STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 367

# Driving with Markov programming

by

Prof.dr. G. de Leve

and

ir. P.J. Weeda

augustus 1966

The Mathematical Centre at Amsterdam, founded the 11th of February, 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

## 1. Introduction

Suppose we consider the following problem:

A motorist has decided to effect an accident insurance under the following conditions. The insurance runs for one year. The premium for the first year amounts $E_0$. If no damages are claimed during the first year, a premium reduction is granted. The premium to be paid will amount $E_1 < E_0$. After two years without claiming a second reduction is granted and after three years a third. After three years no further premium reduction is permitted. The premium is due on the first day of the year. The own risk amounts $a_0$.

The number of accidents of our motorist during a time period T obeys a Poissondistribution with parameter $\lambda T$. The extent of the damage $\underline{s}$ has distribution function $F(s)$ with finite mean and variance.

The problem of our motorist will be to decide whether to claim a damage or not. He will have to develop a strategy that specifies his decisions in every possible situation. His strategy will be optimal if it minimizes his expected loss in the long run.

We may expect that in view of the premium reduction, it will be unprofitable to claim damages which are not much larger than $a_0$. Once a damage is claimed, it will be profitable to claim all following damages that exceed $a_0$ during the remaining part of the year. Hence his decisions will also depend on the time in the year.
If i denotes the index of the lastly paid premium $E_i$, $i = 0, 1, 2, 3$ and if $\tau$ denotes the time elapsed since the last payment of premium we should find a function $s(i,\tau)$ with the following property: If at time $\tau$ after the last payment of premium $E_i$ an accident happens with damage s and no damages are claimed during $\tau$ then s should be claimed if $s > s(i,\tau)$. Our task will be to determine the function $s(i,\tau)$, which is most profitable for our motorist. Under the assumptions stated above this optimal strategy can be obtained by means of a method called Markov-programming. A treatment of this method is now on order.

## 2. Markov programming

In the type of problems to which Markovprogramming can be applied we have a physical system. In our case the system is the car and the accident insurance.

At some point in time t, the system is in some state x. The state x will be a point in a finite dimensional Cartesian space, called the state space X. In general the state space X will consist of:

a) The set of transient states. This set has the property that the system will never return to this set once it has left it.

b) One or more simple *) ergodic sets. These have the property that once a state of such a set is reached the system will remain in that set.

The description is not complete. We will not go into the socalled cyclically moving sets, because in all known practical cases this description is sufficient.

The system is subject to random transitions between successive states. It performs a (stochastic) walk through the state space X. If the decisionmaker does not intervene such a walk is a realisation of the natural process. Condition for application of Markovprogramming is that the underlying natural process of the decisionproblem should be a stationnary strong Markovprocess. Such a process is characterized by the following properties: Suppose the system is in state x at time t after starting the process then the probability of being in some set of states A at time $t+\tau$ depends only on A, x and $\tau$, or in formula:

$$P(A, t+\tau; x, t) = P(A, \tau; x).$$

The states, assumed by the system before the present state x is reached, are irrelevant for future transitions when x is completely specified. Further the distribution of the transition probabilities is independent of the absolute time t.

---

*) We will also consider the set of all ergodic states. The word simple ergodic set is used to distinguish it from the set of all ergodic states.

In our motorists problem, the natural process is given by the motorist, who always claims damages. The assumption that the number of accidents in a time interval T is Poisson distributed with constant parameter and the assumed independence between successive damages together imply that the natural process in our case is a stationnary strong Markov process.

In most cases the decisionmaker will try to influence the natural process by interventions, basically a finite number in a finite interval. After an intervention the system is transformed into some other state. Between interventions the system is subject to the natural process.

It is convenient to assume that at every point of time a decision is taken. The decision will be primarily to decide whether to intervene or not and secondly which intervention to choose. In the case the decision is not to intervene we will speak of a null-decision. Once it is decided to intervene in some state x we will have to decide among the many possibilities which transformation is going to be effected. We shall assume that in every state x there exists a set D(x) of possible decisions (transformations) d in the decisionspace D. Once to every state a decision (including null-decision) is attached we have a strategy. Hence, a strategy should specify the set of states where the decisionmaker should intervene and the transformation from the interventionstate to the new state.

The resultant of the natural process and the transformations dictated by the strategy will be called the decision process.

The set of interventionstates plays a prominent role in the decision process. It will be convenient to consider the sequence of interventionstates $I_n$ (n = 1, ...) assumed by the process. For a strategy z the set of interventionstates will be denoted by $A_z$. The sequence $\underline{I}_n$ (n = 1, ...) constitutes a stationnary Markov process with a discrete time parameter. The probability distribution of $\underline{I}_n$, starting in state x will be denoted by

$$p^{(n)}(A; z; x) \quad n = 1, 2, \ldots .$$

It can be proved that the stationnary distribution of the interventionstates exists and is given by

$$p(A;\ z;\ x) = \lim_{n\to\infty} p^{(n)}(A;\ z;\ x).$$

Also one can prove that for two starting states $x_1$ and $x_2$, which are elements of the same ergodic set, the stationnary distributions are identical:

$$p(A;\ z;\ x_1) = p(A;\ z;\ x_2).$$

The optimal strategy has to be chosen according to some criterion. Which criterion has to be chosen? Let us consider a realisation of the system. Such a walk $\omega$ through the state space X may be represented by a point $\omega$ in the space of all possible walks $\Omega$. Applying strategy z during walk $\omega$ we denote the costs in a time period T by $k_T(\omega;\ z)$. If $T \to \infty$, obviously $\lim_{T\to\infty} k_T(\omega;\ z) = \infty$.
Another disadvantage is that the walk $\omega$ is not known in advance. By considering the average costs per timeunit we can overcome these objections. For this criterion one can prove the following theorem:

If $\omega$ denotes a walk of the system, starting in $x_0$, $x_0$ being an ergodic state, then

$$\lim_{T\to\infty} \mathcal{E}\left\{\frac{k_T(\underline{\omega}_i\ z)}{T}\right\}$$

exists and is equal to:

$$\frac{\int p(dI;\ z;\ x_0)\ k(I;\ z)}{\int p(dI;\ z;\ x_0)\ t(I;\ z)}.$$

$k(I;\ z)$ denotes the expected costs and $t(I;\ z)$ the expected length of the time period between two succeeding interventionstates. We will denote the average costs per timeunit by the function $r(z;\ x_0)$, where z is the applied strategy and $x_0$ is the starting state. Then:

$$r(z;\ x_0) = \frac{\int p(dI;\ z;\ x_0)\ k(I;\ z)}{\int p(dI;\ z;\ x_0)\ t(I;\ z)} \quad .$$

For two states $x_1$ and $x_2$ of the same simple ergodic set we have:

$$r(z;\ x_1) = r(z;\ x_2).$$

If $x_0$ is a transient state, the function $r(z;\ x_0)$ is still a random variable. In that case our criterion will be the expected average costs per time unit given by:

$$r(z;\ x_0) = \int p(dy;\ z;\ x_0)\ \frac{\int p(dI;\ z;\ y)\ k(I;\ z)}{\int p(dI;\ z;\ y)\ t(I;\ z)}$$

where $\underline{y}$ obeys the stationnary distribution of ergodic states, corresponding to the starting state $x_0$.

Although the function $r(z;\ x_0)$ could help us to determine which ergodic set to prefer, it does not have the ability to evaluate our preference for certain states within the set.

One can define a function $c(z;\ x)$ with the property that for two states $x_1$ and $x_2$ in one simple ergodic set the difference in total expected costs is given by:

$$c(z;\ x_2) - c(z;\ x_1).$$

If the state space consists of the set of transient states and m disjunct ergodic sets and the states $(e_j;\ j = 1, \ldots, m)$ are arbitrary interventionstates, then for every intervention state, a function $c(z;\ I_1)$ is given by

$$c(z;\ I_1) = k(I_1;\ z) - r(z;\ I_1) \cdot t(I_1;\ z) + \int p^{(1)}(dI_2;\ z;\ I_1)\ c(z;\ I_2)$$

$$c(z;\ e_j) = 0 \qquad\qquad j = 1, 2, \ldots, m.$$

Hence for a system that is considered only in the sequence of interventionstates we compare the total expected costs of the system itself, starting in state I and the case that per time unit an amount $r(z; I)$ has to be paid. The $c(z; I)$ function will also arise in the following physical situation: Suppose a manager has to control a system under the condition that between every pair of successive interventionstates a compensation per time unit $r(z; I)$, depending on the first interventionstate I of the pair, is given him by the owner of the system. To meet the total expected costs of the system the manager will have to pay an additional amount himself equal to $c(z; I)$.

The c-functions are not uniquely determined by the functional equation only. This can be seen by addition of a constant to the c-function and substitution in the functional equation. That is why we have to define the c-function equal to zero for one interventionstate in each ergodic set.

Also for non-intervention states a definition of $c(z; x)$ can be given. To determine the c-function by solving the functional equation we have to determine the functions $k(x; z)$ and $t(x; z)$. These two functions still depend on the strategy z. It is possible by a slightly different definition to make then dependent only on the decision d taken in state x, rather than the complete strategy z.

To this end we consider the interventionset $A_0$. This set $A_0$ consists of the states in which for every possible strategy z, the decisionmaker will intervene. If $A_z$ is the interventionset for an arbitrary strategy z then:

$$A_z \supset A_0.$$

For every state x and decision $d \in D(x)$ we consider two walks denoted by $W^0$ and $W^d$ respectively. During $W^0$ the system will be subject to the natural process until a state of the set $A_0$ is reached. During $W^d$ the decision d transforms the system to state $\underline{u}$, say. From state $\underline{u}$ the walk $W^d$ will be subject to the natural process. It will reach the set $A_0$ via interventionstate $\underline{I} \in A_z$, if z is the applied strategy.

We define the functions $k(x; d)$ and $t(x; d)$ to be the difference in expected costs and expected duration between the walks $W^d$ and $W^0$. For null decisions holds obviously:

$$k(x; d) = 0$$

$$t(x; d) = 0.$$

$W^0$ and $W^d$ are represented in figure 2.1

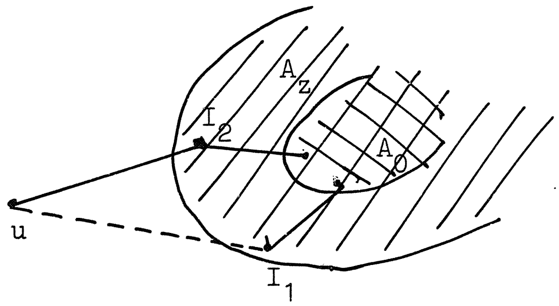


figure 2.1

Consequently, it follows from their definitions that $k(I; z)$ and $t(I; z)$ are identical to the expected cost and duration of the part of the walk $W^d$ between interventionstates $\underline{I}_1$ and $\underline{I}_2$. Both states $\underline{I}_1$ and $\underline{I}_2$ are distributed according to the same limiting distribution $p(A; z; y)$. Hence the expected costs from $\underline{I}_1$ and $\underline{I}_2$ to $A_0$ are equal.

$$\int p(dx; z; y)\, h(x; z) = \int p(dx; z; y)\, k(x; z(x))$$

$$\int p(dx; z; y)\, t(x; z) = \int p(dx; z; y)\, t(x; z(x))$$

where $z(x)$ denotes the decision dictated by strategy z in state I. In the relations for $r(z; x_0)$ and $c(z; x_0)$ the functions $k(x; z)$ and $t(x; z)$ can be replaced by $k(x; z(x))$ and $t(x; z(x))$. If $\underline{I}^*$ denotes the first interventionstate reached by the system, we have for arbitrary strategy z and starting state x:

(2.1) $$r(z;\ x) = \mathcal{E}\{r(z;\ \underline{I}^{*})/x;\ z\}$$

(2.2) $$c(z;\ x) = k(x;\ z(x)) - r(z;\ x)t(x;\ z(x)) + \mathcal{E}\{c(z;\ \underline{I}^{*})/x;\ z\}$$

$$c(z;\ e_j) = 0 \qquad j = 1,\ 2,\ \dots,\ m$$

where $e_j$ are arbitrarily chosen states in the m disjoint ergodic sets.

Finally we will construct an iteration procedure that yields a sequence of strategies $z^i$, i = 1, 2, 3, ... . The properties of this iteration cycle are:

1) every step results in a better strategy or:

$$r(z^i;\ x) \geq r(z^j;\ x) \quad j \geq i$$

2) the sequence converges to the optimal strategy $z_0$ or:

$$\lim_{i \to 0} r(z^i;\ x) = r(z_0;\ x).$$

The proofs are given in [1], but will be omitted here. We will restrict ourselves to some definitions and the iteration cycle.

Suppose we have obtained a strategy z. To improve this strategy we need the following functions. Let in the startingstate x the decision d be taken. If decision d transforms the system into state $\underline{u}$ and after d the strategy z is applied then the functions $r(d \cdot z;\ x)$ and $c(d \cdot z;\ x)$ are defined as follows:

(2.3) $$r(d \cdot z;\ x) \overset{\text{def}}{=} \sum \{r \cdot z;\ \underline{u}/d\}$$

(2.4) $$c(d \cdot z;\ x) \overset{\text{def}}{=} k(x;\ d) - r(d \cdot z;\ x) \cdot t(x;\ d) + \sum \{c(z;\underline{u}/d)\}.$$

Suppose we are in state x and let $\underline{v}$ be the first state in the set A assumed by the system, then we define the functions

[1] G. de Leve: Generalized Markovian Decision Processes, Dissertation, 1964, Mathematical Centre, Amsterdam.

$$r(A \cdot z;\ x) \overset{\text{def}}{=} \sum \{r(z;\ \underline{v})/x;\ A\}$$

$$c(A \cdot z;\ x) \overset{\text{def}}{=} \sum \{c(z;\ \underline{v})/x;\ A\}.$$

$\underline{v}$ obeys the probabilitydistribution of the first state assumed by the system in set A, after starting in state x. Besides we define the class $K_z$ of all closed sets $A \supset A_0$ Satisfying

$$\overline{A} = \{x/r(A \cdot z;\ x) < r(z;\ x)\} \cup \{x/r(A \cdot z;\ x) = r(z;\ x);$$
$$c(A\ z;\ x) \leq c(z;\ x)\}.$$

This means that for all states satisfying this definition it is profitable to delay the intervention according to strategy z until the first state of A is reached.

Finally we define the following subsets:

$$D_z(x) \overset{\text{def}}{=} \{d/d \varepsilon D(x);\ r(d \cdot z;\ x) = \min_{d^* \varepsilon D(x)} r(d^* \cdot z;\ x)$$

and $$A'_z = \bigcap_{A \varepsilon K_z} A.$$

The set $D_z(x)$ contains the decisions d for which the expected average costs per time unit are minimized. The set $A'_z$ is the minimizing subset for strategy z.

The iteration cycle will be as follows:

<u>Preparatory part</u>

Determine the (x; d) functions k(x; d) and t(x; d).

<u>Iterative approach</u>

Let $z^{(n-1)}$ be the strategy obtained at the $(n-1)^{th}$ step of the iteration cycle.

1) Determine the functions $r(z^{(n-1)};\ x)$ and $c(z^{(n-1)};\ x)$ by the relations (2.1) and (2.2).

2) a) Determine the functions $r(d\cdot z^{(n-1)};\ x)$ and $c(d\cdot z^{(n-1)};\ x)$ by using relations (2.3) and (2.4).

b) Determine for every n the subset of minimizing decisions $D_{z^{(n-1)}}(x)$.

c) Minimize the d-function $c(d\cdot z;\ x)$ for every x under the condition $d \varepsilon D_{z^{(n-1)}}(x)$.

3) Determine the functions $r(z_1^{(n-1)};\ x)$ and $c(z_1^{(n-1)};\ x)$ by using relations (2.1) and (2.2).

4) Determine the minimizing subset $A'_{z^{(n-1)}}$ and the new strategy $z^{(n)}(x) \overset{\text{def}}{=} z_1^{(n-1)}(x)$ for $x \varepsilon A'_{z_1^{(n-1)}}$ null decision for $x \notin A'_{z_1^{(n-1)}}$.

The functions $r(z^{(n-1)};\ x)$ and $c(z^{(n-1)};\ x)$ are determined by functional equations. If these equations cannot be solved analytically they often can be solved numerically by Monte Carlo methods.

The way in which the set $A'_{z_1^{(n-1)}}$ can be determined depends heavily on the structure of the decision problem considered. In the boundary points of the minimizing set $A'_z$ it will often be indifferent whether to intervene or not. In the motorists problem, by example, we will see that this property leads to a differential equation for the optimal boundary of the set of states in which claims should be suppressed.

## 3. Application to the motorists problem

Now we shall examine the implications of the framework of Markov programming on the motorists problem and solve it.

Primarily we shall have to define in detail the statespace, the natural process, the interventionstates and the set $A_0$. Secondly we shall determine the k- and t-functions and from these the c-functions. Finally we will show how the optimal strategy is obtained.

### 3.1 Definition of the system

At each point of time the following information will be of interest:

1) whether a damages are covered or not
2) whether an accident happens or not
3) the amount of the last paid premium $E_i$, $i = 0, 1, 2, 3$
4) the point in time considered
5) the extent of the damage
6) whether a damage has been claimed since the last payment of premium.

If we define a time variable it has to specify

a) the amount of the last paid premium
b) the time elapsed since the last paid premium.

We will choose the origin of our time variable at the moment that the premium $E_0$ is paid. If somewhere in a future year a damage is claimed the system will return to the origin at the end of that year. After 4 years of damage-free driving the situation in the fifth year will be the same as in the fourth year, because the same amount of premium is paid. Consequently, taking a year as time unit, the time variable will only assume values in the interval $0 \leq t < 4$.

Because we have assumed that the motorist has effected the insurance, it may be surprising to consider the question whether a damage is covered or not. But in view of the definition of interventionstate, the payment of premium will be an intervention. The inclusion of these interventionstates will result in a convenient choice of the set $A_0$.

By this choice of $A_0$, the computations are considerably reduced. Also we shall have to distinguish between situations in which already a damage is claimed during the same year and the situations in which this is not done. These two facts we will include in our time variable by an enlargement of its definition:

a) $t = i$, $i = 0, 1, 2, 3$, if damages are no longer covered and the premium due amounts $E_i$. (the end of the year just before premium is paid)

b) $t = 1i+\tau$, $i = 0, 1, 2, 3$, $0 \leqq \tau < 1$: the time elapsed since the last paid premium $E_i$ is $\tau$ and one or more damages are claimed during $\tau$

c) $t = 2i+\tau$, $i = 0, 1, 2, 3$, $0 \leqq \tau < 1$: same like b, but no damages claimed during $\tau$.

Summarizing, our time variable is defined in the points $t = 0, 1, 2, 3$ and in the intervals $10 \leqq t < 14$ and $20 \leqq t < 24$.

To specify the state space completely we need two other state variables, u and s. By u we denote the time elapsed since the first claimed damage after the last payment of premium. If no damage is claimed, u will be zero. The state variable s will denote the extent of the last claimed damage. This completes the definition of the state space.

We will describe now the transition mechanism of the process. Two types of transitions are of particular importance. The first type is the transition at the end of the year. Just before paying premium the system will transfer from state ($t = 1i+1$, $s > 0$, $u > 0$), if one or more damages are claimed during the year, to state ($t = 0$, $s = 0$, $u = 0$). By the intervention of paying premium $E_0$ the system will be transformed into state ($t = 20$, $s = 0$, $u = 0$). If no damage is claimed during the year the system will transfer from ($t = 2i+1$, $s = 0$, $u = 0$) to ($t = i+1$, $s = 0$, $u = 0$) if $i < 3$ and ($t = 3$, $s = 0$, $u = 0$) if $i = 3$. By the intervention of paying the relevant premium the system will be transformed respectively to the states ($t = 2i$, $s = 0$, $u = 0$) and ($t = 23$, $s = 0$, $u = 0$). The second type of transition occurs when an accident happens. If one or more damages are claimed already, the system will be in state ($t = 1i+\tau$, $s > 0$, $u > 0$).

The accident at time $\tau$ transfers the system into state ($t = 1i+\tau$, $s = s_1$, $u > 0$) where $s_1$ denotes the extent of the damage. If no damage is claimed since the last payment of premium the system is in state ($t = 2i+\tau$, $s = 0$, $u = 0$). The accident transfers the system into ($t = 1i+\tau$, $s = s_1$, $u = 0$). If the strategy dictates an intervention the system will be transformed to state ($t = 2i+\tau$, $s = 0$, $u = 0$). If no intervention is dictated the system will stay in state ($t = 1i+\tau$, $s = s_1$, $u = 0$). From that moment on u will be positive. Between these types of transitions t will increase linearly and u will remain zero or increase linearly with time.

The natural process of this problem will be such that all accidents will be claimed if the motorist is covered. At the end of the year no premium is going to be paid, so the system remains in state (i,0,0) forever. The physical interpretation of the natural process is given by the motorist, who is a fanatical opponent of insurance. If (by accident!) he has paid premium, he will claim all damages. As soon as the year terminates he will not pay premium anymore.

The interventions of our intelligent motorist will be performed by the payment of premium and the suppression of claims. He will always intervene when premium has to be paid. If no damage is claimed ($u = 0$) before an accident occurs he will intervene by suppressing claims that are smaller than $s(t)$, the lower bound of the extent of the damage that should be claimed according to his strategy. Hence the intervention set $A_z$ of states in which he will intervene according to his strategy z is given by:

1) $10 \leqq t < 14$, $u = 0$, $s \leqq s(t)$
2) $t = 0, 1, 2, 3$, $u = 0$, $s = 0$

The set $A_0$ of states in which he will intervene under all possible strategies is given by:

1) $10 \leqq t < 14$, $u = 0$, $s \leqq a_0$
2) $t = 0, 1, 2, 3$, $u = 0$, $s = 0$

The statespace and the sets $A_0$ and $A_z$ are presented in figure 3.1.

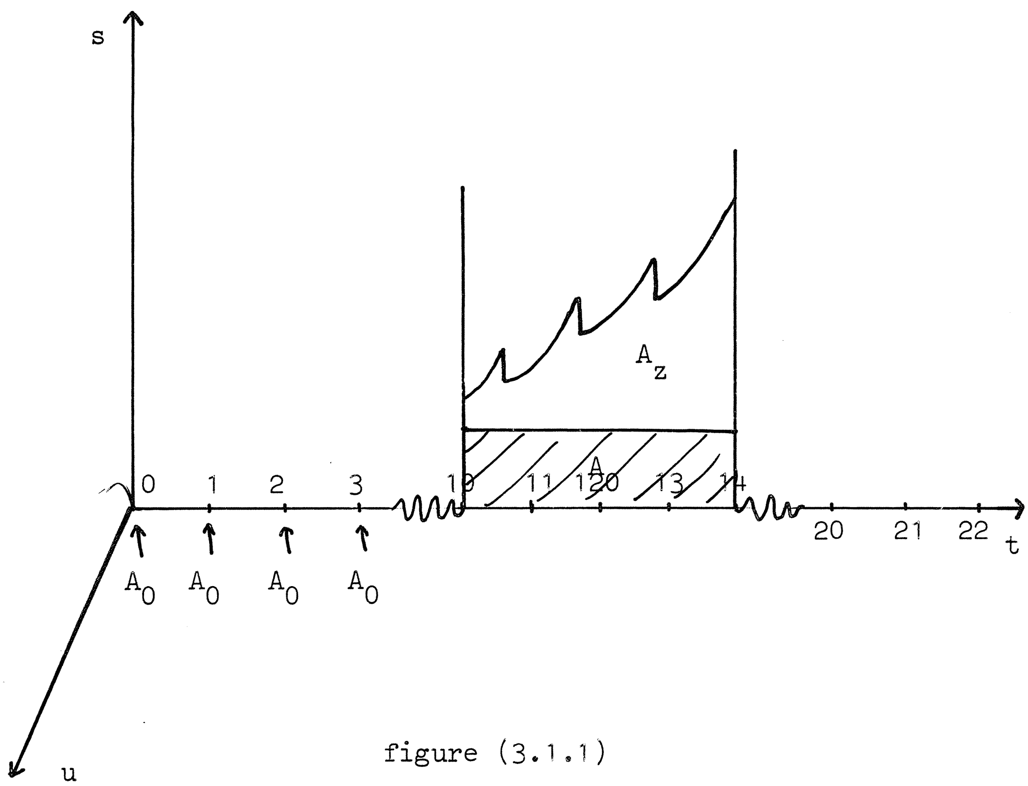


figure (3.1.1)

## 3.2 The determination of the k(x; d) and t(x; d) functions

Primarily we shall consider interventions when an accident occurs at time $\tau$ after the last payment of premium $E_i$. No damage is claimed during the time interval $\tau$. The system will be in state $x = (t = 1i+\tau, s, u = 0)$ and the intervention transforms the system into state $(t = 2i+\tau, s = 0, u = 0)$. Consequently $d = (2i+\tau, 0, 0)$. Consider the walk $W^0$. In this walk the system is subject to the natural process starting in x. If $s \leqq a_0$ then $x \varepsilon A_0$, hence the expected duration as well as the expected costs are zero. If $s > a_0$ the walk reaches the set $A_0$ at the end of the year in state (0, 0, 0). The duration of this walk is $1-\tau$ and the expected costs will be the product of the expected number of accidents $\lambda(1-\tau)$ during the time period $(1-\tau)$ and the expected costs per accident $k(a_0) = \int_0^{a_0} s\, dF(s) + a_0 \int_{a_0}^{\infty} d\,F(s)$.

Consequently, for $x = (t = 1i+\tau,\ s,\ u = 0)$ and $d = (2i+\tau,\ 0,\ 0)$

$$t_0(x;\ d) = \begin{cases} 0 & \text{if } s \leqq a_0 \\ 1-\tau & \text{if } s > a_0 \end{cases}$$

$$k_0(x;\ d) = \begin{cases} 0 & \text{if } s \leqq a_0 \\ \lambda(1-\tau)k(a_0) & \text{if } s > a_0 \end{cases}$$

During the walk $W^d$ the system is subject to the natural process after the decision $d = (2i+\tau,\ 0,\ 0)$. The decision itself involves the decisioncosts $s-a_0$ if $s > a_0$ and zero if $s \leqq a_0$.
The walk terminates as soon as $A_0$ is reached. Suppose the next accident occurs at $\tau+\tau_1$ with damage $s_1$. The system will transfer to state $(1i+\tau+\tau_1,\ s_1,\ 0)$. If $s_1 \leqq a_0$ the walk is terminated at $\tau+\tau_1$. If $s_1 > a_0$ the walk terminates at the end of the year in state (0, 0, 0) by means of a $W^0$ walk starting in $(1i+\tau+\tau_1,\ s_1,\ 0)$. In the case no accident occurs during the remaining period $(1-\tau)$ the walk will terminate in state $(i+1,\ 0,\ 0)$ if $i < 3$ and in $(3,\ 0,\ 0)$ if $i = 3$.
It follows that for $x = (1i+\tau,\ s,\ 0)$ and $d = (2i+\tau,\ 0,\ 0)$

$$t_1(x;\ d) = (1-\tau)e^{-\lambda(1-\tau)} + \int_0^{1-\tau} \lambda\tau_1 e^{-\lambda\tau_1}\, d\tau_1$$

$$+ \{1 - F(a_0)\} \int_0^{1-\tau} \lambda e^{-\lambda\tau_1}(1 - \tau - \tau_1)d\tau_1$$

$$k_1(x;\ d) = k(a_0) \int_0^{1-\tau} \lambda e^{-\lambda\tau_1}\, d\tau_1$$

$$+ \{1 - F(a_0)\}k(a_0) \int_0^{1-\tau} \lambda e^{-\lambda\tau_1}\, \lambda(1-\tau-\tau_1)d\tau_1$$

$$+ \begin{cases} s-a_0 & \text{if } s > a_0 \\ 0 & \text{if } s \leqq a_0 \end{cases}$$

Finally, after some simplifications we have for $x = (1i+\tau, s, 0)$ and $d = (2i+\tau, 0, 0)$:

$$t(x;\, d) = t_1(x{=}\, d) - t_0(x;\, d) =$$

(3.2.1)

$$-F(a_0)\left[(1-\tau) - \frac{1}{\lambda}\left\{1 - e^{-\lambda(1-\tau)}\right\}\right] + \begin{cases}(1-\tau) & s \leqq a_0\\ 0 & s > a_0\end{cases}$$

$$k(x;\, d) = k_1(x;\, d) - k_0(x;\, d) =$$

(3.2.2) $$F(a_0)k(a_0)\left\{1 - e^{-\lambda(1-\tau)}\right\} + \lambda(1-\tau)(1-F(a_0))k(a_0)$$

$$+ \begin{cases} s - a_0 - \lambda(1-\tau)k(a_0) & s > a_0\\ 0 & s \leqq a_0\end{cases}$$

Secondly we consider the interventionstates $x = (i, 0, 0)$ with decision $d = (2i, 0, 0)$. Because $x \varepsilon A_0$ the expected duration and expected costs of the $W^0$ walk are zero. After the decision the system will be subject to the natural process, starting at $\tau = 0$. The decisioncosts are $E_i$. Consequently:

(3.2.3) $$t(x;\, d) = 1 - F(a_0)\left[1 - \frac{1}{\lambda}(1 - e^{-\lambda})\right]$$

(3.2.4) $$k(x;\, d) = E_i + F(a_0)k(a_0)(1 - e^{-\lambda}) + \lambda(1-F(a_0))k(a_0).$$

## 3.3 Determination of the optimal strategy

It is easily verified that for all strategies $z \varepsilon Z$ the Markov process in $A_z$ has only one simple ergodic set. Consequently for every strategy $z_0$ and feasible decision d, we have:

$$r(dz_0;\, x) = r(z_0;\, x) = r(z_0).$$

Hence, we need only to consider the functional equation for $c(z_0;\, x)$:

$$c(z_0;\ x) = k(x;\ z_0(x)) - r(z_0;\ x)t(x;\ z_0(x))$$

(3.3.1)

$$+ \int P^{(1)}_{A_{z_0}} (dI;\ x;\ z_0)c(z_0;\ I).$$

In order to obtain a unique solution we put

$$c(z_0;\ x = (0,\ 0,\ 0)) = 0$$

For $x = (1i+\tau,\ s,\ 0)$ and $d = (2i+\tau,\ 0,\ 0)$ and $s \geqq a_0$, $c(dz_0;\ x)$ is given by

$$c(dz_0;\ x) = (s-a_0) + F(a_0)k(a_0)\{1 - e^{-\lambda(1-\tau)} - \lambda(1-\tau)\}$$

(3.3.2)

$$+ r(z_0) \cdot F(a_0)\left[1 - \tau - \frac{1}{\lambda}(1 - e^{-\lambda(1-\tau)}\right]$$

$$+ c(z_0;\ (2i+\ ,\ 0,\ 0)).$$

Note that $c(dz_0;\ x)$ for $s \geqq a_0$ is a monotone increasing function of s. If z is an arbitrary strategy dictating a null decision in $x = (1i+\tau,\ s,\ 0)$ then the next interventionstate will be $I = (0,\ 0,\ 0)$. Hence,

(3.3.3) $c((z)z_0;\ x) = c((z)z_0;\ (0,\ 0,\ 0)) = c(z_0;\ (0,\ 0,\ 0)) = 0.$

If $z_0$ is an optimal strategy and z is arbitrary, it follows from the optimality property that:

$$c((z)z_0;\ x) \geqq c(z_0;\ x).$$

Consequently, for $x \varepsilon A_{z_0}$ it follows from (3.3.2) that

$$c(z_0;\ x) \leqq c(z_0;\ (0,0,0)) = 0.$$

Let the boundary points of $A_{z_0}$ with $10 \leqq t < 14$ and $s \geqq a_0$ be denoted by the function $s = s(t)$. Since for an intervention d the function $c(dz_0;\ x)$ is a linear monotone increasing function of s, we have for the boundary points x of $A_{z_0}$

(3.3.4) $c(z_0;\ x) = c(z_0;\ (0,\ 0,\ 0)) = 0.$

Now it follows from (3.3.2) and (3.3.4) that for $a_0 \leqq s \leqq s(t)$ and for $10 \leqq t < 14$

$$c(z_0;\ (t,s,0)) = s - s(t).$$

Also from (3.3.2) and (3.3.4) we have

$$\begin{aligned} c(z_0;\ (2i+\tau,\ 0,\ 0)) = a_0 - s(t) - \\ - F(a_0)k(a_0)\left[1 - e^{-\lambda(1-\tau)} - \lambda(1-\tau)\right] \\ - r(z_0)F(a_0)\left[(1-\tau) - \frac{1}{\lambda}\left(1 - e^{-\lambda(1-\tau)}\right)\right] \end{aligned} \tag{3.3.5}$$

If $x = (1i+\tau,\ s,\ 0)$ and $d = (2i+\tau,\ 0,\ 0)$ then for $s < a_0$ we have

$$\begin{aligned} c(dz_0;\ x) &= k(x;\ d) - r(z_0)t(x;\ d) + c(z_0;\ (2i+\tau,0,0)) \\ &= F(a_0)k(a_0)(1 - e^{-\lambda(1-\tau)}) + (1-F(a_0))k(a_0)\lambda(1-\tau) \\ &\quad -r(z_0)\left[(1-\tau) - F(a_0)\left[(1-\tau) - \frac{1}{\lambda}\left(1 - e^{-\lambda(1-\tau)}\right)\right]\right] \\ &\quad + c(z_0;\ (2i+\tau,0,0)). \end{aligned} \tag{3.3.6}$$

For $x \varepsilon A_{z_0}$ : $c(dz_0;\ x) = c(z_0;\ x)$, hence by (3.3.5) and (3.3.6)

$$c(z_0;\ x) = a_0 - s(t) + k(a_0)\lambda(1-\tau) - r(z_0)(1-\tau).$$

Further we have for $i = 1,\ 2,\ 3$:

$$\begin{aligned} c(z_0;\ (i,0,0)) &= \lim_{\tau \uparrow 1} c(z_0;\ (0,\ 2i+\tau,\ 0)) \\ &= a_0 - \lim_{t \uparrow 1i+1} s(t). \end{aligned}$$

Summarizing we have

(3.3.7) $c(z_0;\ x) =$

$\lceil$ 0; for $x = (0,\ 0,\ 0)$ or if $x$ satisfies $s \geqq s(t)$

$$
\left|\begin{array}{l}
a_0 - \lim\limits_{t\uparrow 1i+1} s(t); \text{ for } x = (i, 0, 0) \text{ with } i = 1, 2, 3 \\[2ex]
a_0 - s(t) + k(a_0)\lambda(1-\tau) - r(z_0)(1-\tau); \text{ for } s \leqq a_0, \ t = 1i+\tau \\[2ex]
s - s(t); \text{ for } a_0 \leqq s \leqq s(t) \text{ and } t = 1i+\tau \\[2ex]
a_0 - s(t-10) - F(a_0)k(a_0)\{1 - e^{-\lambda(1-\tau)} - \lambda(1-\tau)\} \\[1ex]
\quad - r(z_0)F(a_0)\{(1-\tau) - \frac{1}{\lambda}(1 - e^{-\lambda(1-\tau)}); \text{ for } x = (t,0,0), \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t = 2i+\tau
\end{array}\right.
$$

From the functional equation (3.3.1) for $c(z_0; (i,0,0))$ we have

$$
\begin{aligned}
c(z_0; (i,0,0)) &= k((i,0,0); (2i,0,0)) - r(z_0)t((i,0,0); (2i,0,0)) \\
&\quad + c(z_0; (2i,0,0)).
\end{aligned} \tag{3.3.8}
$$

By (3.2.3), (3.2.4) and (3.3.8) we have:

$$
\begin{aligned}
&c(z_0; i,0,0) - c(z_0; (2i,0,0)) = \\
&E_i + F(a_0)k(a_0)(1 - e^{-\lambda}) + \lambda(1 - F(a_0))k(a_0) \\
&- r(z_0)\{1 - F(a_0)\left[1 - \frac{1}{\lambda}(1 - e^{-\lambda})\right]\}.
\end{aligned} \tag{3.3.9}
$$

By using the relation (3.3.7) for $c(z_0; (0,0,0))$ and $c(z_0; (20,0,0))$ we find by substitution in (3.3.9)

$$
s(10) = E_1 + \lambda k(a_0) - r(z_0) + a_0. \tag{3.3.10}
$$

In the same way for $i = 1, 2, 3$:

$$
s(1i) = \lim_{t\uparrow 1i} s(t) + E_i + \lambda k(a_0) - r(z_0). \tag{3.3.11}
$$

Further we have

$$
\lim_{t\uparrow 13} s(t) = \lim_{t\uparrow 14} s(t). \tag{3.3.12}
$$

For $x = (t,s,0)$ with $s \geqq a_0$ and $t = 1i+\tau$, it follows from the functional equation for $c(z_0;\, x)$:

$$c(z_0;\ (1i+\tau,s,0)) = k((1i+\tau,s,0);\ (2i+\tau,0,0))$$

$$- r(z_0) \cdot t((1i+\tau,s,0);\ (2i+\tau,0,0))$$

$$+ \int_{1-\tau}^{\infty} \lambda e^{-\lambda\tau_1} d\tau_1 \cdot c(z_0;\ (i+1,0,0))$$

$$+ \int_0^{1-\tau} \lambda e^{-\lambda\tau_1} d\tau_1 \int_0^{s(t+\tau_1)} c(z_0;\ (t+\tau_1;\ y,0)dF(y)$$

$$+ \int_0^{1-\tau} \lambda e^{-\lambda\tau_1} d\tau_1 \int_{s(t+\tau_1)}^{\infty} c(z_0;\ (0,0,0))dF(y).$$

By (3.2.1), (3.2.2) and (3.3.7) it follows

$$c(z_0;\ (1i+\tau,s,0)) = s - a_0 + e^{-\lambda(1-\tau)}\cdot(a_0 - \lim_{t\uparrow 1i+1} s(t))$$

$$+ \int_0^{1-\tau} \lambda e^{-\lambda\tau_1} d\tau_1 \int_{a_0}^{s(t+\tau_1)} (y-s(t+\tau_1))dF(y) \tag{3.3.13}$$

$$+ \int_0^{1-\tau} \lambda e^{-\lambda\tau_1} \cdot F(a_0) \cdot (a_0-s(t+\tau_1))d\tau_1$$

If we choose $s = s(t)$ then $c(z_0;\ (1i+\tau,s,0)) = 0$. After substitution of this condition in equation (3.3.13) and differentiation with respect to t, we arrive at the following functional equation:

$$\frac{ds(t)}{dt} = \lambda \int_{a_0}^{\infty} (y-a_0)dF(y) - \lambda \int_{s(t)}^{\infty} (y-s(t))dF(y). \tag{3.3.14}$$

The boundary of the intervention set $A_{z_0}$ of the optimal strategy obeys this functional equation. Together with the relations (3.3.10), (3.3.11) and (3.3.12) it determines the solution of the motorists problem completely.

To demonstrate this, the solution for the case that the extent of the damage is exponentially distributed is presented in figure (3.3.1).

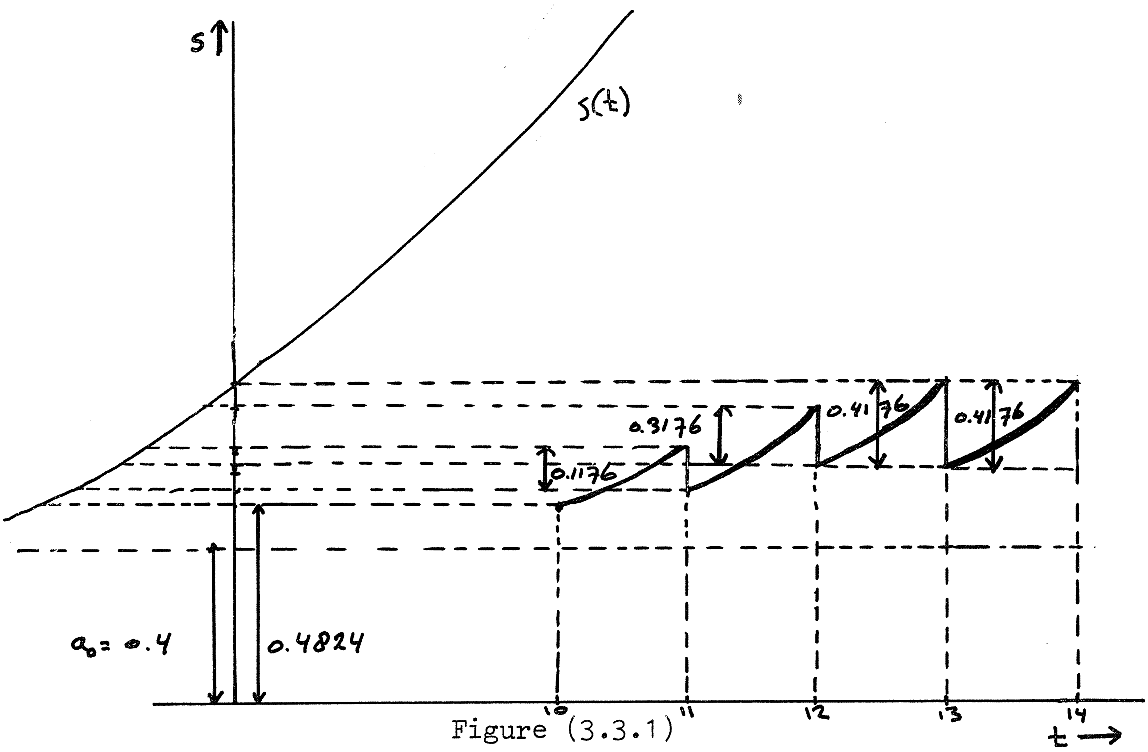


Figure (3.3.1)

The numerical data are:

$a_0 = 0.4$

$\lambda = 2$

$F(s) = 1-e^{-s}$

$E_1 = 1.6 \quad E_3 = 1.2$

$E_2 = 1.4 \quad E_4 = 1.1$

In the exponential case the solution of (3.3.14) can be obtained analytically and is given by:

$$s(t) = 0.4 + \ln\{1 + e^{2e^{-0.4}(t+c_i)}\}, \quad 1i \leq t < 1i+1.$$

The boundaries of the intervention set $A_{z_0}$ for the different years are given by translations in the time variable of this curve which are determined by the constants $c_i$.
These constants as well as $r(z_0)$ can be obtained from the relations (3.3.10), (3.3.11) and (3.3.12).