands
b.deschutter@its.tudelft.nl

B. De Moor
ESAT/SISTA, K.U.Leuven
Kasteelpark Arenberg 10, B-3001 Leuven (Heverlee), Belgium
bart.demoor@esat.kuleuven.ac.be

Abstract

We reconsider the first-order macroscopic traffic flow model of Lighthill and Whitham and we propose some modifications that solve some of the deficiencies related to this model, while at the same time preserving one of the main advantages of this model compared to higher-order models, namely its low computational complexity, which is important when the model is used in on-line traffic control applications.

1 Introduction

In this paper we reconsider the first-order macroscopic traffic flow model of Lighthill and Whitham [4]. Although this model has since long been complemented by second-order models [3, 7, 8] in order to avoid several deficiencies and to improve the modeling accuracy [6], it is still used in practice especially in on-line control applications [2] for large networks where the computational complexity of second-order models is too high.

In this paper we will briefly revisit the model of Lighthill and Whitham, and then illustrate some of the problems that can occur when this model is used for simulation. Next we present some modifications for the model which take care of these problems without increasing the order of the model so that the computational requirements of the improved model stay low.

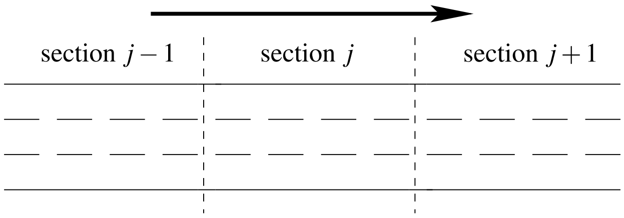

Figure 1: Schematic figure of a highway divided in several consecutive sections.

2 The traffic flow model of Lighthill and Whitham

Traffic models are not only used to estimate the impact of different traffic control measures such as ramp metering, velocity harmonization, etc. on the traffic situation, but they are also used to design or optimize traffic control strategies. There exist various models to describe traffic on highways going from microscopic models (in which individual cars are simulated) over mesoscopic models to macroscopic models (which work on a much more aggregated level and characterize the traffic streams by average speeds, densities and flows).

We will consider a macroscopic flow model that has been introduced by Lighthill and Whitham [4]. For the sake of simplicity of the exposition we will not consider on-ramps and off-ramps. Note however that the model can easily be adapted to include on-ramps and off-ramps.

In a macroscopic traffic flow model the highway is divided in several sections, which typically have a length of about 500 m (see Figure 1). Besides the discretization in space, we also have a discretization in time. The time step or sampling is denoted by Δt. A typical value for Δt is 15 s.

The traffic situation in section j at sample step k, i.e., at time $k\Delta t$, is described by the traffic density or concentration $c_j(k)$ (expressed in vehicles/km/lane), the average speed $v_j(k)$ (expressed in km/h), and the traffic flow or volume $q_j(k)$ (expressed in vehicles/h). These state variables are updated at each sampling time instant. Next we will describe the difference equations that give the relation between the state variables at the current and the next sampling time instant. Let $q_{\mathrm{in},j}(k)$ and $q_{\mathrm{out},j}(k)$ be respectively the in-flow and the out-flow for section j in the period $[k\Delta t,(k+1)\Delta t)$. The flow conservation principle then results in

$$c_j(k+1) = c_j(k) + \frac{\Delta t}{l_j n_j}(q_{\mathrm{in},j}(k) - q_{\mathrm{out},j}(k)) \tag{1}$$

where l_j is the length of the section and n_j the number of lanes. Equation (1) states that the net in-flow in a section over a time interval results in an increase (or a decrease if the net in-flow is negative) of the density of that section. The factor $\frac{1}{l_j n_j}$ is the geometry dependent storage capacity of section j.

The average speed in a section is known to be a function of the density [5]. Although this function is highly nonlinear, it is common practice to approximate it by a two-segment linear function in order to reduce the computational complexity of the model [1, 9]. However, the behavior of the speed-density relation at the transfer between both segments is highly nonlinear and cannot be captured by the piece-wise linear approximation [2]. Therefore, we use the following empirical formula [5] which takes the nonlinearity fully into account:

$$v_j(k) = F\left(c_j(k)\right) \stackrel{\mathrm{def}}{=} v_{\mathrm{f}}\left[1-\left(\frac{c_j(k)}{c_{\mathrm{jam}}}\right)^{\alpha}\right]^{\beta} \tag{2}$$

where v_{f} is the free flow speed and c_{jam} is the jam density. The free flow speed v_{f} is the average speed of a vehicle that is not influenced by other vehicles. This average speed is reached by a vehicle as the density in the section goes to zero. The jam density c_{jam} is the traffic density at which the traffic flow stalls and the vehicles are standing still in the section. The relation between $v_j(k)$ and $c_j(k)$ is represented graphically in Figure 2 for typical values of the parameters v_{f}, c_{jam}, α and β. Note that in general v_{f} and c_{jam} may be different for each section since they depend on the curvature of the section, the width of the lanes, etc. The parameters v_{f} and c_{jam} are generally fitted on traffic data for the given section using regression techniques while the parameters α and β are usually taken from literature [5] since they have a smaller impact on the quality of the fit.

Given the traffic density and the average speed in a section, we can calculate the flow through that section:

$$q_j(k) = c_j(k) v_j(k) n_j \ . \tag{3}$$

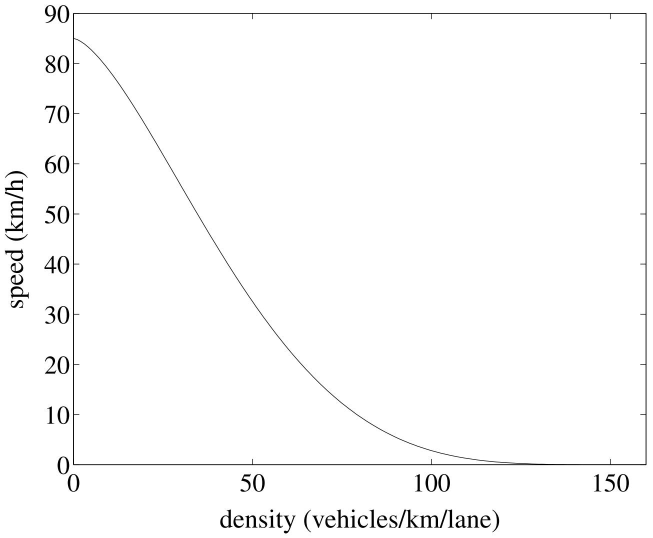

Figure 2: The nonlinear speed density relation in a highway section according to the empirical formula (2). For this plot we have used the values $v_{\mathrm{f}} = 85$ km/h, $c_{\mathrm{jam}} = 160$ vehicles/km/lane, $\alpha = 1.5$, $\beta = 5$ (cf. [5, p. 309]).

When a complete highway or highway network is simulated, consecutive sections have to be coupled together. This can be done in two ways:

- the in-flow of section j can be set equal to the flow in the jth section:

$$q_{\mathrm{in},j}(k) = q_j(k) \ . \tag{4}$$

  This will be called *upstream coupling* as will be explained later on.

- the out-flow of section j can be set equal to the flow of the jth section:

$$q_{\mathrm{out},j}(k) = q_j(k) \ . \tag{5}$$

  This will be called *downstream coupling*.

Furthermore, we have $q_{\mathrm{in},j}(k) = q_{\mathrm{out},j-1}(k)$.
We shall see later on that expressions (4) and (5) will both result in a first-order model for the highway or highway network. Note that it is possible to define more complex coupling relations. However, this leads to more complex or to higher-order models.

If the coupling relations have been defined, then the average speed and the flow in each section can be derived from the density in the section via (2) and (3). The difference between the coupling equations (4) and (5) becomes clearer when we rewrite the nonlinear state space model for both choices by eliminating the average speed and the flow from the evolution equations of the traffic system.

If we combine (1), (2) and (4) we obtain the following nonlinear difference equation for the traffic density in section

j:

$$c_j(k+1) = c_j(k) + \frac{\Delta t}{l_j n_j}\Big[c_j(k)F\big(c_j(k)\big)n_j - c_{j+1}(k)F\big(c_{j+1}(k)\big)n_{j+1}\Big] \ . \quad (6)$$

So $c_{j+1}(k)$ and $c_j(k)$ determine $c_j(k+1)$. Therefore, we call (4) *upstream coupling*. This term becomes clearer when we look at Figure 3 where we see successive highway sections at two consecutive time instances. In this figure we have added thick arrows that represent the data flows due to the coupling. Clearly, the arrows point in the upstream direction of the highway. Note that (6) is a first-order model in which we only need one variable to completely characterize the traffic situation in the given section.

When combining (1), (2) and (5) we obtain the following nonlinear difference equation for the traffic density in section j:

$$c_j(k+1) = c_j(k) + \frac{\Delta t}{l_j n_j}\Big[c_{j-1}(k)F\big(c_{j-1}(k)\big)n_{j-1} - c_j(k)F\big(c_j(k)\big)n_j\Big] \ . \quad (7)$$

So $c_{j-1}(k)$ and $c_j(k)$ determine $c_j(k+1)$. Therefore, we call (5) *downstream coupling*. In Figure 4 we see a graphical representation of the data flows in time and space for the downstream case.

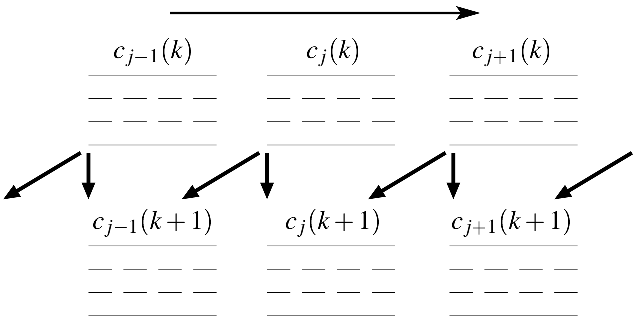

Figure 3: Schematic presentation of upstream coupling. The figure shows the same highway at two consecutive sampling steps k and $k+1$. The thick arrows between the sections indicate the dependencies of the states of the sections through time as described by (6).

3 Model behavior and deficiencies

In this section we discuss some specific cases to illustrate some properties and deficiencies of the models (6) and (7). In Section 4 we will then present some improvements of the model that will address the problems encountered in this section.

3.1 Upstream coupling

Consider the situation depicted in Figure 3 where information about the state of a section is provided to the upstream section. Now suppose that section j is empty at sample step k. So $c_j(k) = 0$. Let us now look what happens in sections j and $j+1$ at sample step $k+1$.

- Section j: Since $c_j(k) = 0$, we have $q_j(k) = 0$ by (3) and thus also $q_{\text{in},j}(k) = 0$ according to (4). This means that section j stays empty independently of what happens upstream of j. This is an unrealistic situation which will not occur in practice unless a road is completely closed.
- Section $j+1$: According to (4) we have $q_{\text{in},j+1}(k) = q_{j+1}(k)$. Hence, $q_{\text{out},j}(k) = q_{\text{in},j+1}(k) = c_{j+1}(k)v_{j+1}(k)n_{j+1}$, which is not necessarily zero. So in this case vehicles are assumed to leave section j while j is empty. This would imply that cars are "created" out of nothing and that the flow conservation property is violated as well as the physical constraint that the density in j must be nonnegative.

It is clear from the two presented cases that for upstream coupling not only the state of the considered section has to be taken into account but also the state of the preceding section. Sometimes this is implemented by using a more complex coupling equation such as

$$q_{\text{in},j}(k) = \alpha q_{j-1}(k) + (1-\alpha)q_j(k) \ . \quad (8)$$

The estimation of α and the lack of a physical motivation are two drawbacks of this method. Moreover, in the next subsection we shall show that downstream coupling does not exhibit the problems connected to empty sections that have been discussed above.

3.2 Downstream coupling

Now consider the situation of Figure 4 where the information of the state of a section is provided to the downstream section. We illustrate the properties of the downstream coupled model with some cases.

When we assume that section j is empty at sample step k, we see that the following phenomena occur:

- If $q_{j-1}(k) \neq 0$ then cars can enter the empty section j from the upstream section $j-1$ since $q_{\text{in},j}(k) = q_{\text{out},j-1}(k) = q_{j-1}(k)$ by (5). So in that case section j starts filling up at sample step $k+1$.
- If section j is empty at sample step k, we have $c_j(k) = 0$ and thus also $q_j(k) = 0$ by (3). Hence, $q_{\text{out},j}(k) = q_j(k) = 0$ by (5).

So for the case where $c_j(k) = 0$ the problems of blocking and car generation out of the void that occurred with upstream coupling have disappeared by using downstream coupling. This does not mean that all problems have disappeared though. The following two cases illustrate that also with downstream coupling problems can occur. Assume that section j is congested at sample step k, which means that $c_j(k) = c_{\text{jam}}$. Note that this implies that $v_j(k) = 0$ and $q_j(k) = 0$ as can be seen from (2) and (3). Let us now look what happens in sections j and $j+1$ at sample step $k+1$.

- Section j: By (5) the in-flow in section j is given by $q_{\text{in},j}(k) = q_{\text{out},j-1}(k) = q_{j-1}(k) = c_{j-1}(k)v_{j-1}(k)n_{j-1}$. So the in-flow in section j depends on the out-flow of section $j-1$ but not on the state of section j. This means that in the model with downstream coupling as it was presented before, the flow of vehicles leaving a section is always fed into the next section. This poses problems when the state of the section receiving the flow is such that the section is not able to accept the cars. E.g., in our example here, cars enter into section j even though the density $c_j(k)$ is already maximal and $q_{\text{out},j}(k) = q_j(k) = 0$. This implies a violation of the flow conservation property since the entering cars cannot be accounted for by an increased density. It seems as if they disappear. The state of the downstream section j should not only be taken into account when calculating the flow out of j, but it should also be used to check whether the inflow does not exceed the maximum allowable value. A discussion of how this can be done is provided in the Section 4.

- Section $j+1$: According to (5) the in-flow in section $j+1$ is given by $q_{\text{in},j+1}(k) = q_{\text{out},j}(k) = q_j(k) = 0$. This has two consequences. First, the density in section $j+1$ can only decrease since the in-flow is always zero. Second, once a section, say j, becomes congested it stays congested throughout the whole simulation since the out-flow stays zero indefinitely.

It is clear that the behavior of the model in the two cases above does not mimic the behavior of real traffic flows in an analogous situation. Therefore, we will have to modify the model such that the behavior of the model corresponds better to real-life traffic flows. However, on the other hand we want to keep the complexity of the model low since we want to use it in on-line traffic control applications. The resulting modifications are discussed in the next section.

4 An improved first-order traffic flow model

There exist different solutions to overcome the model deficiencies presented in the previous section.

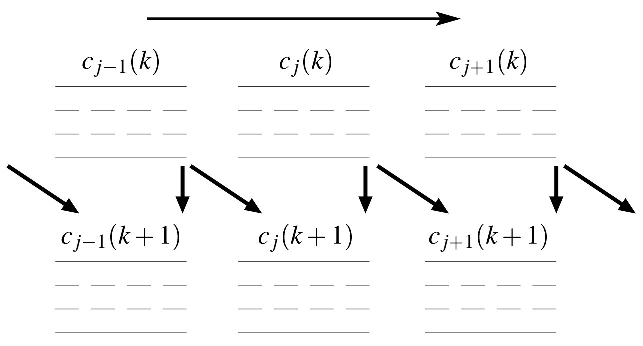

Figure 4: Schematic representation of downstream coupling. The figure shows the same highway at two consecutive sample steps k and $k+1$. The thick arrows between the sections indicate the dependencies of the states of the sections through time as described by (5).

A first solution consists in choosing a more advanced model. As an example we mention the second-order models where the speed in the section is taken as an extra state in the model and a difference equation replaces (2). These models are able to deal with the problems concerning the flow conservation law as well as with the problem of the out-flow of a congested section. Information about the traffic density and the average velocity in the neighboring section is processed in the model. However, these more advanced models require more parameters to be tuned when calibrating the model. Moreover, they have a higher computational complexity which is disadvantageous when we want to use the model in on-line applications where computation speed is important. Since in our own research we are mainly interested in traffic models that can be used in on-line traffic control, we will trade off higher accuracy for lower computational complexity, and focus on simple modifications to the first-order model presented in Section 2.

In the case of upstream coupling, we could take the state of the upstream section into account in order to prevent cars from being created and to allow an empty section to fill. When applying coupling equation (8), we introduce a new parameter which has to be tuned and which has no physical motivation. Moreover, in order to prevent cars from being created at an empty section we must have $\alpha = 1$ which is equivalent to (5). Therefore, we start from downstream coupling and we address the following two problems that occur with downstream coupling:

- Overflow in a section

- Inability to simulate a disappearing congestion

To prevent overflow in section j we have to prevent the net in-flow from becoming so large that the density in section j becomes larger than the maximal density c_{jam}. This implies that the following condition should always be satisfied:

$$c_j(k) + \frac{\Delta t}{l_j n_j}\left[q_{\text{in},j}(k) - q_{\text{out},j}(k)\right] \leq c_{\text{jam}}$$

or equivalently

$$q_{\text{in},j}(k) \leq q_{\text{out},j}(k) + \frac{l_j n_j}{\Delta t}\left[c_{\text{jam}} - c_j(k)\right] \ . \tag{9}$$

Equation (9) can be interpreted as follows. The maximal in-flow in a section j is equal to the flow leaving the section, $q_{\text{out},j}(k)$, plus a flow of cars which are absorbed by the section and which bring the traffic density to its maximal value, c_{jam}. We can decompose the in-flow of a section in two parts: a flow leaving the section again (throughput) and a flow changing the density in the section (storage capacity of the section).

Equation (9) provides an upper bound for the in-flow in section j. By combining this upper bound (9) with (5) we realize a coupling between sections $j-1$ and j which takes the states of both sections into account. The resulting expression for the in-flow in section j is given by

$$q_{\text{in},j}(k) = \min\Bigg(q_{j-1}(k), \\ q_{\text{out},j}(k) + \frac{l_j n_j}{\Delta t}\left[c_{\text{jam}} - c_j(k)\right]\Bigg) \ . \tag{10}$$

This equation prevents cars from disappearing (overflow of the density in j) while it needs only one state, the density $c_j(k-1)$.

Finally we have to deal with the inconsistent behavior of the model when congestion occurs. Once c_{jam} is reached in a section the out-flow of this section becomes zero according to (2) and stays zero indefinitely. This does not correspond to the behavior that can be observed in practice. In order to prevent the out-flow from becoming zero we modify (2) as follows:

$$v_j(k) = \begin{cases} v_{\text{f}}\left[1 - \left(\frac{c_j(k)}{c_{\text{jam}}}\right)^{\alpha}\right]^{\beta} & \text{if } c_j(k) \leq (1-\varepsilon)c_{\text{jam}} \\ \zeta v_f & \text{if } c_j(k) > (1-\varepsilon)c_{\text{jam}} \end{cases}$$

where ζ and ε are small positive real numbers. Note that we can avoid a discontinuity of v_j for $c_j = (1-\varepsilon)c_{\text{jam}}$ by imposing the following relationship between ε and ζ:

$$\zeta = \left(1 - (1-\varepsilon)^{\alpha}\right)^{\beta} \tag{11}$$

The two corrections presented in this section enable us to simulate the traffic flows in a highway network using a first-order model. Even though this model is still first-order, it does no longer exhibit the problems of the original model that have been discussed in the previous section.

5 Conclusion

In this paper we have considered the first-order macroscopic traffic flow model of Lighthill and Whitham. First we have shown that in some cases this model exhibits some deficiencies. Next we have proposed some modifications that solve these deficiencies while at the same time preserving the low computational complexity of the model, which is an important advantage in comparison with higher-order traffic flow models if we want to use the model in on-line traffic control applications.

One of the main topics for future research is an extensive determination of the trade-off between accuracy and computational complexity of the improved first-order traffic flow model compared to higher-order traffic flow models that can be found in literature.

Acknowledgments
Bart De Moor is a senior Research Associate with the F.W.O. (Fund for Scientific Research-Flanders).
This work is supported by several institutions: the Flemish Government (Research Council K.U.Leuven: Concerted Research Action Mefisto-666; FWO projects: G.0240.99 and G.0256.97; FWO-Research Communities: ICCoS and ANMMM; IWT projects: EUREKA 1562-SINOPSYS, EUREKA 2063-IMPACT, STWW), the Belgian State, Prime Minister's Office-OSTC (IUAP P4-02 (1997-2001) and IUAP P4-24 (1997-2001); Sustainable Mobility Programme-Project MD/01/24 (1997-2000) "Traffic Congestion Problems in Belgium: Mathematical Models, Simulation, Control and Actions"), and the European Commission (TMR Networks: ALAPEDES and System Identification; Brite/Euram Thematic Network: NICONET).

References

[1] G.L. Chang, J. Wu, and S.L. Cohen, "Integrated real-time ramp metering model for nonrecurrent congestion: Framework and preliminary results," *Transportation Research Record*, no. 1446, pp. 56–65, 1994.

[2] O.J. Chen, A.F. Hotz, and M.E. Ben-Akiva, "Development and evaluation of a dynamic ramp metering control model," in *Proceedings of the 8th IFAC/IFIP/IFORS Symposium on Transportation Systems*, Chania, Greece, pp. 1162–1168, June 1997.

[3] M. Cremer and A.D. May, "An extended traffic flow model for inner urban freeways," in *Proceedings of the 5th IFAC/IFIP/IFORS International Conference on Control in Transportation Systems*, Vienna, Austria, pp. 383–388, 1986.

[4] M.J. Lighthill and G.B. Whitham, "On kinematic waves: II. A theory of traffic flow on long crowded roads,"

Proceedings of the Royal Society of London, vol. 299A, pp. 317–345, May 1955.

[5] A.D. May, *Traffic Flow Fundamentals*. Englewood Cliffs, New Jersey: Prentice-Hall, 1990.

[6] M. Papageorgiou, "Some remarks on macroscopic traffic flow modelling," *Transportation Research Part A*, vol. 32, no. 5, pp. 323–329, Sept. 1998.

[7] M. Papageorgiou, J.M. Blosseville, and H. Haj-Salem, "Modelling and real-time control of traffic flow on the southerm part of Boulevard Périphérique in Paris – Part I: Modelling," *Transportation Research Part A*, vol. 24, pp. 345–359, 1990.

[8] H.J. Payne, "Models of freeway traffic and control," in *Mathematical Models of Public Systems* (G.A. Bekey, ed.), vol. 1, no. 1 of *Simulation Council Proceedings Series*, pp. 51–61, La Jolia California, 1971.

[9] H. Zhang, S. Ritchies, and W. Recker, "Some general results on the optimal ramp control problem," *Transportation Research Part C*, vol. 4, no. 2, pp. 51–69, 1996.