stichting
mathematisch
centrum

AFDELING MATHEMATISCHE BESLISKUNDE BD 2/73 OCTOBER

G. DE LEVE and H.C. TIJMS

A GENERAL MARKOV PROGRAMMING METHOD, WITH APPLICATIONS

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

<u>Contents</u>

1. Model and method

1.1. Introduction

The well known models of Howard [2] and Jewell [3] assume a finite state and action space. Also, these models assume that the costs incurred during the time interval between two successive epochs at which the decision maker intervenes depend only on the state at the beginning of that interval and the action taken in that state. A similar assumption is made for the transition times and transition probabilities.

In many decision problems these assumptions are not satisfied (cf. the example below). In this paper we shall treat techniques that may attack these problems, where we assume an infinite planning horizon and the long-run average cost as criterion. These techniques are not "ready-made" techniques and their final form depends heavily on the structure of the decision problem considered.

In the next section we shall carry point by point the elements of our model. In this presentation we shall not pursue mathematical rigourousness everywhere. The various concepts will be illustrated on the basis of the next example.

Production problem*)

The production of a continuous product can be realised on a finite number of production levels ℓ_i, $i = 0, 1, \ldots, N$ with $\ell_0 = 0$. The product is kept in stock. The storage capacity is limited to a quantity M.

Orders arrive according to a Poisson-process with a mean of λ per unit of time. The order size $\underline{y}$ is a non-negative random variable with a given distribution function F(y) with finite mean and variance. The order size is assumed to be independent of the arrival process. Orders are fulfilled immediately by the available stock. If the size of an order exceeds the available stock then the supply is replenished by an emergency purchase.

The production can be controlled by switching over to another production level. There is no lead time needed to perform a change of production level. The following costs are involved in the operation of this system:

*) This problem has been solved by P.J. Weeda in Chapter 6 of [5].

a) production costs $c_p(i)$ per unit of time for level ℓ_i, $i = 0, 1, \ldots$ $\ldots, N$ with $c_p(0) = 0$.

b) costs $c_q(i,j)$ of switching over from level ℓ_i to level ℓ_j, $i, j = 0, \ldots, N$.

c) costs c_r per unit of product of an emergency purchase.

d) stockholding costs c_s per unit of time per unit of product.

1.2. The elements of the model

(A) *In our model the state of the system is described by a point in a finite dimensional Cartesian space* X. *The state space* X *is chosen in such a way that at each point of time the state of the system can be described by a point of* X.

(B) *We define some basic process that describes for each initial state the evolution of the system when the decision maker takes no interventions. This process is called the natural process and is assumed to be a strong Markov process.*

(C) *For each state* $x \in X$, *there is a set* $D(x)$ *of feasible decisions in state* x, *where* $D(x)$ *is a closed subset of a finite dimensional Cartesian space. We distinguish between null-decisions and interventions. A null-decision does not interrupt the natural process. An intervention is a decision which interrupts the natural process and causes an instantaneous (possibly random) change of the state of the system (since the state of the system is specified at each point of time, it is no restriction to assume that the change of the state takes no time).*

When an intervention has a deterministic effect we shall often identify the intervention with the state which results from that intervention. The elements (A) – (C) must be chosen in such a way that the following property holds.

(D) *There is a non-empty closed set* A_0*) *consisting of states in which the null-decision is not feasible, i.e. in each state of* A_0 *the decision-maker has to intervene always. Any feasible intervention in a state of*

---

*) In this paper any set $A \subseteq X$ is assumed to be a Borelset.

A_0 *has to result into a state that does not belong to* A_0 *with probability* 1. *Also, there must hold that from each initial state* $x \in X$ *the natural process reaches the set* A_0 *within a finite time with probability* 1.

Before we proceed, we illustrate the points (A) - (D).

Productionexample. We specify the state of the system by two state variables, an integer i for the production level and a real variable s for the stocklevel, where $i = 0,\ldots, N$ and $0 \le s \le M$. So the state space is given by

$$X = \{x = (i,s) \mid i = 0,1,\ldots, N \ ; \ 0 \le s \le M\}.$$

We define the natural process as follows. When the initial state is (0,s) with $0 < s \le M$, the system remains in this state until the next epoch at which an order arrives. At that epoch the system assumes state $(0,\max(0,s - \underline{v}))$ where $\underline{v}$ is the size of that order. Once the natural process is in state (0,0), the natural process remains in that state forever. When the initial state is (i,s) with $i > 0$ and $s < M$, the system continues to produce on the productionlevel i, where a decrease of the stock level occurs when an order arrives (an emergence purchase is done, when the size of the order exceeds the available stock). The natural process remains in state (i,M) forever as soon as this state is taken on.

A decision in this problem has a deterministic effect. We identify any decision with the state into which the decision results. We forbid the null-decision in the states (0,0) and (i,M) $(1 \le i \le N)$. In state (i,M) we only permit the intervention (0,M). Hence

$$D(0,0) = \{(0,j) \mid 1 \le j \le N\}; \ D(i,M) = \{(0,M) \mid \text{ for } 1 \le i \le N\};$$
$$D(i,s) = \{(j,s) \mid 0 \le j \le N\}, \text{ otherwise}.$$

Clearly, the set $A_0 = \{(0,0)\} \cup \{(i,M) \mid 1 \le i \le N\}$ satisfies point (D).

(E) *We will consider only the class* Z *of stationary strategies. A strategy* $z \in Z$ *adds to each state* x *a decision* $z(x) \in D(x)$. *Let* A_z *be the set of states in which strategy* z *dictates an intervention. It is required that* A_z *is a closed set and that for any intervention the resulting state does not belong to* A_z *with probability* 1. *)

In fig. 1 we present for the productionproblem with N = 1 a particular strategy

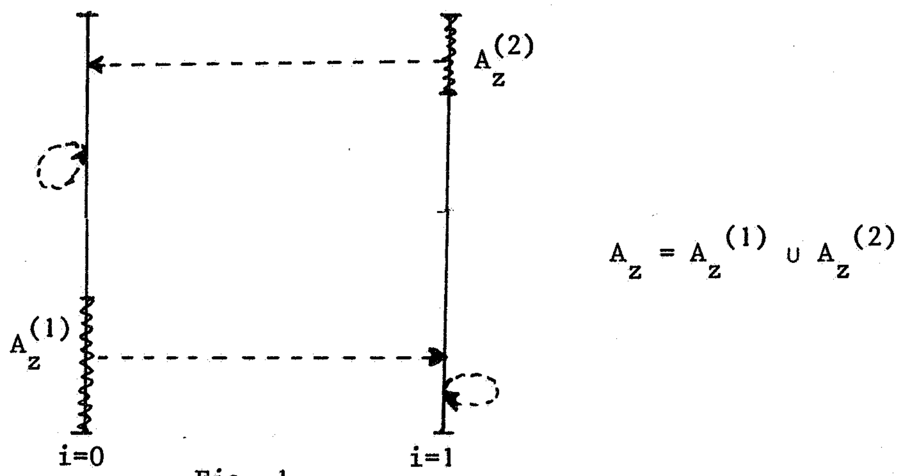

Fig. 1

(F) *When the natural process is controlled by a strategy* $z \in Z$, *then the resulting process is called the decision process. It is required that for any decision process the number of interventions in a finite time is finite with probability* 1.

<u>Remark 1.</u> In the decision process corresponding to strategy z an intervention occurs when the system assumes a state of A_z. By $A_z \supseteq A_0$ and point (D), any intervention transfers instantaneously the system into a state which does not belong to A_0 with probability 1. Hence the decision process is independent of the definition of the natural process on A_0. In other words, we have some freedom in defining the natural process. However, the result of the natural process and a strategy must agree with "reality". This remark may be useful for the determination of the functions k(x;d) and t(x;d) which will be introduced in point (H).

---

*) This latter requirement may be weakened somewhat.

(G) *For any strategy* $z \in Z$ *and any* $z \in X$, *let* $\underline{I}_n$ *be the nth future intervention state* *) *when strategy* z *is used and the initial state is* x *(observe* $\underline{I}_1$ *need not be* x *when* $x \in A_z$). *The process* $\{\underline{I}_n, n \geq 1\}$ *is called the decision process in* A_z.

It can be shown under general conditions that $\{\underline{I}_n\}$ is a Markov process with discrete time parameter.

(H) *Choose now two empty subsets*

$$A_{01} \subseteq A_0 \text{ and } A_{02} \subseteq A_0$$

such that in the natural process with probability 1 *each of these subsets is reached from each initial state within a finite time. We now associate to each state* $x \in X$ *and decision* $d \in D(x)$ *random walks* $\underline{w}_{0i}$ *and* $\underline{w}_{di}$, $i = 1,2$. *The walk* $\underline{w}_{01}$ (resp. $\underline{w}_{02}$) *has* x *as initial state and during this walk the system is subjected to the natural process. The walk* $\underline{w}_{01}$ (resp. $\underline{w}_{02}$) *ends as soon as the system assumes a state of* A_{01} (resp. A_{02}). *The walk* $\underline{w}_{d1}$ (resp. $\underline{w}_{d2}$) *has* x *as initial state too. But now in state* x *decision* d *is made, by which the system is transferred (instantaneously!) into some state and from that state on the system is subjected to the natural process. The walk* $\underline{w}_{d1}$ (resp. $\underline{w}_{d2}$) *ends as soon as the system assumes a state of* A_{01} (resp. A_{02}).

Let k_0 (x) *and* k_1 (x;d) *be the expected costs* **) *incurred during* $\underline{w}_{01}$ *and* $\underline{w}_{d1}$, *respectively. For the determination of the costs the walks are assumed to be "left open and right closed". However, the costs of the decision* d *are included in* $k_1(x;d)$. *Let* t_0 (x) *and* t_1 (x;d) *be the expected duration of* $\underline{w}_{02}$ *and* $\underline{w}_{02}$, *respectively. Now, for* $x \in X$ *and* $d \in D(x)$, *define*

$$(1) \qquad k(x;d) = k_1(x;d) - k_0(x) \text{ and } t(x;d) = t_1(x;d) - t_0(x).$$

---

*) In this paper we underline random variables.

**) It is supposed that there is a given cost structure on the natural process and that there are associated direct decision costs with each intervention. For convenience the costs are assumed to be non-negative.

Observe that $k(x;d) = t(x;d) = 0$ if d = null-decision. By a proper choice of A_{01} and A_{02} the determination of $k(x;d)$ and $t(x;d)$ may become very simple. Also, observe that the functions $k(x;d)$ and $t(x;d)$ not depend on any particular strategy. Hence we need only once and for all to determine these functions.

Finally, it is assumed that the functions $k(x;d)$ and $t(x;d)$ are bounded.

1.3. The criterion function and the functional equations

In point (G) we have defined for a strategy $z \in Z$ the imbedded process $\{\underline{I}_n\}$. This process will play an important role in our considerations. For ease of presentation we assume throughout this paper that for each strategy $z \in Z$ the associated Markov process $\{\underline{I}_n\}$ has no two disjoint ergodic sets.

For each $z \in Z$, $x \in X$ and any set $A \subseteq X$, let

$p^{(n)}(A;z;x)$ = probability that $\underline{I}_n$ belongs to the set A when the strategy z is used and x is the initial state.

We make the next assumption.

Assumption. For each strategy $z \in Z$ the associated Markov process $\{\underline{I}_n\}$ satisfies the Doeblin condition.

We refer to Doob [1] for the definition of the Doeblin condition.

Remark 2. A useful criterion can be given to verify the Doeblin condition. Let $z \in Z$. In case there is a finite set $E \subseteq X$, an integer $\upsilon \geq 1$ and a positive number ρ such that

$$p^{(\upsilon)}(E;x;z) \geq \rho \qquad \text{for all } x \in A_z, \tag{2}$$

then for strategy z the process $\{\underline{I}_n\}$ satisfies the Doeblin condition.

In the problems we encountered the condition (2) appeared to be satisfied.

Now, by the theory of Markov processes (cf. Doob), there is a probability measure $q(A;z)$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p^{(k)}(A;z;x) = q(A;z) \qquad \text{for all } x\epsilon X.$$

Further, for any $A \subseteq X$,

(3) $$q(A;z) = \int_{A_z} p^{(1)}(A;z;x)\, q(dx;z).$$

THEOREM 1. For each strategy $z\epsilon Z$, the long-run average cost is independent of the initial state and equals with probability 1

(4) $$\int_{A_z} k(x;z(x))q(dx;z) \Big/ \int_{A_z} t(x;z(x))q(dx;z).$$

Proof. We will briefly sketch the proof. Suppose that strategy $z\epsilon Z$ is applied and that at epoch 0 the system is in state $x\epsilon X$. Let $T(x;z)$ be the expected duration of the time until the next epoch at which the system reaches a state of A_z. Let $K(x;z)$ be the expected costs incurred during this time interval, where we only count any decision cost incurred at epoch 0 and where we consider this interval left open and right closed for the other costs. Then, by using standard ergodic theorems, it can be shown (see De Leve [4])that the long-run average cost equals with probability 1

(5) $$\int_{A_z} K(x;z)q(dx;z) \Big/ \int_{A_z} T(x;z)q(dx;z).$$

Now, by the definitions of $K(x;z)$ and $k(x;z(x))$,

(6) $$k(x;z(x)) = K(x;z) + \int_{A_z} k_0(I)p^{(1)}(dI;z;x) - k_0(x), \quad x\epsilon X.$$

By an interchange of the order of integration and (3),

(7) $$\int_{A_z} q(dx;z) \int_{A_z} k_0(I)p^{(1)}(dI;z;x) = \int_{A_z} k_0(I)q(dI;z)$$

From (6) and (7),

(8) $$\int_{A_z} k(x;z(x))q(dx;z) = \int_{A_z} K(x;z)q(dx;z).$$

Similarly, $\int_{A_z} t(x;z(x))q(dx;z) = \int_{A_z} T(x;z)q(dx;z)$ From this,

(5) and (8) the theorem now follows

The average cost can also be found by solving a set of functional equations. This set of functional equations will play a basic role in our considerations.

THEOREM 2. Let $z \in Z$. Then

(a) the set of equations

(9) $$c(z;x) = k(x;z(x)) - r(z)t(x;z(x)) + Ec(z;\underline{I}_1), \qquad x \in X$$

has a solution.

(b) For any solution of (9) holds that $r(z)$ equals the expression given in (4), i.e., $r(z)$ represents the long-run average cost.

(c) Let y be an arbitrary state in X. If the condition $c(z;y) = 0$ is added to (9), then the resulting set of equations has a unique solution.

Proof. We refer to De Leve [4] for a proof of (a). The proofs of (b) and (c) are standard (cf. Howard [2]) and are omitted.

Remark 3. Since $k(x;z(x)) = t(x;z(x)) = 0$ for $x \notin A_z$, we have

(10) $$c(z;x) = Ec(z;\underline{I}_1) \qquad \text{for } x \notin A_z.$$

Hence, in fact it is sufficient to solve (9) for $x \in A_z$.

Definition 1. A strategy z^* is called optimal if $r(z^*) \le r(z)$ for all $z \in Z$.

1.4. Basic tools

In this section we shall give a number of results which justify the solution techniques that will be treated in the final section.

Let us fix a strategy $z_1 \epsilon Z$ and a solution $(r(z_1), c(z_1;x))$ of the set of functional equations (9) with $z = z_1$. Our first goal is to improve strategy z_1. We need the following definitions.

Definition 2. For any $x \epsilon X$ and $d \epsilon D(x)$,

(11) $$c(d.z_1;x) = k(x;d) - r(z_1)\ t(x;d) + Ec(z_1;\underline{u})$$

where $\underline{u}$ is the state in which the system is transferred by decision d in state x.

It is not difficult to verify that

(12) $c(z_1(x).z_1;x) = c(z_1;x)$ and $c(d.z_1;x) = c(z_1;x)$ for d = null-decision.

The proof of the first part of (12) is a simple exercise in the use of conditional expectations. The other part of (12) follows simply from $k(x;d) = t(x;d) = 0$ for d = null-decision.

Definition 3. For any $z \epsilon Z$,

(13) $$c([z]z_1;x) = \begin{cases} c(z(x).z_1;x) & \text{for } x \epsilon A_z , \\ Ec([z]z_1;\underline{I}) & \text{for } x \notin A_z , \end{cases}$$

where $\underline{I}$ is the first state the system takes on in the set A_z when the system is subjected to the natural process and x is the initial state.

The following theorem is basic (for its proof, see De Leve [4]).

THEOREM 3. Let z be a strategy such that

$$c([z]z_1;x) \leq c(z_1;x) \qquad \text{for all } x \epsilon X ,$$

then $r(z) \leq r(z_1)$. When the inequality signs are reversed, the theorem remains true.

<u>COROLLARY.</u> Let z^* be a strategy such that

$$c(z^*;x) = \min_{z \in Z} c([z]z^*;x) \qquad \text{for all } x \in X$$

then strategy z^* is optimal.

If we start with a strategy z_1 and we want to find a better strategy, then this corollary suggests to look for a strategy z_2 satisfying

$$(14) \qquad c([z_2]z_1;x) = \min_{z \in Z} c([z]z_1;x) \qquad \text{for all } x \in X.$$

In case z_2 appears to be identical to z_1 we have found an optimal strategy. The suggestion above will be performed in two steps. First we improve strategy z_1 by the Policy Improvement Operation.

<u>Policy Improvement Operation.</u> For each $x \in X$, determine

$$(15) \qquad c^*(z_1;x) = \min_{d \in D(x)} c(d.z_1;x).$$

Strategy z_1' is constructed by adding to each state x a minimizing decision, where we agree that we choose $z_1'(x) = z_1(x)$ when $z_1(x)$ is a minimizing decision.

By the above agreement and (12), we have

$$(16) \qquad A_{z_1'} \supseteq A_{z_1}$$

<u>THEOREM 4.</u> $r(z_1') \leq r(z_1)$.

<u>Proof.</u> By Theorem 3 it suffices to show that $c([z_1']z_1;x) \leq c(z_1;x)$ for all x. Since $c^*(z_1;x) = c(z_1'(x) \cdot z_1;x)$, it follows from (13) that

$$c([z_1']z_1;x) \leq c(z_1;x) \qquad \text{for all } x \in A_{z_1'}.$$

Hence, by (13), for $x \notin A_{z_1}'$

$$c([z_1']z_1;x) = E\, c([z_1']z_1;\underline{I}) \leq E\, c(z_1;\underline{I}) = c(z_1;x) \ ,$$

where the last equality follows from (10) and (16) by conditioning on the first entry in the set A_{z_1}'. This ends the proof.

By (16) it will be clear that we need a mechanism which may reduce ppp the set A_{z_1}' . We give the following definition

Definition 4. For any closed set A with $A \supseteq A_0$,

$$(17) \qquad c(A.[z_1']z_1;x) = Ec^*(\underline{z}_1,\underline{a}), \qquad x \in X,$$

where $\underline{a}$ is the first state in A taken on by the system when the initial state is x and the system is subjected to the natural process.

Observe $c(A.[z_1']z_1;x) = c^*(z_1;x)$ for $x \in A$. We now introduce:

Cutting mechanism (optimal stopping) Let F be the class of all closed sets A satisfying

$$(18) \qquad A_0 \subseteq A \subseteq A_{z_1}' \quad \text{and } c(A.[z_1']z_1;x) \leq c^*(z_1;x) \text{ for all } x \in A_{z_1}'.$$

It can be shown that $A \cap B \in F$ when $A, B \in F$. Define now

$$A_{z_1}'' = \bigcap_{A \in F} A,$$

and suppose that this intersection set belongs to F. Define strategy z_2 as follows

$$z_2(x) = \begin{cases} z_1'(x) & \text{for } x \in A_{z_1}'' \\ \text{null-decision,} & \text{otherwise.} \end{cases}$$

THEOREM 5. The strategy z_2 satisfies the relation (14) *).

---

*) By Theorem 3, $r(z_2) \leq r(z_1)$. More general, let A be a set satisfying (18), and let $z_A(x) = z_1'(x)$ for $x \in A$ and $z_A(x)$=null-decision, otherwise. Then $r(z_A) \leq r(z_1)$, by $c(A.[z_1']z_1;x) = c([z_A]z_1;x)$ for all x, (18) and Theorem 3.

We refer to De Leve [4] for the proof of this theorem.

Remark 4. The determination of the set A'_{z_1} can be seen as an optimal stopping problem. Let us consider the natural process.
Suppose that the natural process may be stopped in each state of the set A_{z_1}, where we have to stop the natural process as soon as it takes on a state in A_0. The natural process may not be stopped outside A_{z_1}. When the natural process is stopped in state x a cost $c^*(z_1;x)$ is incurred. Under general conditions we have that the set A'_{z_1} is equal to the smallest optimal stopping set (cf.[6]).

1.5. The solution techniques

First we specify the elements (A) - (D) and we determine the (x;d) - functions k(x;d) and t(x;d).

I DIRECT APPROACH. Determine $z^* \in Z$ such that

$$c(z^*;x) = \min_{z \in Z} c([z]z^*;x) \qquad \text{for all } x \in X .$$

Then, by the Corollary of Theorem 3, strategy z^* is optimal. However in most cases an optimal strategy can be determined only in an iterative way.

II ITERATIVE APPROACH

Let $z = z^{(n-1)}$ be the strategy obtained at the end of the (n-1)th step of the iteration procedure (start in step 1 with an arbitrary strategy of Z). The nth step runs as follows

(a) Functional equations. Choose a state $y \in X$. Determine the unique solution $(r(z), c(z;x))$ of (cf. (9) and (10))

$$\begin{aligned} c(z;x) &= k(x;z(x)) - r(z)t(x;z(x)) + Ec(z;\underline{I}_1), \qquad x \in X \\ c(z;y) &= 0. \end{aligned}$$

(b) Policy Improvement Operation. For each $x \in X$, determine (cf.(11))

$$c^*(z;x) = \min_{d \in D(x)} \{k(x;d) - r(z)t(x;d) + Ec(z;\underline{u})\}.$$

Construct strategy z' by adding to each state x a minimizing decision, where we choose $z'(x) = z(x)$ when $z(x)$ is a minimizing decision.

(c) Cutting mechanism (optimal stopping) Determine the smallest set A satisfying (18). Denote this set by $A'_{z'}$, and define strategy $z^{(n)}$ by

$$z^{(n)}(x) = \begin{cases} z'(x) & \text{for } x \in A'_{z'}, \\ \text{null-decision}, & \text{otherwise}. \end{cases}$$

End of the nth step.

It can be shown (see [4]) that

$$\lim_{n \to \infty} r(z^{(n)}) = \inf_{z \in Z} r(z).$$

In [6] the theory given in [4] has been specialized to the case of a finite number of states and a finite number of decisions. For this case the iterative method converges after a finite number of steps (see [6]).

2. The motorist problem (Chapter 5 in [5]).

5.1. Problemformulation

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 have been claimed during i successive years i = 1, 2 or 3 the premium is reduced to E_i. After four years of damagefree driving no further premium reduction is granted, so the premium remains E_3. The premium is due on the first day of the premium year.*) The own risk amounts a_0.

The number of accidents is assumed to be Poisson-distributed with a mean of λ per year. It is assumed that the damages caused by the accidents are mutually independent random variables, which have a common distribution function F(s) with finite mean and variance. Furthermore the damages are assumed to be independent of the Poisson-process, which generates the accidents.

The problem of the motorist will be to decide whether to claim a damage or not. The solution of the problem will be a strategy that specifies his decisions in every possible situation. This strategy will be optimal if it minimizes the expected average costs per year in the long run.

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 damages that exceed a_0 during the remaining part of the year. Hence his decisions will also depend on the time of the year and the premium paid at the beginning of that year. So we distinguish between four types of year, for each premium one.

Our task will be to determine for each premium year a function s(t) with the following property: If at time t an accident occurs with damage s and no damages have been claimed since the last payment of premium, then s

*) It is no restriction to assume this is January 1.

should be claimed if $s > s(t)$. The strategy is completely fixed by this function. The optimal strategy will be the function $s(t)$ that minimizes the (expected) average costs per year in the long run.

The solution of this problem by our method will start with the application of the strategy-independent notions in section 5.2. In this section the state space, the natural process, the feasible decisions, the set A_0 and the functions $k(x;d)$ and $t(x;d)$ will be determined. In section 5.3 the functional equation (9) is specified to the situation met in this problem, after which the optimal strategy is determined using the direct approach given by (19). Finally in section 5.4 some numerical results will be given.

5.2. The strategy-independent notions

In order to define the state space in this problem the relevant information at each point of time is considered. The following information will be of interest:

(1) whether an eventual damage is 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 date and time of the day considered;

(5) the extend of the damage;

(6) whether a damage has been claimed since the last payment of premium or not.

In figure 5.1. the state space is presented.

At the t-axis we distinguish:

a) Four points: E_i, $i = 0, 1, 2, 3$. In these states the corresponding premium has to be paid; damages are no longer covered by insurance.

b) Four intervals of one year *) : $1i \leq t < 1i + 1$, $i = 1, 2, 3, 4$. The t-component of the state runs through $1i \leq t < 1i + 1$ if and only if

*) $1i = 11, 12, 13, 14$ if $i = 1, 2, 3$ and 4 respectively.

the last premium paid was E_{i-1}, one or more damages have been claimed that year and coming damages are still covered by insurance.

c) Four intervals of one year: $2i \leq t < 2i + 1$, $i = 1, 2, 3, 4$. The t-component of the state runs through $2i \leq t < 2i + 1$ if and only if the last premium paid was E_{i-1}, no damages have been claimed up to t since the last payment of premium and coming damages are still covered by insurance.

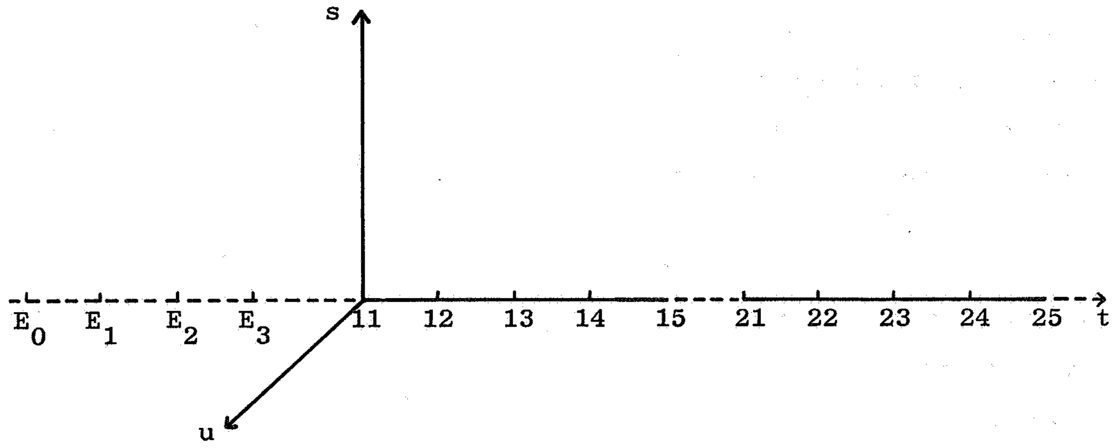

Figure 5.1 The state space

The s-variable is zero unless at least one damage has been claimed that year and moreover the coming damages are still covered by insurance. In that case the s-component denotes the extend of the last claim.

The u-variable is zero unless at least one damage has been claimed that year and coming damages are still covered by insurance. In that case the u-component denotes the time elapsed since the first claim that year.

Note that the s-component can only be different from zero if $1i \leq t < 1i + 1$, $i = 1, 2, 3, 4$. Consequently the state space consists of:

a) 4 points E_i $i = 0, 1, 2, 3$;

b) a 3-dimensional subspace (t,s,u) with $11 \leq t < 15$;

c) a 1-dimensional interval $21 \leq t < 25$.

Next the natural proces is described. This process can start in each state of the state space. In accordance with the premium paid the system runs through one of the time intervals $2i \leq t < 2i + 1 \quad i = 1, 2, 3, 4$, if no damage has been claimed that year. If no accident happens during the rest of the year the system is transferred to E_i. Since in the natural process no premiums are paid the system will stay there forever. However, if at time t' during the year an accident occurs the system is transferred to (t'-10,s',0) where s' denotes the damage. Since during the natural process irrespective of their extends all damages are claimed the system will stay in the 3-dimensional part of the state space for the remaining part of the year. Then the u-component is increasing with time. The s-component only changes if a second, third, etc. accident happens. At the end of the year the system is transferred to E_0 where it stays forever.

The two feasible decisions in the states E_i $i = 0, 1, 2, 3$ are the null-decision and the decision involving the payment of the premium E_i. The respective transformations are $E_i \rightarrow E_i$ and $E_i \rightarrow (2i+1,0,0)$. In states (t,s,0) an accident has just occured and the decisionmaker can suppress the claim if he wants. In that case the respective transfermation is $(t,s,0) \rightarrow (t+10)$. Note that a claim corresponds with a null-decision. This is in accordance with the fact that in the natural process all damages are claimed. In the states (t,s,u) with $u > 0$ only null-decisions are feasible. If an accident occurs in a state with $u > 0$ the decision not to claim is of course a bad decision and is considered to be in-feasible for that reason. Also in the states t with $21 \leq t < 25$ only null-decisions are feasible. In figure 5.2 states have been marked with two feasible decisions.

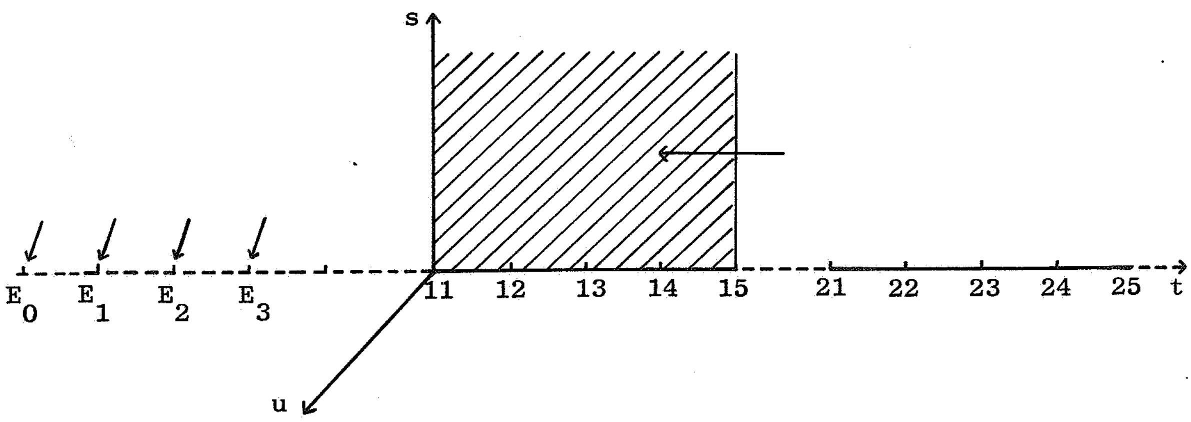

Figure 5.2 States with more than one feasible decision

From now on only strategies are considered which dictate payment of premium in the states E_i, i = 0, 1, 2, 3. By its definition the set A_0 consists of the states in which each strategy dictates an intervention. In this problem the states E_i, i = 0, 1, 2, 3 and the states (t,s,u) with $11 \leq t < 15$, $s \leq a_0$ and u = 0 constitute the set A_0 because each strategy dictates the payment of premium and suppression of the claim if the damage does not exceed the own risk. So we have:

$$(5.1) \qquad A_0 = \{E_i, i = 0,1,2,3\} \cup \{11 \leq t < 15, s \leq a_0, u = 0\}.$$

The non-empty subsets $A_{0,1}$ and $A_{0,2}$ of the set A_0 are chosen in such a way that the most simple expressions for the functions k(x;d) and t(x;d) are obtained. Here we choose:

$$(5.2) \qquad A_{0,1} = A_{0,2} = \{E_i, i = 0,1,2,3\}$$

and consequently for the associated stochastic walks it follows:

$$(5.3) \qquad \underline{w}_{0,1} = \underline{w}_{0,2}$$

and

$$(5.4) \qquad \underline{w}_{d,1} = \underline{w}_{d,2}.$$

To abbreviate the notation we write $\underline{w}_0$ and $\underline{w}_d$ respectively.

Consider the $\underline{w}_0$-walk having $(11+\tau,s,0)$ as initial state. During the walk $\underline{w}_0$ the system is subjected to the natural process. In the natural process each damage is claimed. The damage s at time τ is thus claimed and the costs $\min(s,a_0)$ are incurred. For each damage which occurs in the natural process we have the expected costs

$$(5.5) \qquad k(a_0) = \int_0^{a_0} s\, F(ds) + a_0 \int_{a_0}^{\infty} F(ds)$$

The expected number of accidents in a fraction $1 - \tau$ of a year is equal to $\lambda(1-\tau)$. Hence the expected costs incurred during the walk $\underline{w}_0$ are given by:

(5.6) $$k_0(1i+\tau,s,0) = \lambda(1-\tau)\ k(a_0) + \min(s,a_0).$$

The expected duration of the walk $\underline{w}_0$ is obviously:

(5.7) $$t_0(1i+\tau,s,0) = 1 - \tau.$$

Since decisions lead to deterministic transitions in this problem they will be denoted by the resulting states.

During the $\underline{w}_d$-walk starting in $x = (1i+\tau,s,0)$ the claim is suppressed and the system is transformed to state $d = 2i + \tau$. After this transformation the system is subjected to the natural process up to the end of the year. At that moment either state E_0 is taken on if a second accident occurred or state E_i if no seoond accident occurred. The expected duration $t_1(x;d)$ of the $\underline{w}_d$-walk for $x = (1i+\tau,s,0)$ and $d = 2i + \tau$ is given by:

(5.8) $$t_1(1i+\tau,s,0;2i+\tau) = 1 - \tau,$$

and the expected costs by:

(5.9) $$k_1(1i+\tau,s,0;2i+\tau) = \lambda(1-\tau)\ k(a_0) + s.$$

By (5.6) ... (5.9) and referring to (1) the following relations are obtained for the functions $k(x;d)$ and $t(x;d)$ with $x = (1i+\tau,s,0)$ and $d = 2i + \tau$:

(5.10) $$k(1i+\tau,s,0;2i+\tau) = \max(s-a_0,0),$$

(5.11) $$t(1i+\tau,s,0;2i+\tau) = 0.$$

Finally the k- and t-functions for the states E_{i-1}, i = 1, 2, 3, 4 are determined. The $\underline{w}_0$-walk having E_{i-1} as initial state ends immediately in that state because $E_{i-1} \in A_{0,1} = A_{0,2}$. In the natural process no premiums are paid. Hence $k_0(E_{i-1}) = t_0(E_{i-1}) = 0$. Next we consider the $\underline{w}_d$-walk having E_{i-1} as initial state. The payment of premium in state E_{i-1} transforms the system into state 2i. The $\underline{w}_d$-walk is from state 2i on subjected to the natural process. At the end of the year the walk ends either in state E_0 or state E_i. The expected duration of the $\underline{w}_d$-walk is thus one year. The expected costs of the $\underline{w}_d$-walk consists of the premium E_{i-1} and the expected costs incurred during the year in the natural process. So we obtain

(5.12) $$t(E_{i-1},2i) = 1,$$

(5.13) $$k(E_{i-1},2i) = E_{i-1} + \lambda k(a_0).$$

Suppose that z^* is an optimal strategy satisfying

(5.14) $$c(z^*;x) = \min_{z \in Z} c([z]z^*;x) \qquad \text{for all } x \in X,$$

where $(r(z^*),c(z^*;x))$ denotes a solution of (9) with $z = z^*$. It is no restriction to assume that (cf. Theorem 2(b)),

(5.15) $$c(z^*;E_0) = 0.$$

Consequently, by (10),

(5.15a) $$c(z^*;x) = 0 \qquad \text{for } x \notin A_{z^*},$$

since E_0 is the next intervention state when in state x the null-decision is made (claim!). We shall now demonstrate that a function s(t) determines A_{z^*}. To do this, we fix state x =(s,t,0) with $11 \le t < 15$ and $s > a_0$. Let z be a strategy which dictates the null-decision in state x. Then, by (13) and $z(E_0) = z^*(E_0)$, we have $c([z]z^*;x) = c(z^*;E_0) = 0$. From this and (5.14), (5.16) $$c(z^*;x)\le 0.$$

An intervention in state x (do not claim!) transfers the system into state t+10. Now, by (9), (10), (5.10), (5.11),

(5.17) $$c(z^*;\ t,s,0) = s-a_0 + c(z^*;t+10) \qquad \text{for } x = (t,s,0)\in A_{z^*}\ .$$

From (5.16) and (5.17),

(5.18) $$s-a_0 + c(z^*;t+10)\le 0 \qquad \text{for } x = (t,s,0)\in A_{z^*}$$

Next let $\bar{z}$ be a strategy which dictates an intervention in state x. Then, by (13), (11), (5.10) and (5.11),

(5.18a) $$c([\bar{z}]z^*;t,s,0) = s-a_0 + c(z^*;t+10).$$

Now, by (5.14) and (5.18a),

(5.19) $$c(z^*;t,s,0) \le s-a_0 + c(z^*;t+10).$$

From (5.15a) and (5.19),

(5.20) $$s-a_0 + c(z^*;t+10) \ge 0 \qquad \text{for } x = (t,s,0)\notin A_{z^*}.$$

Since the left side of (5.20) is a linear increasing function of s, it follows from (5.18) and (5.20) that A_{z^*} is determined by a function s(t) ($>a_0$) satisfying

(5.21) $$s(t)-a_0 + c(z^*;t+10) = 0 \qquad \text{for all } 11\le t<15\ ,$$

where x=(t,s,0) belongs to A_{z^*} if and only if s≤s(t). Observe that it is indifferent on the boundary s(t) of A_{z^*} to claim or not to claim.

Next we shall demonstrate how s(t) can be found. From (5.21),

(5.22) $$c(z^*;t+10) = a_0-s(t),$$

and from (5.17) and (5.22),

$$(5.23) \qquad c(z^*;t,s,0) = s - s(t) \qquad \text{for } a_0 < s \le s(t).$$

For $s \le a_0$ by (5.10) and (5.11) we have (c.f. (9)),

$$(5.24) \qquad c(z^*;t,s,0) = c(z^*;t+10) = a_0 - s(t).$$

Furthermore holds for $c(z^*;E_i)$, $i = 1, 2, 3$:

$$(5.25) \qquad c(z^*;E_i) = \lim_{t\uparrow 2i+1} c(z^*;t),$$

or by (5.22) and (5.25):

$$(5.26) \qquad c(z^*;E_i) = a_0 - \lim_{t\uparrow 1i+1} s(t).$$

Summarizing our results:

$$(5.27) \qquad c(z^*;x) = \begin{cases} 0 & \text{for } x \in E_0 \cup \{11 \le t < 15, s > s(t), u = 0\} \\ & \quad \cup \{11 \le t < 15, s \ge 0, u > 0\}, \\ a_0 - \lim\limits_{t\uparrow 1i+1} s(t) & \text{for } x \in \bigcup\limits_{i=1}^{3} E_i, \\ a_0 - s(t) & \text{for } x \in \{11 \le t < 15, s \le a_0, u = 0\}, \\ s - s(t) & \text{for } x \in \{11 \le t < 15, a_0 < s < s(t), u = 0\}, \\ a_0 - s(t-10) & \text{for } x \in \{21 \le t < 25\}. \end{cases}$$

From functional equation (9) it follows for $x = E_{i-1}$, $i = 1, 2, 3, 4$:

$$(5.28) \qquad c(z^*;E_{i-1}) = k(E_{i-1};2i) - r(z^*)\, t(E_{i-1};2i) + c(z^*;2i).$$

By substitution of (5.12) and (5.13) in (5.28) it follows:

(5.29) $$c(z^*;E_{i-1}) - c(z^*;2i) = E_{i-1} + \lambda k(a_0) - r(z^*).$$

From (5.27) and (5.29) we obtain:

(5.30) $$s(1i) = \begin{cases} E_0 + \lambda k(a_0) - r(z^*) + a_0 & \text{for } i = 1, \\ E_{i-1} + \lambda k(a_0) - r(z^*) + \lim\limits_{t\uparrow 1i} s(t) & \text{for } i = 2,\ 3,\ 4. \end{cases}$$

Furthermore we have the relation:

(5.31) $$\lim_{t\uparrow 14} s(t) = \lim_{t\uparrow 15} s(t).$$

For $x = (t,s,0)$ with $a_0 < s \le s(t)$ and $t = 1i+\tau$ it follows (c.f. (9)),

(5.32) $$c(z^*;t,s,0) =$$

$$k(t,s,0;t+10) - r(z^*)\ t(t,s,0;t+10) +$$

$$+ \int_{1i+1-t}^{\infty} c(z^*;E_i)\ \lambda e^{-\lambda\tau_1}\ d\tau_1 +$$

$$+ \int_{0}^{1i+1-t} \lambda e^{-\lambda\tau_1}\ d\tau_1 \int_{0}^{s(t+\tau_1)} c(z^*;t+\tau_1,y,0)\ F(dy) +$$

$$+ \int_{0}^{1i+1-t} \lambda e^{-\lambda\tau_1}\ d\tau_1 \int_{s(t+\tau_1)}^{\infty} c(z^*;E_0)\ F(dy).$$

Substitution of (5.10), (5.11) and (5.27) in (5.32) leads to:

(5.33) $$c(z^*;t,s,0) =$$

$$s - a_0 + e^{-\lambda(1i+\tau-t)}\ (a_0 - \lim_{t\uparrow 1i+1} s(t)) +$$

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

$$+\int_0^{li+1-t} \lambda e^{-\lambda\tau_1}\, d\tau_1 \int_0^{a_0} (a_0-s(t+\tau_1))\, dF(y).$$

After substitution of $s = s(t)$ and (5.29) the differentiation of (5.33) with respect to t leads to:

$$\text{(5.34)} \qquad \frac{ds(t)}{dt} = \lambda \int_{a_0}^{\infty} (y-a_0)\, dF(y) - \lambda \int_{s(t)}^{\infty} (y-s(t))\, dF(y).$$

By partial integration this functional equation can be written in the more simple form:

$$\text{(5.35)} \qquad \frac{ds(t)}{dt} = \lambda \int_{a_0}^{s(t)} (1-F(y))\, dy.$$

Except a translation along the t-axis the boundary $s(t)$ is determined by (5.35). In other words the boundary of A_{z^*} for i = 1, 2, 3, 4 are in the t-direction translated parts of one curve satisfying (5.35). The location of each part on this curve has to be determined from the relations (5.30) and (5.31).

Suppose that $r(z^*)$ is known, then $s(11)$ is solved from (5.30). From the curve $s = s(t)$ we find $\lim_{t\uparrow 12} s(t)$. From (5.30) we obtain $s(12)$. Similarly we compute $\lim_{t\uparrow 13} s(t)$, $s(13)$, $\lim_{t\uparrow 14} s(t)$, $s(14)$ and $\lim_{t\uparrow 15} s(t)$. If $r(z^*)$ is not known its value is determined by relation (5.31).

It should be noted that the functional equation (5.35) has an analytical solution in the case the damage per accident is exponentially distributed. We have then for $F(s) = 1 - e^{-\mu s}$:

(5.36) $$\frac{ds(t)}{dt} = \frac{\lambda}{\mu} e^{-\mu a_0} (1-e^{-\mu(s(t)-a_0)}).$$

The solution of (5.36) is given by:

(5.37) $$s(t) = a_0 + \frac{1}{\mu} \ln \{1+e^{\lambda(t+c_i)e^{-\mu a_0}}\},$$

where the c_i $i = 1, 2, 3, 4$ are integration constants each corresponding to the time intervals $1i \leq t < 1i + 1$, $i = 1, 2, 3, 4$. If the distribution of the damage is not exponential we have to solve (5.35) numerically in most cases.

5.4 Some numerical results.

The following numerical data are used:

$$E_0 = 1.6$$
$$E_1 = 1.4$$
$$E_2 = 1.2$$
$$E_3 = 1.1$$
$$a_0 = 0.4.$$

For these data and $\lambda = 2$ accidents per year five distributions with the same expectation were investigated. The type of distribution, its expectation and coefficient of variation are given in the following table:

Number of curve	Type of distribution	Expectation	Coefficient of variation
1	exponential	1	1
2	gamma	1	1/3
3	log normal	1	1
4	log normal	1	1/3
5	log normal	1	3

The densityfunctions are sketched in figure 5.3

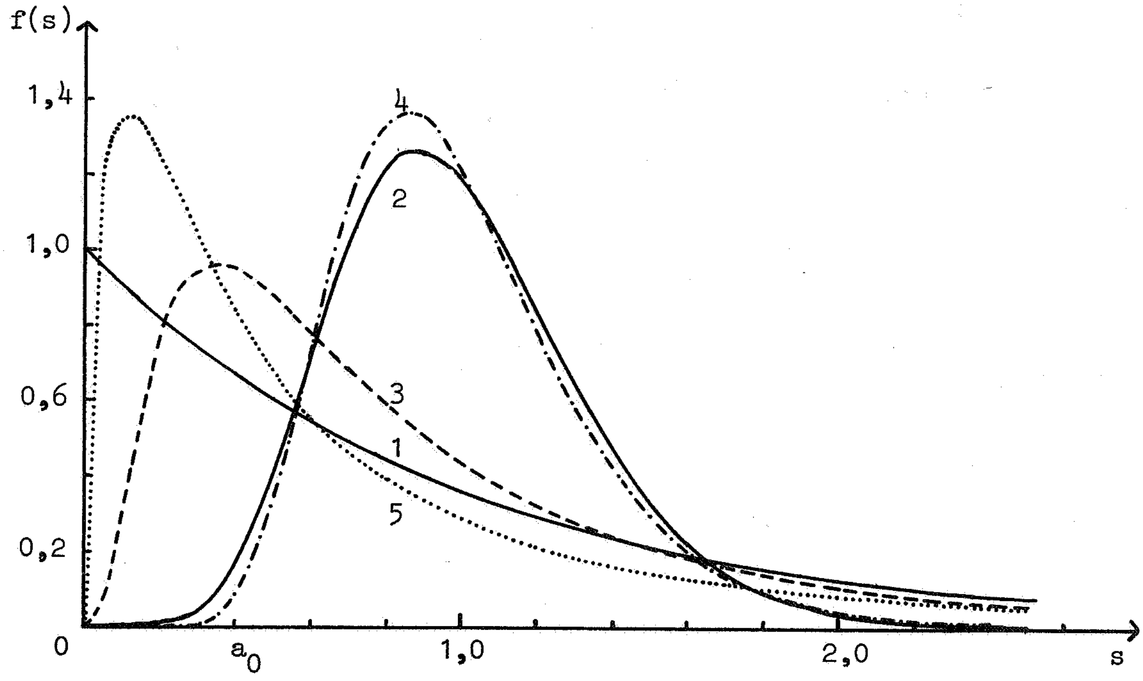

Figure 5.3. The five used damage distributions

The corresponding optimal strategies are presented in figure 5.4. From these results it can be deduced that for distributions with the same mean and variance the optimal strategy are nearly the same. Further, if the variance increases the boundary of A_{z*} moves upwards. The results were obtained by a computer program especially written for this problem.

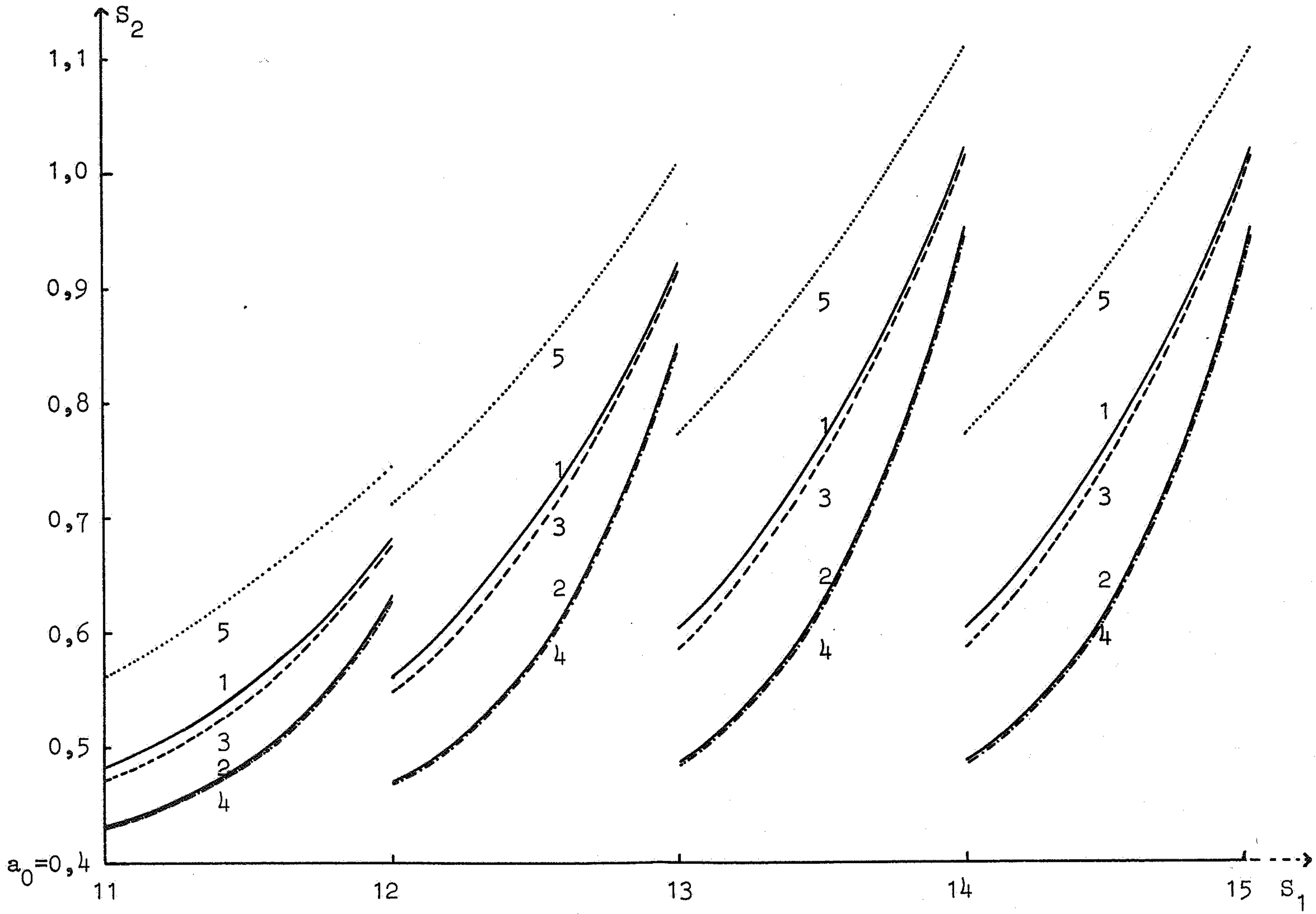

Figure 5.4 The optimal strategy for these distributions

A quality control problem*)

Consider a machine that produces each day a product whose quality will be identified with one of the integers 1,...,M. At the beginning of each day the machine may be inspected. The costs of an inspection are equal to J. An inspection is done when the decision maker wants to find out the quality of the product that will be produced that day. After an inspection he knows this quality. When he thinks that this quality (say i) is not acceptable, he decides to a revision of the machine. The revision costs are R(i). Both the time needed to inspect the machine and the time needed to revise the machine may be neglected. After a revision the machine produces that day a product of quality M. The production costs incurred at a day are given by p(i) when the machine produces that day a product of quality i, $(1 \le i \le M)$.

If the machine produces at day t a product of quality i, then, with probability p_{ij}, at the beginning of day t + 1 the machine will be in a condition to produce a product of quality $j(1 \le j \le M)$, and, with probability p_{i0}, the machine will be defect at the beginning of day t + 1. We assume $p_{i0}<1$ and $p_{i0} + p_{i1} + \ldots + p_{iM} = 1$ for all $1 \le i \le M$. When the machine is defect at the beginning of a day, the machine is repaired. The repair costs are R(0) and the repair time may be neglected. After a repair the machine produces that day a product of quality M.

It is assumed that the machine becomes eventually defect when never a revision occurs.

The decision maker wants to determine a strategy for inspecting and revising the machine such that the long-run average costs per day are minimal.

---

*) The elaboration of this problem will be sent to the participants before november 1, 1973.

References.

1. J.L. DOOB, stochastic Processes, Wiley, New York, 1953.
2. R.A. HOWARD, Dynamic Programming and Markov Processes, M.I.T. Press, Cambridge, Massachusetts, 1960.
3. W.S. JEWELL, "Markov-Renewal Programming. I: Formulation, Finite Return Models. II: Infinite Return Models". Operations Research 10 (1963), pp. 938-972.
4. G.DE LEVE, Generalized Markovian Decision Processes, Part I: Model and Method, Part II: Probabilistic Background, Mathematical Centre Tracts, No. 3 and 4, Amsterdam, 1964.
5. G.DE LEVE, H.C. TIJMS AND P.J.WEEDA, Generalized Markovian Decision Processes, Applications, Mathematical Centre Tract, No. 5, Amsterdam, 1970.
6. P.J. WEEDA, " Generalized Markov-Programming Applied to Semi-Markov Decision Problems and two Algorithms for its Cutting Operation", Report BW 24/73, Mathematisch Centrum, Amsterdam, 1973.