Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

J.N.T. Schuit

Numerical solution of the Hamilton-Jacobi-Bellman equation for a freeway traffic flow control problem

Department of Operations Research, Statistics, and System Theory   Note BS-N8901   January

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

Numerical Solution of the Hamilton-Jacobi-Bellman Equation for a Freeway Traffic Flow Control Problem

J.N.T. Schuit

Department of Applied Mathematics of Twente University,
P.O. Box 217, 7500 AE Enschede, The Netherlands
and
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

In this report the optimal control laws for two freeway traffic flow models, described by non-linear stochastic differential equations, are determined by discretization of the state space and application of dynamic programming techniques. The objective is to maximize the flow of freeway traffic. The optimal stochastic control problem leads to the Hamilton-Jacobi-Bellman equation. A discretization method is used to approximate the stochastic differential equations by finite-state continuous-time Markov processes. The Hamilton-Jacobi-Bellman equation of the discretized Markov processes can be solved by policy iteration, successive approximation or modified policy iteration methods. A block-iterative version of modified policy iteration is proposed which seems very appropriate for a two-dimensional model. Finally the resulting optimal control law is discussed.

1980 Mathematics Subject Classification: 93E20.
Key Words & Phrases: optimal stochastic control, Hamilton-Jacobi-Bellman equation, numerical solution, freeway traffic.

Note BS-N8901
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

TABLE OF CONTENTS

1. Introduction

In this report an investigation is described on how to determine an optimal control law for the Dutch Motorway Control and Signalling System. The objective of control is to make the freeway traffic flow more homogeneous and to delay the start of congestion. The system consist of a consecutive series of detection stations and portals. The detection stations consist of measuring loops embedded in the road surface. From each detection station one may obtain the passage times of cars and their passage speeds. On the portals are electronic boards which can display an advisory speed to drivers. A description of this system is given in Appendix A.

Congestion often sets in when about 85% to 90% of the traffic capacity is reached. A reason for this is the inhomogeneity of the traffic flow. The Traffic Engineering Division of Rijkswaterstaat has made experiments with the objective to make the traffic flow more homogeneous. This is done by a homogenising control: displaying an advisory speed on the signal boards of the control and signalling system to drivers. The same speed is shown on all boards, independent of lane or position. An analysis of these experiments leads to the conclusion that this homogenising control is useful, applying this kind of control makes the traffic flow more homogeneous and delays the start of congestion.

At the Centre for Mathematics and Computer Science a research project is carried out by S.A. Smulders in which the possibility of using homogenising control for freeway traffic. The control law should switch the advisory speed signs on or off depending on the information gained from the measurements. In [9] a model for freeway traffic flow is described that consists of a series of consecutive sections. A section is regarded as a piece of freeway with at each end a detection station. Based on this model a filter was developed in [10]. In [11] two models for one section were derived from the former model: a one-dimensional model and a two-dimensional model. The first is a simplification of the second. The effect of the homogenising control on the traffic flow has been determined, and incorporated in the models. Investigations into useful control laws were made.

In [11] a suggestion is made for a method to compute the optimal control law for the one-section model. This method will be used in this report to determine the optimal control for the one-dimensional and the two-dimensional model. Also a comparison of the performance between optimal and suboptimal controls of different classes is made.

The models are described by stochastic differential equations. The one-dimensional model is for the density only, the two-dimensional model is for both density and mean speed of a section. The proposed method is described in [6] and [3]. According to this method one approximates a stochastic differential equation by a finite-state continuous-time Markov process. For this process an optimal control policy can be determined by several methods. The policy iteration method is described in [5], the successive approximation method in [2] and a modified policy iteration method in [7]. In this report another version of the modified policy iteration method is proposed which is very useful for Markov processes coming from approximated stochastic differential equations.

After the determination of the optimal control it will be demonstrated that the class of one switch-controls based on the density only has a suboptimal control law which is both useful and simple.

The problem of determining the optimal control for the one-dimensional and the two-dimensional model will be formulated in section 2.1. In section 2.2. the discretization procedure is formulated. The methods for finding the optimal solution of the finite-state continuous-time Markov process are described in section 2.3. Results of these methods and an investigation into classes of suboptimal controls are presented in section 3 for the one-dimensional model, and in section 4 for the two-dimensional model. Section 5 contains general conclusions and suggestions for research.

2. Problem formulation

In this chapter two models will be presented, both describing the flow of freeway traffic in a section of a freeway. Both models are based on a freeway traffic flow model for a series of consecutive sections presented in [9]. The one-dimensional model consists of a stochastic differential equation for the density only, the two-dimensional model contains also a stochastic differential equation for the mean speed.

By discretization of the state space the stochastic differential equation can be approximated by a finite-state continuous-time Markov process. Both in [3] and in [6] techniques are proposed for doing this. For the Markov process the optimal control can be found using dynamic programming.

The two basic computational methods for dynamic programming which will be considered are policy iteration and successive approximation. Combining those two methods gives a range of methods each of which will be called a *modified policy iteration method.*

2.1. Models for freeway traffic flow

The two models to be considered are based on the model proposed in [9]. The simplification to a model for one section was done in [11]. The models are based on two state variables:

ρ_t: the density (number of veh/km/lane) in the section at time t.

v_t: the mean speed (km/h) of the vehicles in the section at time t.

The one-dimensional model is given by:

$$d\rho_t = \frac{1}{lL}(\lambda_0 - l\rho_t v_t)dt + \sigma dw_t \qquad (1.1)$$

$$v_t = v^e(\rho_t) \qquad (1.2)$$

The two-dimensional model is given by:

$$d\rho_t = \frac{1}{lL}(\lambda_0 - l\rho_t v_t)dt + \sigma dw_t$$

$$dv_t = -\frac{1}{T}(v_t - v^e(\rho_t))dt + \mu dz_t \qquad (1.3)$$

Here $v^e(\rho_t)$ denotes the equilibrium speed:

$$v^e(\rho) = \begin{cases} v_{free} - \alpha\rho & \text{for } 0 \leqslant \rho \leqslant \rho_{crit} \\ d(\frac{1}{\rho} - \frac{1}{\rho_{jam}}) & \text{for } \rho_{crit} < \rho \leqslant \rho_{jam} \end{cases} \qquad (1.4)$$

where

$$d = \frac{v_{free} - \alpha\rho_{crit}}{\frac{1}{\rho_{crit}} - \frac{1}{\rho_{jam}}}$$

to assure continuity at ρ_{crit}. The definitions of the parameters are:

- l: the number of lanes in the section;
- L: the length of the section (km);
- λ_0: the intensity at the entrance (veh/h);
- σ: the standard deviation of the noise in the density equation;
- w_t: a standard Brownian motion;
- T: the relaxation time (h);
- μ: the standard deviation of the noise in the speed equation;
- z_t: a standard Brownian motion;

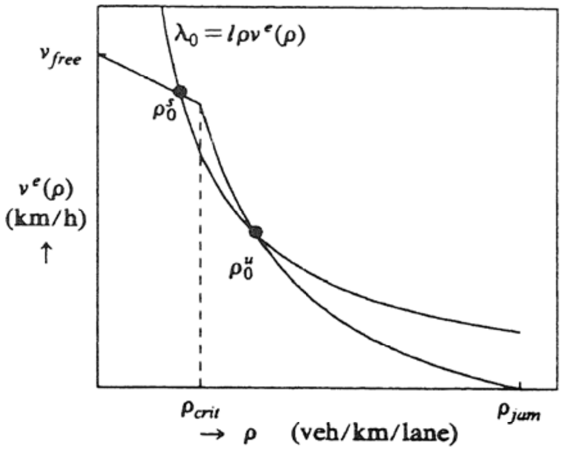

Figure 1. Equilibrium relation between density and average speed.

v_{free}: the free speed (km/h);
ρ_{crit}: the critical density (veh/km/lane);
ρ_{jam}: the jam density, at which v^e is zero (veh/km/lane);

A description of the stability properties of both models is presented in [11]. Some conclusions from this study are presented below.

Proposition 1. *Suppose* $\sigma=0$ *and* $\lambda_0<\lambda_{cap}=2\rho_{crit}v^e(\rho_{crit})$ *then the one-dimensional model has two equilibrium points.*

$$\rho_0^s = \frac{v_{free}}{2\alpha} - \sqrt{(\frac{v_{free}}{2\alpha})^2 - \frac{\lambda_0}{l\alpha}}$$

$$\rho_0^u = (1-\frac{\lambda_0}{ld})\rho_{jam}$$

where ρ_0^s *is stable and* ρ_0^u *is unstable. The domain of attraction of* ρ_0^s *is* $[-\infty,\rho_0^u)$. *For* $\rho>\rho_0^u$ ρ *converges to* ∞.
Proof. See [11, Theorems 3.1 and 3.2].

Proposition 2. *Suppose* $\sigma=\mu=0$ *and* $\lambda_0<\lambda_{cap}=2\rho_{crit}v^e(\rho_{crit})$. *Then the two-dimensional model has two equilibrium points* $(\rho_0^s,v^e(\rho_0^s))$ *and* $(\rho_0^u,v^e(\rho_0^u))$, *where* $(\rho_0^s,v^e(\rho_0^s))$ *is stable and* $(\rho_0^u,v^e(\rho_0^u))$ *is unstable. The state space is divided into two components by a curve through* $(\rho_0^u,v^e(\rho_0^u))$, *the separator. The part with* $(\rho_0^s,v^e(\rho_0^s))$ *in it is the domain of attraction of* $(\rho_0^s,v^e(\rho_0^s))$. *The separator consists of two trajectories ending in* $(\rho_0^u,v^e(\rho_0^u))$. *Trajectories starting in the other component of the state space are attracted by* "$(\infty,0)$".
Proof. See [11, page 15].

Realistic values for the parameters are:

v_{free} = 105 km/h
ρ_{crit} = 27 veh/km/lane
α = 0.58 /*fRkm*^{2}/h
σ = 14,000
ρ_{jam} = 110 veh/km/lane

A section was chosen as follows:

$$l = 2 \text{ lanes}$$
$$L = 0.5 \text{ km}$$

Control can be exerted by showing advisory speed signs to drivers. This control can either be on (1) or off (0). The effect of the homogenising control put on was modeled as a reduction of the variance σ^2, a change in the equilibrium speed $v^e(\rho)$ and a small increase in the intensity λ_0. The control dependent parameters take the values presented in table 1.

no control	control
$v^0_{free} = 105$ km/h	$v^1_{free} = 102$ km/h
$\rho^0_{crit} = 27$ veh/km/lane	$\rho^1_{crit} = 29$ veh/km/lane
	$\lambda^1_0 = \lambda^0_0 + 0.01\lambda^0_0$ veh/h
$\sigma^2_0 = 14{,}000$	$\sigma^2_1 = 11{,}000$

Table 1. Values of control dependent parameters.

The objective of control is to maximize the throughput. The conditional cost to go is therefore defined as

$$V(x_0) = E\left[\int_0^\infty e^{-ct} l\rho_t v_t dt \mid x_0\right] \tag{1.5}$$

where $x = \rho$, respectively $x = (\rho, v)$ for the one- and two-dimensional model. A discounted cost criterion was chosen to favour immediate throughput above future throughput. In this way expected congestion in the distant future does not have a major influence. This congestion is sure to occur, and delaying it from 9 to 10 years for example is not useful.

With this value function the problem can be mathematically formulated. For the one dimensional model find $u(\rho)$ $u:[0,\rho_{jam}] \to \left\{0\,,1\right\}$ that maximizes

$$V(\rho_0) = E\left[\int_0^\infty e^{-ct} l\rho_t v_t dt \mid \rho_0\right]$$

where

$$d\rho_t = \frac{1}{lL}(\lambda_0^{u(\rho)} - l\rho_t v_t)dt + \sigma_{u(\rho)} dw_t \tag{1.6}$$

$$v_t = v^e_{u(\rho)}(\rho_t) \tag{1.7}$$

The optimal control $u^*(\rho)$ and the value function $V^*(\rho)$ satisfy the Hamilton-Jacobi-Bellman equation.

$$\max_{u(\rho)=k\in\left\{0,1\right\}} \left\{ \frac{\sigma_k^2}{2}\frac{d^2V^*(\rho)}{d\rho^2} + \frac{1}{Ll}(\lambda_0^k - l\rho v_k^e(\rho))\frac{dV^*(\rho)}{d\rho} + l\rho v_k^e(\rho) \right\} - cV^*(\rho) = 0 \tag{1.8}$$

with boundary conditions

$$dV^* \frac{(0)}{d\rho} = 0, \quad V^*(\rho_{jam}) = 0.$$

For the two-dimensional model find $u(\rho,v)$ $u:[0,\rho_{jam}]\times[0,v_{\max}]\rightarrow\left\{0\,,1\right\}$ that maximizes

$$V(\rho_0,v_0) = E\left[\int_0^{\infty} e^{-ct} l\rho_t v_t dt \mid \rho_0,v_0\right]$$

where

$$\begin{aligned} d\rho_t &= \frac{1}{lL}(\lambda_0^{u(\rho,v)} - l\rho_t v_t)dt + \sigma_{u(\rho,v)}dw_t \\ dv_t &= -\frac{1}{T}(v - v^e_{u(\rho,v)}(\rho_t))dt + \mu dz_t \end{aligned} \tag{1.9}$$

Again the optimal control $u^*(\rho)$ and the value function $V^*(\rho,v)$ satisfy the Hamilton-Jacobi-Bellman equation.

$$\max_{u(\rho,v)=k\in\left\{0,1\right\}} \left\{ \frac{\sigma_k^2}{2}\frac{\partial^2 V^*(\rho,v)}{\partial\rho^2} + \frac{1}{Ll}(\lambda_0^k - l\rho v)\frac{\partial V^*(\rho,v)}{\partial\rho} + \frac{\mu^2}{2}\frac{\partial^2 V^*(\rho,v)}{\partial v^2} - \frac{1}{T}(v - v_k^e(\rho))\frac{\partial V^*(\rho,v)}{\partial v} + l\rho v \right\} - cV^*(\rho,v) = 0 \tag{1.10}$$

with boundary conditions

$$\frac{\partial V^*}{\partial v}(\rho,0)=0, \quad \frac{\partial V^*}{\partial v}(\rho,v_{\max})=0, \quad \frac{\partial V^*}{\partial\rho}(0,v)=0,$$

$$V^*(\rho_{jam},v)=0 \ \text{ for } \ 0\leqslant v\leqslant\frac{\lambda_0}{l\rho_{jam}}, \quad \frac{\partial V^*}{\partial\rho}(\rho_{jam},v)=0 \ \text{ for } \ \frac{\lambda_0}{l\rho_{jam}}<v\leqslant v_{\max}.$$

2.2. An approximating finite-state continuous-time Markov process

The partial differential equations with a maximization (1.8) and (1.10) are not solvable by ordinary techniques for partial differential equations. Discretization of the state spaces ρ_t or (ρ_t,v_t) leads to an approximation of these processes. This approximation gives a continuous-time finite-state Markov process. The Hamilton-Jacobi-Bellman equation of this Markov process can be solved by various techniques.

The processes $\{\ \rho_t\,|\,t\geqslant 0\ \}$ and $\{\ (\rho_t,v_t)\,|\,t\geqslant 0\ \}$ can be defined by their infinitesimal generator $\mathfrak{L}$.

For the one-dimensional process:

$$\mathfrak{L}(u)V(\rho) = \frac{\sigma_{u(\rho)}^2}{2}\frac{d^2V(\rho)}{d\rho^2} + (\lambda_0^{u(\rho)} - l\rho v^e_{u(\rho)}(\rho))\frac{dV(\rho)}{d\rho} \tag{2.1}$$

For the two-dimensional process:

$$\begin{aligned} \mathfrak{L}(u)V(\rho,v) = {} & \frac{\sigma_{u(\rho,v)}^2}{2}\frac{\partial^2 V(\rho,v)}{\partial\rho^2} + (\lambda_0^{u(\rho,v)} - l\rho v)\frac{\partial V(\rho,v)}{\partial\rho} \\ & + \frac{\mu^2}{2}\frac{\partial^2 V(\rho,v)}{\partial v^2} - \frac{1}{T}(v - v^e_{u(\rho,v)}(\rho))\frac{\partial V(\rho,v)}{\partial v} \end{aligned} \tag{2.2}$$

The Hamilton-Jacobi-Bellman equation for both processes can now be written as

$$\max_{u(x)\in\{0,1\}} \left\{ \mathfrak{L}(u)V(x) + l\rho v \right\} = cV(x) \tag{2.3}$$

where $x=\rho$ or $x=(\rho,v)$ and $\mathfrak{L}(u)$ the appropriate infinitesimal generator.
Approximation of

$$\frac{\partial^2 V}{\partial \rho^2}, \quad \frac{\partial V}{\partial \rho}, \quad \frac{\partial^2 V}{\partial v^2} \quad \text{and} \quad \frac{\partial V}{\partial v}$$

will lead to a discrete operator L on a finite state space G. L is the infinitesimal generator of a finite-state continuous-time Markov process. In [6] and [3] this is proved.

The one-dimensional process can be approximated in a similar way as the two-dimensional process. Below the discretization of the two-dimensional operator $\mathfrak{L}(u)$ will be shown. The discretization of the one-dimensional model is just a simplification of the following derivation.

Approximations:

$$\frac{\partial V}{\partial \rho}(\rho,v) \approx \frac{V(\rho+h_1,v)-V(\rho,v)}{h_1} \quad or \quad \frac{\partial V}{\partial \rho}(\rho,v) \approx \frac{V(\rho,v)-V(\rho-h_1,v)}{h_1} \tag{2.4}$$

$$\frac{\partial V}{\partial v}(\rho,v) \approx \frac{V(\rho,v+h_2)-V(\rho,v)}{h_2} \quad or \quad \frac{\partial V}{\partial v}(\rho,v) \approx \frac{V(\rho,v)-V(\rho,v-h_2)}{h_2} \tag{2.5}$$

$$\frac{\partial^2 V}{\partial \rho^2}(\rho,v) \approx \frac{V(\rho+h_1,v)-2V(\rho,v)+V(\rho-h_1,v)}{{h_1}^2} \tag{2.6}$$

$$\frac{\partial^2 V}{\partial v^2}(\rho,v) \approx \frac{V(\rho,v+h_2)-2V(\rho,v)+V(\rho,v-h_2)}{{h_2}^2} \tag{2.7}$$

Taking

$$(\lambda_0 - l\rho v)\frac{\partial V}{\partial \rho}(\rho,v) \approx \begin{cases} (\lambda_0 - l\rho v)\dfrac{V(\rho+h_1,v)-V(\rho,v)}{h_1} & \text{if } \lambda_0 - l\rho v \geqslant 0 \\ (\lambda_0 - l\rho v)\dfrac{V(\rho,v)-V(\rho-h_1,v)}{h_1} & \text{if } \lambda_0 - l\rho v < 0 \end{cases} \tag{2.8}$$

$$-\frac{1}{T}(v-v^e(\rho))\frac{\partial V}{\partial v}(\rho,v) \approx \begin{cases} -\dfrac{1}{T}(v-v^e(\rho))\dfrac{V(\rho,v+h_2)-V(\rho,v)}{h_2} & \text{if } v-v^e(\rho) \leqslant 0 \\ -\dfrac{1}{T}(v-v^e(\rho))\dfrac{V(\rho,v)-V(\rho,v-h_2)}{h_2} & \text{if } v-v^e(\rho) > 0 \end{cases} \tag{2.9}$$

it follows that

$$\begin{aligned} (\lambda_0 - l\rho v)\frac{\partial V}{\partial \rho} &\approx \frac{[\lambda_0 - l\rho v]^+}{h_1}V(\rho+h_1,v) - \frac{|\lambda_0 - l\rho v|}{h_1}V(\rho,v) \\ &+ \frac{[\lambda_0 - l\rho v]^-}{h_1}V(\rho-h_1,v) \end{aligned} \tag{2.10}$$

$$-\frac{1}{T}(v-v^e(\rho))\frac{\partial V}{\partial v} \approx \frac{[-\frac{1}{T}(v-v^e(\rho))]^+}{h_2}V(\rho,v+h_2) - \frac{|-\frac{1}{T}(v-v^e(\rho))|}{h_2}V(\rho,v)$$

$$+\frac{[-\frac{1}{T}(v-v^e(\rho))]^-}{h_2}V(\rho,v-h_2) \tag{2.11}$$

$$\frac{\sigma^2}{2}\frac{\partial^2 V}{\partial\rho^2} \approx \frac{\sigma^2}{2h_1{}^2}V(\rho+h_1,v)-\frac{\sigma^2}{h_1{}^2}V(\rho,v)+\frac{\sigma^2}{2h_1{}^2}V(\rho-h_1,v) \tag{2.12}$$

$$\frac{\mu^2}{2}\frac{\partial^2 V}{\partial v^2} \approx \frac{\mu^2}{2h_2{}^2}V(\rho,v+h_2)-\frac{\mu^2}{h_2{}^2}V(\rho,v)+\frac{\mu^2}{2h_2{}^2}V(\rho,v-h_2) \tag{2.13}$$

and so

$$\begin{aligned}
\mathfrak{L}V(\rho,v) \approx & \left[\frac{\mu^2}{2h_2{}^2}+\frac{[-\frac{1}{T}(v-v^e(\rho))]^+}{h_2}\right]V(\rho,v+h_2) \\
&+ \left[\frac{\sigma^2}{2h_1{}^2}+\frac{[\lambda_0-l\rho v]^-}{h_1}\right]V(\rho-h_1,v) \\
&- \left[\frac{\sigma^2}{h_1{}^2}+\frac{\mu^2}{h_2{}^2}+\frac{|\lambda_0-l\rho v|}{h_1}+\frac{|-\frac{1}{T}(v-v^e(\rho))|}{h_2}\right]V(\rho,v) \\
&+ \left[\frac{\sigma^2}{2h_1{}^2}+\frac{[\lambda_0-l\rho v]^+}{h_1}\right]V(\rho+h_1,v) \\
&+ \left[\frac{\mu^2}{2h_2{}^2}+\frac{[-\frac{1}{T}(v-v^e(\rho))]^-}{h_2}\right]V(\rho,v-h_2) \\
&= \left[LV\right]_{(\rho,v)}
\end{aligned} \tag{2.14}$$

For points on the boundary the following holds. At the reflecting boundary the points outside the boundary are replaced by their reflections. So in particular

$$\begin{aligned}
&V(\rho-h_1,v) \to V(\rho+h_1,v) && \text{if } \rho-h_1<0 \\
&V(\rho+h_1,v) \to V(\rho-h_1,v) && \text{if } \rho+h_1>\rho_{jam} \text{ and } v>\frac{\lambda_0}{l\rho_{jam}} \\
&V(\rho,v-h_2) \to V(\rho,v+h_2) && \text{if } v-h_2<0 \\
&V(\rho,v+h_2) \to V(\rho,v-h_2) && \text{if } v+h_2>v_{\max}
\end{aligned}$$

At the absorbing boundary all values of V are replaced by 0. This means

$$[LV]_{(\rho,v)}=0 \quad \text{if } \rho+h_1>\rho_{jam} \text{ and } v\leqslant\frac{\lambda_0}{l\rho_{jam}}$$

Taking a grid of n points, where

$$n=n_1\times n_2, \quad h_1=\frac{\rho_{jam}}{n_1-1}, \quad h_2=\frac{v_{\max}}{n_2-1}$$

a finite state space G is defined. The infinitesimal generator L is an approximation of the infinitesimal generator $\mathfrak{L}$ on G. Counting from $(0,v_{\max})$ to $(\rho_{jam},0)$ the index number i represents the point (ρ,v), where $\rho=h_1(i-1)$ modulo n_1 and $v=h_2[i-1/n_1]$. On this grid the functional $V(\rho,v)$ can be approximated by a n-dimensional vector V. Then

$$\mathfrak{L}V(\rho,\nu) \approx \sum_{j=1}^{n} L_{ij} V(j) \tag{2.15}$$

The choices in (2.8) and (2.9) are made to assure that the approximated process is Markov. If forward differences were always used, then $[\lambda_0 - l\rho\nu]^-$, $|\lambda_0 - l\rho\nu|$ and $[\lambda_0 - l\rho\nu]^+$ in (2.14) would be replaced by 0, $\lambda_0 - l\rho\nu$ and $\lambda_0 - l\rho\nu$. Then

$$L_{i,i+1} = \frac{\sigma^2}{2{h_1}^2} + \frac{\lambda_0 - l\rho\nu}{h_1}$$

which is in general not positive. A similar problem occurs in $L_{i,i+n_1}$. If backward differences were always used, a similar problem comes up in $L_{i,i-1}$ and $L_{i,i-n_1}$. So (2.8) and (2.9) guarantee that the non-diagonal elements in L are non-negative. This is a necessary condition for the existence of a continuous-time Markov process. Together with the fact that

$$\sum_{j=1}^{n} L_{ij} = 0$$

by construction, the infinitesimal generator L properly defines a continuous-time Markov process.

Defining $f_i = l\rho\nu|_i$ then the Hamilton-Jacobi-Bellman equation for the approximated finite-state continuous-time Markov process becomes

$$\max_{u(i)\in\{0,1\}} \left\{ f_i^{u(i)} + \sum_{j=1}^{n} L_{ij}^{u(i)} V(j) \right\} = cV(i) \qquad i = 1 \,..\, n \tag{2.16}$$

2.3. Optimal control of a finite-state Markov process

There are two basic approaches to find the optimal control u^* and the value function V^*. The first method, policy iteration, is described in [5]. This method gives the solution to the problem (2.16) in a finite number of steps. Every step however takes a "large" amount of computation. Successive approximation is described in [2] and needs an infinite number of steps. In every step however a small amount of computation is done. Also bounds for V^* can be computed which can be used in a stopping rule of the algorithm.

Combination of these methods leads to a variety of modified policy iteration methods. The simplest of them is described in [7]. The aim of these methods is to get faster converging bounds for V^* then successive approximation does. When this is possible with a "moderate" amount of computations per step, the method can be faster then policy iteration. Moreover modified policy iteration algorithms need less storage then policy iteration does. The combined method is rather general. Special cases of this algorithm are again policy iteration and successive approximation.

With policy iteration it is possible to solve (2.16) directly. Successive approximation and modified policy iteration methods exist only for discrete-time Markov processes or Markov chains. So it is necessary to transform (2.16) into

$$\max_{u(i)\in\{0,1\}} \left\{ g_i^{u(i)} + \alpha \sum_{j=1}^{n} P_{ij}^{u(i)} V(j) \right\} = V(i) \qquad i = 1 \,..\, n \tag{3.1}$$

where

$$|L| = \max_{i,j,k} |L_{ij}^k| \tag{3.2}$$

$$g_i^k = \frac{f_i^k}{c + |L|} \tag{3.3}$$

$$\alpha = \frac{|L|}{c + |L|} \tag{3.4}$$

$$P_{ij}^k = \frac{L_{ij}^k}{|L|} + \delta_{ij} \tag{3.5}$$

The transformed equation (3.1) is the same as would be obtained for a discrete-time Markov processes. This discrete-time Markov problem has time-step $\Delta t = 1/|L|$, a transition matrix $P_{ij}^{u(i)}$, immediate rewards $g_i^{u(i)}$ and discount factor α. Equation (3.1) is the Hamilton-Jacobi-Bellman equation for the discrete-time Markov problem. The equivalence of (2.16) and (3.1) can easily be verified.

2.3.1. Policy iteration

Policy iteration is described in [5]. For a continuous-time Markov process and a discounted value function policy iteration yields the following.

ALGORITHM 1: POLICY IMPROVEMENT

INITIAL STEP: Choose an initial policy $u_1(i)$.

ITERATION STEP I: Solve for V

$$cV^m(i) = f_i^u + \sum_{j=1}^n L_{ij}^u V^m(j) \Leftrightarrow \sum_{j=1}^n (L_{ij}^u - c\delta_{ij}) V^m(j) = -f_i^u \qquad i = 1 \,..\, n$$

where $u = u_m(i)$.

ITERATION STEP II: For every state i, find the alternative $u_{m+1}(i)$ that satisfies

$$f_i^u + \sum_{j=1}^n L_{ij}^u V^m(j) = \max_{k \in \{0,1\}} \left\{ f_i^k + \sum_{j=1}^n L_{ij}^k V^m(j) \right\}$$

where $u = u_{m+1}(i)$.

STOPPING RULE: End the iterative algorithm when the new found policy is the same as the former policy.

The policy iteration algorithm is such that at every iteration the value of the former policy is determined by solving a linear system of n equations and n unknowns, and finding the best possible policy for this value by n maximizations over the control space. The next propositions will lead to the correctness of this algorithm. The correctness is proved in [5]. The proof is included here for reason of exposition.

PROPOSITION 2. *Let V^m for $m = 1,2,..$ be a sequence of vectors produced by the policy iteration algorithm. Then*

$$V^{m+1}(i) \geqslant V^m(i) \qquad i = 1 \,..\, n$$

PROOF. By construction in iteration step II

$$f_i^{u_{m+1}} + \sum_{j=1}^n L_{ij}^{u_{m+1}} V^m(j) \geqslant f_i^{u_m} + \sum_{j=1}^n L_{ij}^{u_m} V^m(j) \qquad i = 1 \,..\, n$$

So

$$\gamma_i = f_i^{u_{m+1}} + \sum_{j=1}^n L_{ij}^{u_{m+1}} V^m(j) - f_i^{u_m} - \sum_{j=1}^n L_{ij}^{u_m} V^m(j) \geqslant 0 \qquad i = 1 \,..\, n \tag{*}$$

γ_i is the improvement in iteration step II of state i.

By construction in iteration step I

$$cV^m(i) = f_i^{u_m} + \sum_{j=1}^{n} L_{ij}^{u_m} V^m(j) \qquad i = 1 \,..\, n$$

and

$$cV^{m+1}(i) = f_i^{u_{m+1}} + \sum_{j=1}^{n} L_{ij}^{u_{m+1}} V^{m+1}(j) \qquad i = 1 \,..\, n$$

Together we get

$$c(V^{m+1}(i) - V^m(i)) = \gamma_i + \sum_{j=1}^{n} L_{ij}^{u_{m+1}} (V^{m+1}(j) - V^m(j)) \qquad i = 1 \,..\, n$$

Or, equivalently,

$$c\Delta V_i = \gamma_i + \sum_{j=1}^{n} L_{ij}^{u_{m+1}} \Delta V_j \qquad i = 1 \,..\, n$$

So

$$\gamma = (cI - L)\Delta V \qquad \Delta V = (cI - L)^{-1}\gamma \tag{**}$$

And while

$$(cI - L)^{-1} = \frac{1}{c}\sum_{n=1}^{\infty} (\frac{1}{c}L)^n$$

every element of $(cI - L)^{-1}$ is positive. From this, (*) and (**) follows that every element of ΔV is non-negative. This is the same as

$$V^{m+1}(i) \geqslant V^m(i) \qquad i = 1 \,..\, n$$

□

Proposition 3. *Let V^m $m = 1,2,..$ be a sequence of vectors produced by the policy improvement algorithm. Let V^* be the solution of (2.16). Then*

$$V^m(i) \leqslant V^*(i) \qquad i = 1 \,..\, n \qquad m = 0,1,...$$

Proof. By definition of V^* and u^*:

$$f_i^{u_m} + \sum_{j=1}^{n} L_{ij}^{u_m} V^*(j) \leqslant f_i^{u^*} + \sum_{j=1}^{n} L_{ij}^{u^*} V^*(j) \qquad i = 1 \,..\, n$$

and

$$cV^*(i) = f_i^{u^*} + \sum_{j=1}^{n} L_{ij}^{u^*} V^*(j) \qquad i = 1 \,..\, n$$

The rest of the proof is similar to the proof of proposition 1. □

Theorem 4. *Let V^m $m = 1,2,..$ be a sequence of vectors produced by the policy improvement algorithm. Let V^* be the solution of (2.16). Then*

$$V^M(i) = V^*(i) \qquad i = 1 \,..\, n \qquad \text{for } some\ M > 0$$

Proof. From proposition 1 and 2 it follows that $V^m \rightarrow V^*$. Because of the finiteness of the collection of controls, and the fact that at every step a control is chosen with a higher pay-off, in a finite number of steps the algorithm will converge.

□

The linear system that has to be solved in iteration step I requires a large amount of computations. Solving goes directly by using a LU-decomposition. The algorithm for this decomposition, and the solution of the system is taken from [4]. Note that in the one-dimensional case the system is tridiagonal, in the two-dimensional case it is banded with upper and lower band width n_1.

2.3.2. Successive approximation

Successive approximation is based on a contraction mapping. This contraction mapping has as fixed point the value function. Beginning with a vector V^0 the contraction produces a better approximation of V^*. The optimal control policy according to the approximated value function is also computed. Below one finds a summary of the properties of the successive approximation method. The correctness of this can be found in [2, § 5.2]. The contraction mapping is defined as

$$T(V)(i) = \max_{k \in \{0,1\}} \left\{ g_i^k + \alpha \sum_{j=1}^{n} P_{ij}^k V(j) \right\} \qquad i = 1 .. n \tag{3.6}$$

where $u(i)=k$ is the approximated optimal control policy. After every iteration bounds for the fixed point can be computed.

$$\begin{aligned} T^m(V)(i) + c_m &\leqslant T^{m+1}(V)(i) + c_{m+1} \leqslant V^*(i) \\ &\leqslant T^{m+1}(V)(i) + \bar{c}_{m+1} \leqslant T^m(V)(i) + \bar{c}_m \end{aligned} \tag{3.7}$$

where

$$c_m = \frac{\alpha}{1-\alpha} \min_{i=1..n} \left[T^m(V)(i) - T^{m-1}(V)(i) \right] \tag{3.8}$$

$$\bar{c}_m = \frac{\alpha}{1-\alpha} \max_{i=1..n} \left[T^m(V)(i) - T^{m-1}(V)(i) \right] \tag{3.9}$$

The successive approximations algorithm is

ALGORITHM 5: SUCCESSIVE APPROXIMATION

INITIAL STEP: $V^0 \equiv 0$

ITERATION STEP: $V^{m+1}(i) = T(V^m)(i) \qquad i = 1 .. n$

STOPPING RULE: Repeat this iteration step until $\bar{c}_m - c_m < \epsilon$.

Here ϵ is the maximal allowed error of V^*.

It is also possible to estimate the value of a given control policy.

$$T_u(V)(i) = \left\{ g_i^{u(i)} + \alpha \sum_{j=1}^{n} P_{ij}^{u(i)} V(j) \right\} \qquad i = 1 .. n \tag{3.10}$$

2.3.3. Modified policy iteration methods

When comparing both methods, policy iteration and successive approximation, the following similarity can be found. Both methods consist of constructing better approximations of the value function. Subsequently the control policy for this functional is computed. Where policy iteration solves a linear system of equations to find a larger value for the approximation of V^* successive approximation takes just one step of an iterative procedure for solution of the system. The question arises if several iterative steps would still give a contracting mapping. Another question is whether faster iterative scheme's can be found. A general iterative scheme that solves (3.1) has the following form

$$V^{m+1}(i) = g_i^{u(i)} + \alpha \left[\sum_{j=1}^{s} P_{ij}^{u(i)} V^{m+1}(j) + \sum_{j=s+1}^{n} P_{ij}^{u(i)} V^m(j) \right] \tag{3.11}$$

Where s is an integer between 0 and n. The modified policy algorithm becomes

ALGORITHM 6: MODIFIED POLICY ITERATION

INITIAL STEP: $V^0 \equiv 0$

ITERATION STEP I: For every state i, find the control $u_{m+1}(i)$ that satisfies:

$$g_i^u + \alpha \sum_{j=1}^{n} P_{ij}^u V^m(j) = \max_{k \in \{0,1\}} \left\{ g_i^k + \alpha \sum_{j=1}^{n} P_{ij}^k V^m(j) \right\}$$

where $u = u_{m+1}(i)$

ITERATION STEP II: $V^{m+1}(i) = \mathcal{T}_u^M(V^m)(i)$

where M is the number of iterations used for "solution" of the linear system of equations and

$$\mathcal{T}_u(V)(i) = g_i^u + \alpha \left[\sum_{j=1}^{s} P_{ij}^u \mathcal{T}(V)(j) + \sum_{j=s+1}^{n} P_{ij}^u V(j) \right]$$

where $u = u_{m+1}(i)$

When $s = 0$ the iteration can be denoted as

$$V^{m+1}(i) = g_i^{u(i)} + \alpha \sum_{j=1}^{n} P_{ij}^{u(i)} V^m(j)$$

or

$$V^{m+1}(i) = T_u(V^m(i))$$

where T_u as defined in (3.10).

With $M = 1$ this is the successive approximation algorithm again. In [2] this is called the Jacobi form of successive approximation. For $M > 1$ this method can be regarded as policy iteration where the linear system is "solved" iteratively using M steps of Jacobi iteration. For $M > 1$ this method is worked out in [7]. There this method is called *modified policy iteration*. The correctness of this method is given in [7, Theorem 1].

When $s = i - 1$ we can denote

$$V^{m+1}(i) = g_i^{u(i)} + \alpha \left[\sum_{j=1}^{i-1} P_{ij}^{u(i)} V^{m+1}(j) + \sum_{j=i}^{n} P_{ij}^{u(i)} V^m(j) \right]$$

or

$$V^{m+1}(i) = F_u(V^m(i))$$

When $M = 1$ this is called the Gauss-Seidel form of the successive approximation method in [2]. For $M > 1$ this method can be seen as policy iteration where the linear system is solved iteratively using M steps of Gauss-Seidel iteration.

When $s = i$ we can denote

$$V^{m+1}(i) = g_i^{u(i)} + \alpha \left[\sum_{j=1}^{i} P_{ij}^{u(i)} V^{m+1}(j) + \sum_{j=i+1}^{n} P_{ij}^{u(i)} V^m(j) \right]$$

or

$$V^{m+1}(i) = \frac{g_i^{u(i)} + \alpha\left[\sum_{j=1}^{i-1} P_{ij}^{u(i)} V^{m+1}(j) + \sum_{j=i+1}^{n} P_{ij}^{u(i)} V^m(j)\right]}{1-\alpha P_{ii}^{u(i)}}$$

or

$$V^{m+1}(i) = \omega_i F_u(V^m(i)) + (1-\omega_i)V^m(i)$$

where

$$\omega_i = \frac{1}{1-\alpha P_{ii}^{u(i)}}$$

When $M=1$ this is the successive overrelaxation form of the successive approximation method. For $\omega^* = \min \omega_i$ and replacing all ω_i by ω^* correctness of that method is proved in [8, Theorems 1 and 2]. The only necessity for choosing ω is

$$1-\omega+\omega\, \alpha\, P_{ii}^{u(i)} \geqslant 0$$

which hold for ω^* as well as for ω_i. For $M>1$ this method can be seen as policy iteration where the linear system is solved iteratively using M steps of successive overrelaxation iteration.

When $s=i+1$ we can denote

$$V^{m+1}(i) = g_i^{u(i)} + \alpha\left[\sum_{j=1}^{i+1} P_{ij}^{u(i)} V^{m+1}(j) + \sum_{j=i+2}^{n} P_{ij}^{u(i)} V^m(j)\right]$$

The structure of $P_{ij}^{u(i)}$ is such that

$$P_{ij}^{u(i)} = 0 \quad \text{if } j \neq i-n_1, i-1, i, i+1, i+n_1$$

$$P_{ij}^{u(i)} = 0 \quad \text{if } i=ln_1+1 \text{ and } j=i-1 \quad \textit{or} \quad i=(l+1)n_1 \text{ and } j=i+1$$

$$\text{and } l=0..n_2-1$$

Or P^u is a $n \times n$ block-tridiagonal matrix, with $n_2 \times n_2$ blocks of size $n_1 \times n_1$. The diagonal blocks are of tridiagonal form, and the blocks of upper and lower diagonals are diagonal blocks. Then (3.1) simplifies to the formula's

for $l = 0 \,..\, n_2-1$

$$V^{m+1}(i) = g_i^{u(i)} + \alpha\left[P_{i,i-n_1}^{u(i)} V^{m+1}(i-n_1) + \sum_{j=i}^{i+1} P_{ij}^{u(i)} V^{m+1}(j) + P_{i,i+n_1}^{u(i)} V^m(i+n_1)\right]$$

for $i = ln_1+1$

$$V^{m+1}(i) = g_i^{u(i)} + \alpha\left[P_{i,i-n_1}^{u(i)} V^{m+1}(i-n_1) + \sum_{j=i-1}^{i+1} P_{ij}^{u(i)} V^{m+1}(j) + P_{i,i+n_1}^{u(i)} V^m(i+n_1)\right]$$

for $i = ln_1+2 \,..\, (l+1)n_1-1$

$$V^{m+1}(i) = g_i^{u(i)} + \alpha\left[P_{i,i-n_1}^{u(i)} V^{m+1}(i-n_1) + \sum_{j=i-1}^{i} P_{ij}^{u(i)} V^{m+1}(j) + P_{i,i+n_1}^{u(i)} V^m(i+n_1)\right]$$

for $i = (l+1)n_1$

For $M>1$ this method can be seen as a modified policy iteration method where the linear system is solved iteratively using a block iterative scheme described in [12]. For every l a tridiagonal system of size n_1 has to be solved.

Of the four methods presented in this section the convergence rate of the last one is the lowest.

3. THE ONE-DIMENSIONAL MODEL

As stated in chapter 2 the one-dimensional model was formulated as:

$$d\rho_t = \frac{1}{lL}(\lambda_0^{u(\rho)} - l\rho_t v_t)dt + \sigma_{u(\rho)} dw_t$$

$$v_t = v_{u(\rho)}^e(\rho_t)$$

with cost-function:

$$V(\rho_0) = E\left[\int_0^\infty e^{-ct} l\rho_t v_t dt \mid \rho_0\right]$$

The problem is to find the optimal control $u(\rho)$ $u:[0,\rho_{jam}] \rightarrow \{0,1\}$ for this criterion. The optimal control $u^*(\rho)$ and the value-functional $V^*(\rho)$ satisfy the Hamilton-Jacobi-Bellman equation.

$$\max_{u(\rho)\in\{0,1\}} \left\{ \frac{\sigma_k^2}{2}\frac{d^2V^*(\rho)}{d\rho^2} + \frac{1}{Ll}(\lambda_0^k - l\rho v_k^e(\rho))\frac{dV^*(\rho)}{d\rho} + l\rho v_k^e(\rho) \right\} - cV^*(\rho) = 0$$

With boundary conditions:

$$\frac{dV^{*(0)}}{d\rho} = 0, \quad V^*(\rho_{jam}) = 0.$$

After discretization the resulting finite-state continuous-time Markov process has the infinitesimal generator

$$\begin{aligned}[L^uV]_i = &\left(\frac{\sigma_{u(i)}}{2h^2} + \frac{[\lambda_0 - l\rho v]^-}{h}\right)V(i-1) \\ &- \left(\frac{\sigma_{u(i)}}{h^2} + \frac{|\lambda_0 - l\rho v|}{h}\right)V(i) \\ &+ \left(\frac{\sigma_{u(i)}}{2h^2} + \frac{[\lambda_0 - l\rho v]^+}{h}\right)V(i+1)\end{aligned}$$

The Hamilton-Jacobi-Bellman equation for this Markov process is:

$$\max_{u(i)\in\{0,1\}} \left\{ f_i^{u(i)} + \sum_{j=1}^{n} L_{ij}^{u(i)} V(j) \right\} = cV(i) \qquad i = 1 \,..\, n$$

All methods as described in the former chapter to solve this equation were used for different values of λ_0 and h. The one-dimensional model consists of one block, so the block iterative modified policy improvement method is equivalent to policy improvement. All methods led to the same control law. As stopping rule for the successive approximation method and for the modified policy iteration methods a maximum error of $\epsilon = 10$ was allowed in the value function. The tridiagonality of L^u makes the LU-decomposition of $(L^u - cI)$ very fast. This decomposition did not use more memory then the other methods used. The LU-decomposition was done by a method described in [4]. Because of the diagonal dominance of $(L^u - cI)$ pivoting is not necessary. The high value of the contraction factor α, which is about 0,997 makes both successive approximation and modified policy iteration methods very slow. The latter methods however have the advantage of giving bounds for V^*. Policy iteration gives a value function with is numerical not exact because of the rounding errors in the solution of the linear system of equations.

3.1. Optimal control
First of all the estimation of $V^*(\rho_0^\delta)$ will be presented.

λ_0 \ h	1	0.5	0.25
4000	1315	1369	1392
3000	5910	5949	5965
2000	4265	4246	4237

Table 2. The value function for several traffic intensities (λ_0)
and several discretization steps (h).

The optimal control for $h = 0.5$ is indicated in figure 2.

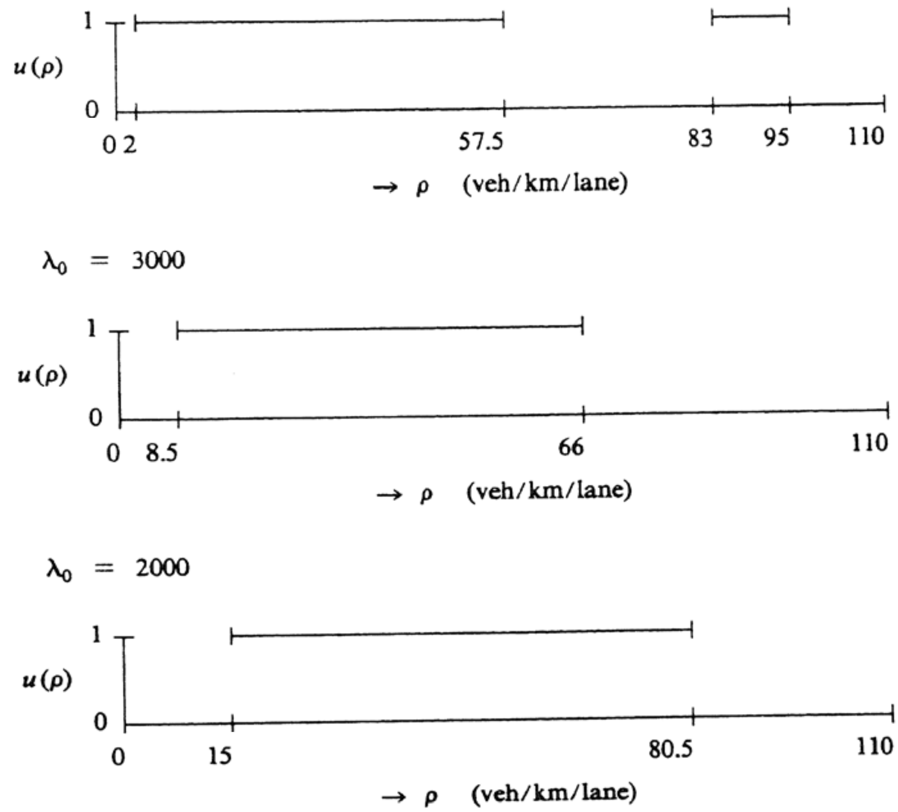

Figure 2. Optimal control for the one-dimensional model.

3.2. Suboptimal control

Our interest is especially directed to the area between the stable equilibrium point ρ_0^s and ρ_{crit}. Most of the time we find the optimal control there in "on" position. But applying control to often is counter productive. The model is based on supposed driver behaviour. Applying control too often can change this behaviour. We want the homogenising control switched off as long as possible. We therefore look for a class of control policies which hopefully are realistic. The class that is used is the class of one-switch policies: turning the control on for densities above or equal to $\overline{\rho}$ and off below $\overline{\rho}$. In this class the control is determined by demanding that the control should be switched on at as high a density $\overline{\rho}$ as possible, while the cost function of this control law in the stable equilibrium point should be as high as 95% of the value function in this equilibrium point. So these policies have to satisfy the following criteria. First $V_{\overline{\rho}}(\rho_0^s) \geqslant 0.95\ V^*(\rho_0^s)$. Second $\overline{\rho}$ must be as large as possible.

Solving the linear system

$$cV = f^u + L^u V \Leftrightarrow (L - cI)V = -f$$

gives us the value V according to the control u. The results for $h = 0.5$ of these computations are shown in the next table.

$\lambda_0 = 4000$	
$\overline{\rho}$	$V(\rho_0^s)$
26	1333
27	1326
28	1312
29	1278
30	1229

Table 3. Cost of one-switch controls for traffic intensity $\lambda_0 = 4000$.

While $V^*(ev) = 1369$, so 95% $V^*(ev) = 1301$ the value of $\overline{\rho}$ that is acceptable is $\overline{\rho} = 28$.

$\lambda_0 = 3000$	
$\overline{\rho}$	$V(\rho_0^s)$
37	5712
38	5690
39	5667
40	5643
41	5619

Table 4. Cost of one-switch controls for traffic intensity $\lambda_0 = 3000$.

While $V^*(ev) = 5949$, so 95% $V^*(ev) = 5652$ the value of $\overline{\rho}$ that is acceptable is $\overline{\rho} = 39$.

$\lambda_0 = 2000$	
$\overline{\rho}$	$V(\rho_0^s)$
79	4243.78946
80	...894
81	...878
82	...892
83	...929

Table 5. Cost of one-switch controls for traffic intensity $\lambda_0 = 2000$.

While $V^*(ev) = 4246$, so 95% $V^*(ev) = 4039$ every value of $\overline{\rho}$ is acceptable, while the minimal value is reached for $\bar{\rho} = 81$. This point is the point where the control is turned off again in the optimal control policy. The minimal value is still high enough.

3.3. Conclusions
For several values of the intensity λ_0 the optimal control is estimated. The optimal control has two properties which in practice could lead to unwanted, and unmodelled effects. The control is turned off when congestion is very likely to occur. The control is turned on even at very low density. To avoid these problems we have considered a class of one-switch policy's. Acceptable control strategies have been found. At low intensity the control is never applied, at high intensity switch-points are determined.

4. THE TWO-DIMENSIONAL MODEL
As stated in chapter 2 the two-dimensional model is formulated as

$$d\rho_t = \frac{1}{lL}(\lambda_0^{u(\rho,v)} - l\rho_t v_t)dt + \sigma_{u(\rho,v)}dw_t$$

$$dv_t = -\frac{1}{T}(v_t - v^e_{u(\rho,v)}(\rho_t))dt + \mu dz_t$$

with cost function

$$V(\rho_0,v_0) = E\left[\int_0^\infty e^{-ct} l\rho_t v_t \mid \rho_0,v_0\right]$$

The problem is to find the optimal control $u(\rho,v)$ $u:[0,\rho_{jam}]\times[0,v_{\max}]\rightarrow \{0,1\}$ such that this criterion is maximized. The optimal control $u^*(\rho,v)$ and the value function $V^*(\rho,v)$ satisfy the Hamilton-Jacobi-Bellman equation

$$\max_{u(\rho,v)\in\{0,1\}}\left\{\frac{\sigma_k^2}{2}\frac{\partial^2 V^*(\rho,v)}{\partial\rho^2} + \frac{1}{Ll}(\lambda_0^k - l\rho v)\frac{\partial V^*(\rho,v)}{\partial\rho}\right.$$

$$\left. + \frac{\mu^2}{2}\frac{\partial^2 V^*(\rho,v)}{\partial v^2} - \frac{1}{T}(v - v_k^e(\rho))\frac{\partial V^*(\rho,v)}{\partial v} + l\rho v\right\} - cV^*(\rho,v) = 0$$

with boundary conditions

$$\frac{\partial V^*}{\partial v}(\rho,0)=0,\quad \frac{\partial V^*}{\partial v}(\rho,v_{\max})=0,\quad \frac{\partial V^*}{\partial\rho}(0,v)=0,$$

$$V^*(\rho_{jam},v)=0 \text{ for } 0\leqslant v\leqslant\frac{\lambda_0}{l\rho_{jam}},\quad \frac{\partial V^*}{\partial\rho}(\rho_{jam},v)=0 \text{ for } \frac{\lambda_0}{l\rho_{jam}}\leqslant v\leqslant v_{\max}.$$

After discretization the resulting finite-state continuous-time Markov process has the infinitesimal generator

$$\left[LV\right]_{(\rho,v)} = \left\{\frac{\mu^2}{2h_2^2} + \frac{[-\frac{1}{T}(v - v^e(\rho))]^+}{h_2}\right\}V(\rho,v+h_2)$$

$$+ \left\{\frac{\sigma^2}{2h_1^2} + \frac{[\lambda_0 - l\rho v]^-}{h_1}\right\}V(\rho-h_1,v)$$

$$- \left\{\frac{\sigma^2}{h_1^2} + \frac{\mu^2}{h_2^2} + \frac{|\lambda_0 - l\rho v|}{h_1} + \frac{|-\frac{1}{T}(v - v^e(\rho))|}{h_2}\right\}V(\rho,v)$$

$$+ \left\{ \frac{\sigma^2}{2h_1{}^2} + \frac{[\lambda_0 - l\rho\nu]^+}{h_1} \right\} V(\rho+h_1,\nu)$$

$$+ \left\{ \frac{\mu^2}{2h_2{}^2} + \frac{[-\frac{1}{T}(\nu - \nu^e(\rho))]^-}{h_2} \right\} V(\rho,\nu - h_2)$$

The Hamilton-Jacobi-Bellman equation for this approximated finite-state continuous-time Markov process is

$$\max_{u(i)\in\{0,1\}} \left\{ f_i^{u(i)} + \sum_{j=1}^{n} L_{ij}^{u(i)} V(j) \right\} = cV(i) \qquad i = 1 \,..\, n$$

To solve this problem policy iteration and modified policy iteration methods were used. For a very small discretization step the policy iteration method could not be used because the LU-decomposition required too much storage. For a larger discretization step, however, it is one of the fastest methods. Only the block iterative version of modified policy iteration was as fast as policy iteration, the other three were slower. Of these three the successive over-relaxation version was satisfactory, the other two were too slow. The amount of storage for all four modified policy methods were of the same order, of the $o(n)$. They all used far less then the policy iteration algorithm, which used a memory of the order $o(n_1{}^2 n_2)$.

4.1. Optimal control
The estimation of $V^*(\rho_0^s, \nu^e(\rho_0^s))$ is presented in the next tables.

$\lambda_0 = 4800$

h_1	h_2	$V^*(\rho_0^s,\nu^e(\rho_0^s))$
10	10	743.7
5	5	775.3
2	5	785
2	2	792
1.25	2	800

Table 6. Value function for traffic intensity $\lambda_0 = 4800$.

$\lambda_0 = 4000$

h_1	h_2	$V^*(\rho_0^s,\nu^e(\rho_0^s))$
10	10	2747
5	5	3677
5	3	3861

Table 7. Value function for traffic intensity $\lambda_0 = 4000$.

The optimal control is as indicated in figure 3.

4.2. Suboptimal control
Again a search is made to find a simple control, which still has a value in ρ_0^s which is at least 95% of the value function in the equilibrium point $(\rho_0^s, \nu^e(\rho_0^s))$. We will consider control policies in which the state space is partioned into two parts by a straight line: at one side homogenising control is applied, at the other side it is not. The following three classes of policies are considered:

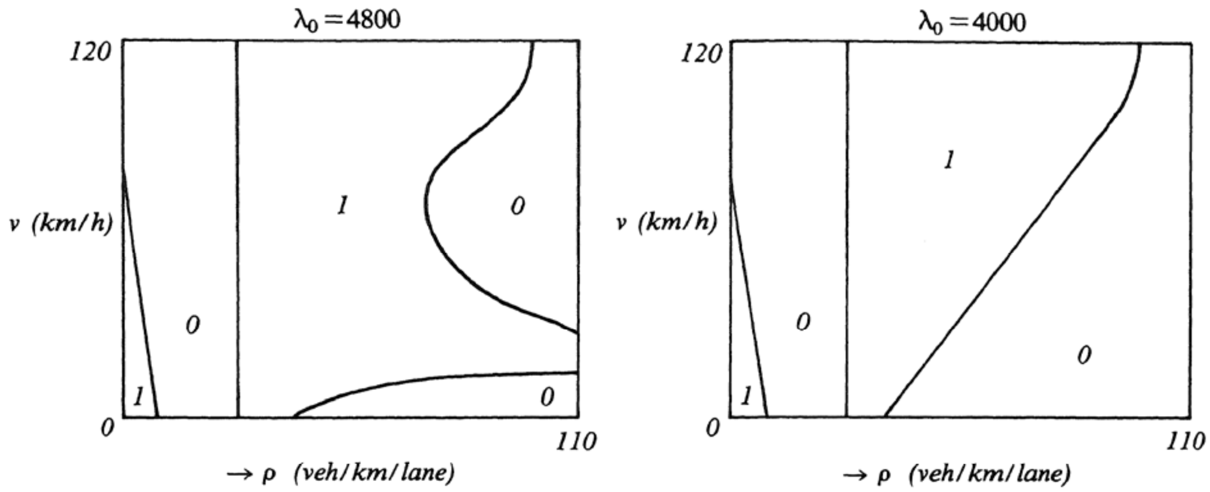

Figure 3. Optimal control for the two-dimensional model.

$$u(\rho,\nu) = \begin{cases} 0 & \rho < \bar{\rho} \\ 1 & \rho \geqslant \bar{\rho} \end{cases} \tag{2.1}$$

$$u(\rho,\nu) = \begin{cases} 0 & \nu < \bar{\nu} \\ 1 & \nu \geqslant \bar{\nu} \end{cases} \tag{2.2}$$

$$u(\rho,\nu) = \begin{cases} 0 & \nu \geqslant \gamma\rho + b \\ 1 & \nu < \gamma\rho + b \end{cases} \tag{2.3}$$

Where γ is defined as the tangent of the separator in the unstable equilibrium point ρ_0^u, see section 2.1 proposition 2. The determination of γ is done in [11], and $\gamma \approx 0.7$ for intensity $\lambda_0 = 4000$. So the third class can be seen as a shift of the linear approximation of the separator. For $\lambda_0 = 4000$ and $h_1 = h_2 = 5$ the results are

class 2.1		class 2.2		class 2.3	
$\bar{\rho}$	$V(\rho_0^s, \nu^e(\rho_0^s))$	$\bar{\nu}$		b	
0	3459.1	40	2596.7	-60	2570.7
5	3459.6	45	2639.9	-50	2570.7
10	3462.1	50	2706.4	-40	2570.4
15	3472.3	55	2797.5	-30	2569.4
20	3508.0	60	2905.8	-20	2566.6
25	3592.7	65	3033.0	-10	2563.2
30	3661.9	70	3160.4	0	2567.2
35	3198.6	75	3281.3	10	2594.2
40	2869.3	80	3390.0	20	2659.8
45	2680.9	85	3463.4	30	2779.2
50	2596.4	90	3483.7	40	2955.2
55	2568.1	95	3474.2	50	3193.9
		100	3463.9	60	3425.8
		105	3460.8	70	3550.3
				80	3515.0

Table 8. Values of cost functions for several classes of suboptimal controls.

In every class we can determine an optimal control. Also the control can be determined which postpones the use of homogenising control, but still has a cost function which is high enough. Results of the considered classes are presented in the next table.

class 2.1	$\bar{\rho}$	$V / V^*(\rho_0^s, \nu^e(\rho_0^s))$
optimal	30	99.6 %
	35	87.0 %
class 2.2	$\bar{\nu}$	
optimal	90	94,8 %
	85	94.2 %
	80	92.2 %
class 2.3	b	
optimal	70	96.6 %
	60	93.2 %
	50	86.9 %

Table 9. Comparison of cost functions for optimal and suboptimal controls.

4.3. Conclusions

For different values of the intensity λ_0 the optimal control is determined. Several classes of controls are investigated, and suboptimal controls are determined in these classes. As the optimal control already suggests, control based on density only yields a high value. So even in the two-dimensional model a useful control may be based on the density only. Seen of the point of view of control the second state variable does not give extra information. This to our surprise as we expected that the separator would play a role in the control.

5. Conclusions

5.1. Control

In this report two models for freeway traffic flow have been investigated. For both models optimal control strategies have been computed. Different classes of control have been investigated and suboptimal controls for these classes have been determined. A general conclusion is that one-switch control based on density is both a useful and simple type of control. The switching value $\overline{\rho}$ still depends on the intensity at the entrance of the freeway section considered.

5.2. The method

The approach used to compute the optimal control of the stochastic differential equations seems to work well. The discretization of the state space, to get a finite-state Markov process was already used in [1] with success. A direct method to solve the linear system of equations was used there to minimize the number of iterative steps in policy improvement. For in [1] a continuous control space was considered which leads to a higher amount of computations in each iterative step. With a finite control space this problem is far less important. Iterative methods can be used which use more modified policy improvement steps, although each step takes less computation time. These modified policy improvement methods need less storage, so larger problems can be solved and a finer discretization may be used. The special structure of the two-dimensional Markov process suggests the use of a block iterative method for solving the linear system of equations. This method indeed converges fast and uses a reasonable amount of memory.

5.3. Suggestions for research

The models which have been investigated in this report were simplifications of the model proposed in [9]. Further research can be done on the effect of control in one section on the traffic flow in another section. A possible approach is to make a two-dimensional, two-section model based on the density in both sections only. Such a model can be described by the stochastic differential equation

$$\begin{aligned} d\rho_{1,t} &= \frac{1}{L_1 l_1}(\lambda_0 - l_1 \rho_{1,t} v_{1,t})dt \ + \ \sigma dw_{1,t} \\ v_{1,t} &= v^e(\rho_{1,t}) \\ d\rho_{2,t} &= \frac{1}{L_2 l_2}(l_1 \rho_{1,t} v_{1,t} - l_2 \rho_{2,t} v_{2,t})dt \ + \ \sigma dw_{2,t} \\ v_{2,t} &= v^e(\rho_{2,t}) \end{aligned}$$

The same methods as used in this report can be applied to investigate this problem. Taking more sections or a extensions of the model with a stochastic differential equation for the mean speed leads to a model with higher dimensions. The same methods are applicable in principal to these models, but they will lead to linear systems of equations of very high order.

Another approach to obtain conclusions about the effect of control in a model of many sections, is by way of simulation. Several control laws can be investigate based on density in one or more sections. In every section control can be based on density in that section, or overall control can be based on density in one special section. The simulations could give a view into the effects on the traffic flow upstream and downstream.

Acknowledgment

I like to thank Stef Smulders and Jan van Schuppen. Without their support I would not accomplished this research. For the revision of this report I want to thank Jan van Schuppen and Henk Nijmeijer. Their suggestions and comment make it possible to present this report.

Sjaak Schuit

REFERENCES

1. S. BELLIZI, R. BOUC, F. CAMPILLO, and E. PARDOUX. *Controle Optimal Semi-actif de Suspension de Vehicule*.
2. D.P. BERTSEKAS (1987). *Dynamic Programming: Deterministic and Stochastic Models*, Prentice-Hall, Englewood Cliffs.
3. F. CAMPILLO (1988). *Resolution Numerique de l'Equation de Hamoliton Jacobi Bellman Ergodique en Dimension 2*, Notes de travail
4. G.H. GOLUB and C.F. VAN LOAN (1983). *Matrix Computations*, John Hopkins University Press, Baltimore.
5. R.A. HOWARD (1960). *Dynamic Programming and Markov Processes*, John Wiley, New York.
6. H.J. KUSHNER (1977). *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Academic Press, New York.
7. M.L. PUTERMAN and MOON CHIRL-SHIN (1978). Modified Policy Iteration Algorithmes for Discounted Markov Decision Problems, *Management Sci*, 24, 1127-1137.
8. D. REETZ (1973). Solution of a Markovian Decision Problem by Successive Overrelaxation, *Zeitschrift fur Operations Research*, 17, 29-32.
9. S.A. SMULDERS (1986). *Modelling and Simulation of Freeway Traffic Flow*, Centre for Mathematics and Computer Science, Amsterdam, Report OS-R8615.
10. S.A. SMULDERS (1988). *Filtering of Freeway Traffic Flow*, Centre for Mathematics and Computer Science, Amsterdam, Report OS-R8806.
11. S.A. SMULDERS (1988). *Control of Freeway Traffic Flow*, Centre for Mathematics and Computer Science, Amsterdam, Report OS-R8817.
12. J. STOER and R. BULIRSCH (1980). *Introduction to Numerical Analysis*, Springer-Verlag, New York.