STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 400

An application of Markov-programming in a one dimensional nondenumerable state space.

by

P.J. Weeda

august 1968.

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 and some basic notions.

The production of a continuous product is considered with a finite number of possible production rates a_i, $i=1, \ldots, N$ with $a_1=0$. The production costs per time unit for production rate a_i are denoted by $c_p(i)$ with $c_p(i) \geq c_p(i-1)$ for $i=2, \ldots, N$ and $c_p(1)=0$.
The product is kept in stock. Stockholding costs are c_s per unit time per unit product. If the stocklevel reaches a given maximum amount M then the production has to be stopped. The arrivals of orders are described by a stationary Poisson-process with parameter λ. The order size $\underline{y}$ is distributed according to a given distribution function $F(y)$. Orders are fulfilled immediately either by the available stock or by purchases elsewhere at a given higher cost c_r per unit product. Furthermore the costs of a transition from production rate a_i to production rate a_j are given by $c_q(i,j)$ with $i,j \in \{1,2, \ldots, N\}$.

We will show in this paper how the method developed in [1] leads to the optimal production strategy in this problem. A survey of the method is given in [2] and [3]. We will only state the definitions of the necessary functions and will derive functional equations for them, specialized for the considered problem. A method of solution for the functional equations for the function $c(z;x)$ will be given. Numerical methods of solving the functional equations for the functions $k(x;d)$, $t(x;d)$ and the probability distributions of entering a set of states within the set of interventionstates from states outside this set for an arbitrary strategy z are considered as a separate subject and will not be given here. Finally the procedure in the strategy-improvement routine will be outlined.

The production manager is allowed to control the system by changing over to another production rate. His interventions will depend on the state of the system, which is specified by two state variables: the production rate a_i and the stock level s.
The state space X of this problem consists of states $x=(i,s)$ with $-\infty < s < \infty$ and $1 \leq i \leq N$.

The state space is presented in figure 1.1.

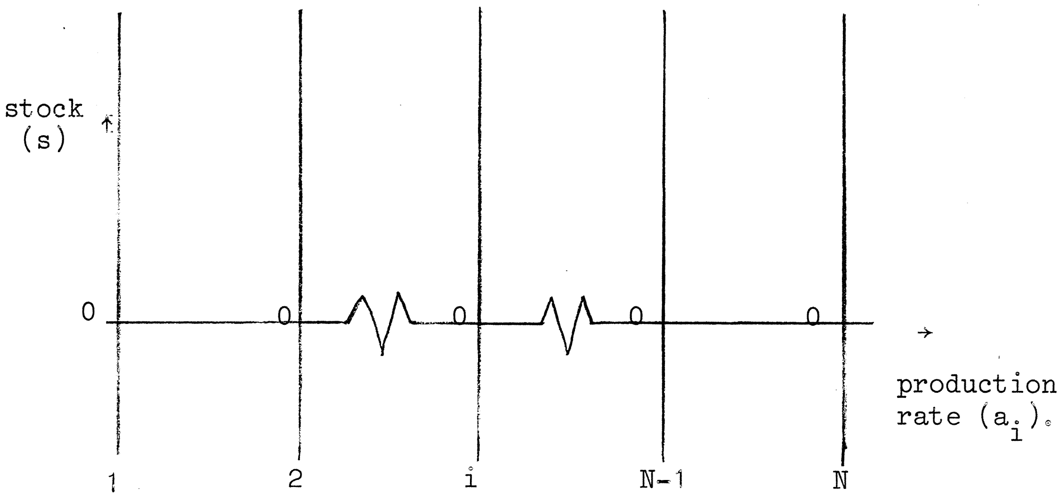

Figure 1.1.: The state space X.

If the production manager does not intervene, the system is subject to the natural process. The natural process is defined for every state x. During the natural process the system remains on the same production rate a_i as it is in the starting state x. Additional purchases are included in the natural process.

If the production manager does intervene then the resulting process will be different from the natural process. It will be called the decision process. In every state of the system x the production manager has to make a choice between the possible decisions in state x. The set of possible decisions will be denoted by D(x), a particular decision by $d \in D(x)$. D(x) includes the decision not to intervene, called the null decision. By an intervention in state x=(i,s) the system is transferred into state (j,s) with $j \neq i$ if $s \geq 0$ and into state (j,0) if $s < 0$. Between interventions the system is subject to the natural process. If to every state a decision is fixed, we have a strategy. We will denote a strategy by ƶ. The decision dictated by strategy ƶ in state x will be denoted by ƶ(x). A strategy is called optimal if it minimizes the average costs per unit time in the long run, denoted by r(ƶ). r(ƶ) does not depend on the starting state if there is only one ergodic set of states in the decisionprocess, as will be the case in this problem.

A strategy $\mathfrak{z}$ will dictate an intervention in the states belonging to a closed set $A_{\mathfrak{z}}$, called the set of intervention states for strategy $\mathfrak{z}$. The state space will consist of two distinct non-empty sets of states: the set of intervention states $A_{\mathfrak{z}}$ and its complement, the set of non-intervention states states where null-decisions are dictated by strategy $\mathfrak{z}$.

Furthermore it is assumed that there exists a non-empty set of states A_0 where every strategy dictates an intervention. Hence for every strategy there holds:

$$A_{\mathfrak{z}} \supset A_0 \qquad (1.1)$$

If the maximum stock level M is reached in this problem then the production is always stopped. Hence the states (i,M) for $i=2, \ldots, N$ are elements of A_0. Also in the states $(1,s)$ with $s < 0$ an intervention will be dictated by every strategy. Hence the states $(1,s)$ with $s < 0$ are elements of A_0 *). The set A_0 will be given by

$$A_0 = \{(i,s) \mid s \geq M, i > 1\} \cup \{(i,s) \mid s < 0, i=1\} \qquad (1.2)$$

and is presented in figure 1.2 by the shaded intervals of s.

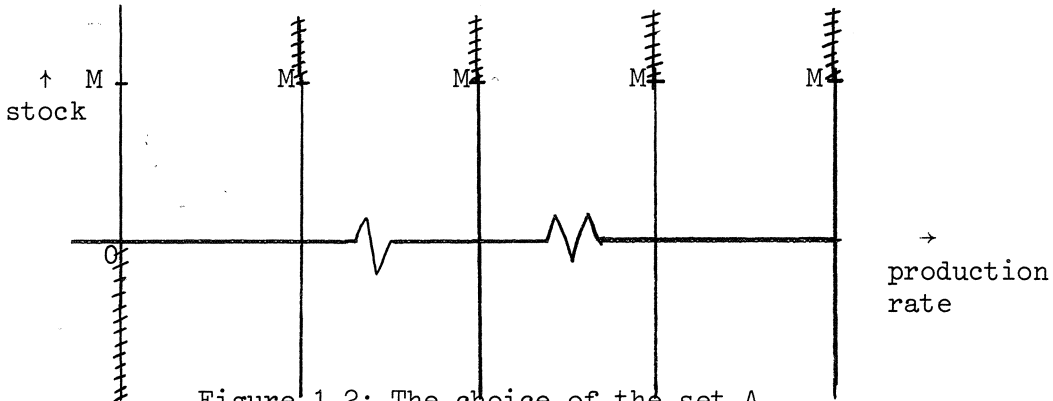

Figure 1.2: The choice of the set A_0.

*) Note that only one strategy is excluded by this choice of the set A_0, namely the strategy which satisfies customer demand by purchases at the cost c_r per unit product.

§2. The determination of the strategy-independent functions.

We will derive now the functional equations for the functions $k(x;d)$ and $t(x;d)$. These functions denote respectively the difference in expected costs and in expected duration between two stochastic walks starting in x. In the first walk the decision d is taken in state x after which the system is subject to the natural process until the first state in A_0 is reached. The second walk is only subject to the natural proces from state x on. Denote by $k_i(s)$ and $t_i(s)$ respectively the expected costs and duration of the second walk starting in state $x=(i,s)$. If the decision d transfers the system from state (i,s) into state (j,s) then we have for $k(x;d)$ and $t(x;d)$:

$$k(x;d) = c_q(i,j) + k_j(s) - k_i(s) \tag{2.1}$$

$$t(x;d) = t_j(s) - t_i(s) \tag{2.2}$$

Because states (i,s) with $s > M$ and $i \geq 1$ are never reached in the decision process, it will be sufficient to determine the functions $k_i(s)$ and $t_i(s)$ for $s \leq M$ and $1 \leq i$

For i=1 thereis no production. For states $(1,s) \varepsilon A_0$ we have:

$$t_1(s)=0 \tag{2.3}$$

For $s \geq 0$ the walk terminates when the stock level drops below zero because then the set A_0 is reached. If we denote the arrival time of the next order by $\underline{\tau}$ and the order size by $\underline{y}$, then we have:

$$\underline{t}_1(s) = \underline{\tau} + \begin{cases} t_1(s-\underline{y}) & s \geq \underline{y} \\ 0 & s < \underline{y} \end{cases} \tag{2.4}$$

Taking expectations leads to the following functional equation for $t_1(s) = \mathcal{E}\,\underline{t}_1(s)$ ($\mathcal{E}\,\underline{\tau}$ being $^1/\lambda$):

$$t_1(s) = \frac{1}{\lambda} + \int_0^s t_1(s-y)\, dF(y). \tag{2.5}$$

For $s < 0$ we have, because $(1,s) \in A_0$:

$$k_1(s)=0 \tag{2.6}$$

For $s \geq 0$ we have

$$\underline{k}_1(s) = c_s . s . \underline{\tau} + \begin{cases} k_1(s-\underline{y}) & s \geq \underline{y} \\ c_r(\underline{y}-s) & s < \underline{y} \end{cases} \tag{2.7}$$

By taking the expectations on both sides of (2.7) we have the following functional equation for $k_1(s) = \mathcal{E}\,\underline{k}_1(s)$ with $s \geq 0$:

$$k_1(s) = c_s\, s\, \frac{1}{\lambda} + c_r \int_s^\infty (y-s)dF(y) + \int_0^s k_1(s-y)dF(y) \tag{2.8}$$

For production rates a_i with $1 < i \leq N$ the stock level is increasing linearly between the arrivals of orders. The walk terminates when the maximum stock is reached because $(i,M) \in A_0$ for $i > 1$.
So wee have:

$$t_i(M)=0 \tag{2.9}$$

$$k_i(M)=0 \tag{2.10}$$

If the stock level drops below zero then the walk continues from state $(i,0)$ after an additional purchase.

The functional equations for $t_i(s)$ and $k_i(s)$ are derived for $i > 1$ by considering the possible events during a small time $\Delta\tau$.
Let the stock level at time 0 be s. Suppose that the first order arrives at time $\underline{\tau}_1$. The ordersize will be a stochastic amount $\underline{y}$.
Consider a small time interval $(0, \Delta\tau]$ and denote the stocklevel at time $\tau + \Delta\tau$ by $\underline{s}'$.
Then we have for $\underline{s}'$:

$$\underline{s}' = \begin{cases} s+a_i\Delta\tau & \underline{\tau}_1 > \Delta\tau & \\ s+a_i\Delta\tau-\underline{y} & \underline{\tau}_1 \leq \Delta\tau & y \leq s+a_i\underline{\tau}_1 \\ a_i(\Delta\tau - \underline{\tau}_1) & \underline{\tau}_1 \leq \Delta\tau & y > s+a_i\,\Delta\tau \end{cases} \qquad (2.11)$$

neglecting the case of more than one arrival in $(\tau,\tau+\Delta\tau]$ which happens with probability $o(\Delta\tau)$.
Furthermore we have:

$$P\{\underline{\tau}_1 > \Delta\tau\} = 1- \lambda\,\Delta\,\tau + o(\Delta\tau) \qquad (2.12)$$

$$P\{\underline{\tau}_1 \leq \Delta\tau\} = \lambda\,\Delta\,\tau + o(\Delta\,\tau) \qquad (2.13)$$

For $\underline{t}_i(s)$ we have:

$$\underline{t}_i(s)=\Delta\tau + t_i(\underline{s}') \qquad (2.14)$$

By taking expectations in both sides of (2.14) we have for $t_i(s) = \mathcal{E}\underline{t}_i(s)$ with $0 \leq s < M$ and $i > 1$:

$$\begin{aligned} t_i(s) = {} & (1-\lambda\Delta\tau + o(\Delta\tau))\,\{\Delta\tau + t_i(s+a_i\Delta\tau)\} \\ & + (\lambda\Delta\tau + o(\Delta\tau))\,\{\Delta\tau + \int_0^{s+a_i\tau_1} t_i(s-y+a_i\Delta\tau)dF(y) \\ & + \int_{s+a_i\tau_1}^{\infty} t_i(a_i(\Delta\tau-\tau_1))\,dF(y)\} \\ & + o(\Delta\tau) \end{aligned} \qquad (2.15)$$

where $0 < \tau_1 \leq \Delta\tau$.
Dividing by $\Delta\tau$, replacing $\Delta\tau$ by $\frac{\Delta s}{a_i}$ and performing the limit operation $\Delta s \to 0$, we arrive at the functional equation:

$$\frac{dt_i(s)}{ds} = \frac{\lambda}{a_i} t_i(s) - \frac{1}{a_i} - \frac{\lambda}{a_i} \int_0^s t_i(s-y)dF(y) - \frac{\lambda}{a_i} t_i(0)(1-F(s)) \tag{2.16}$$

For the stockcosts within $\Delta\tau$, ignoring higher order terms we have:

$c_s \cdot s \cdot \Delta\tau$ if $\underline{\tau}_1 > \Delta\tau$

$c_s \cdot s \cdot \underline{\tau}_1 + c_s \cdot \underline{s}'(\Delta\tau - \underline{\tau}_1)$ if $\underline{\tau}_1 \leq \Delta\tau$ and $\underline{y} \leq s + a_i\tau_1$

$c_s \cdot s \cdot \underline{\tau}_1$ if $\underline{\tau}_1 \leq \Delta\tau$ and $\underline{y} > s + a_i\underline{\tau}_1$

Additional purchases are done only in the case that $\underline{\tau}_1 \leq \Delta\tau$ and $\underline{y}$ s + $a_i\tau_1$. The costs are $c_r(\underline{y} - a_i\tau_1 - s)$. The production costs are $c_p(i) \cdot \Delta\tau$.
These considerations and the arguments used at the derivation of the functional equation for $t_i(s)$ lead to the functional equation:

$$\frac{dk_i(s)}{ds} = \frac{\lambda}{a_i} k_i(s) - \frac{c_s \cdot s}{a_i} - \frac{c_p(i)}{a_i} - \frac{\lambda}{a_i}(1-F(s))k_i(0) - \frac{\lambda}{a_i} c_r \int_s^\infty (y-s)dF(y) - \frac{\lambda}{a_i} \int_0^s k_i(s-y)dF(y). \tag{2.17}$$

for $k_i(s)$ with $0 \leq s < M$ and $i > 1$.
For $s < 0$ we have $(i > 1)$:

$$t_i(s) = t_i(0) \tag{2.18}$$

$$k_i(s) = k_i(0) \tag{2.19}$$

3. Determination of the strategy-dependent functions $c(z;x)$.

According to the method presented in [1] the function $c(z;x)$ for a given strategy z has to be obtained from the following functional equation:

$$c(z;x) \overset{\text{def}}{=} k(x;z(x)) - r(z) \cdot t(x;z(x)) + \int_{A_z} P_{A_z}^{(1)}(du;z;x)\, c(z;u) \qquad (3.1)$$

where $\underline{u} \in A_z$ denotes the first future intervention state assumed by the system if it starts in state x. The probability of $\underline{u}$ is given by $P_{A_z}^{(1)}(\cdot;x)$.

If x is a state where the nulldecision is dictated then $k(x;z(x)) = t(x;z(x)) = 0$ and (3.1) reduces to

$$c(z;x) = \int_{A_z} P_{A_z}^{(1)}(du;z;x)c(z;u) \qquad (3.2)$$

It will depend on the location of the set of intervention states A_z for an arbitrary strategy z in the state space how the functional equation (3.1) specializes in this particular problem. We shall first consider strategies with only one set of non-intervention states for each production rate a_i i=1, ...,N. This is no restriction because strategies with two or more distinct sets of non-intervention states for some production rates can be reduced to the preceding class by an extension of state space.

A representative of the considered class of strategies is sketched in figure 3.1. The set A_z is given by the shaded intervals and is completely specified by the states $(i,b_z^{(1)})$ and $(i,b_z^{(2)})$ for i=1, ...,N as follows:

1. For states (i,s) with $1 < i \leq N$ and $s \geq b_{ƶ}^{(2)}$ an intervention is dictated. Because in states (i,M) with $1 < i \leq N$ always an intervention is dictated (A_0!) we have $b_{ƶ}^{(2)} \leq M$. The subsets of $A_{ƶ}$ with states (i,s) for each i with $1 < i \leq N$ and with $s \geq b_{ƶ}^{(2)}$ will be denoted by $A_{ƶ}^{(i,2)}$. For $i=1$ we put $b_{ƶ}^{(2)} \overset{\text{def}}{=} \infty$ for each possible strategy. Cousequently the subsets $A_{ƶ}^{(i,2)}$ exist only for $i > 1$.

2. For states (i,s) with $s \leq b_{ƶ}^{(1)}$ an intervention is dictated. Let the set of indices of these production rates for which $b_{ƶ}^{(1)} \geq 0$ be denoted by $I_{ƶ}$. For production rates a_i with $i \notin I_{ƶ}$ we put $b_{ƶ} = -\infty$. For each $i \in I_{ƶ}$ we will denote the set of states with $s \leq b_{ƶ}^{(1)}$ by $A_{ƶ}^{(i,1)}$. Note that for every strategy $N \notin I_{ƶ}$ and $1 \in I_{ƶ}$.

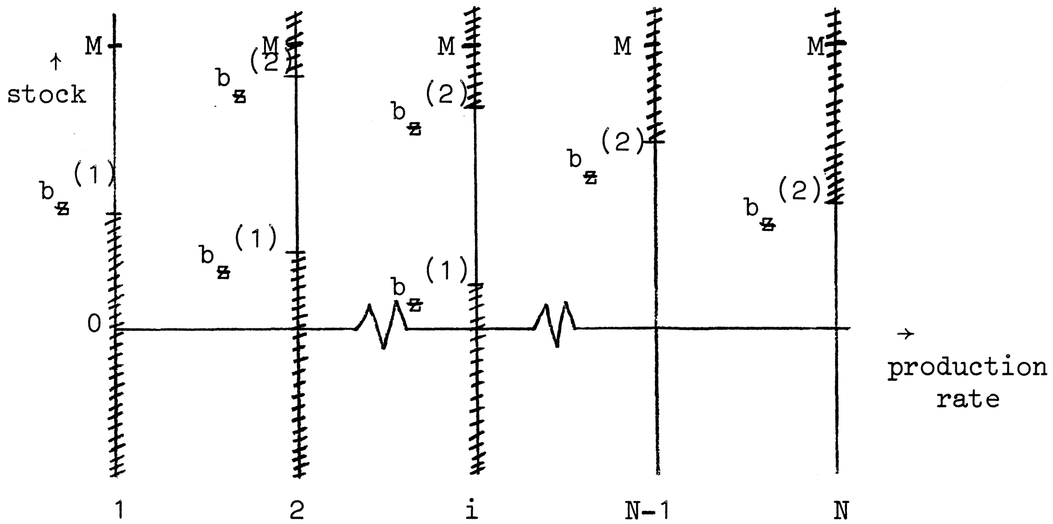

Figure 3.1 The set $A_{ƶ}$ for a strategy ƶ of the considered class of strategies.

A strategy is further specified by the decision $ƶ(i,s)$ in each $(i,s) \in A_{ƶ}$. A decision means a transformation to state (j,s') with $j \neq i$ and $s'=s$, if it is not a null decision. This implicates that except by the states $(i,b_{ƶ}^{(1)})$ and $(i,b_{ƶ}^{(2)})$ a strategy has to be specified by the dictated decision in each intervention state in $A_{ƶ}$.

Usually for the in practice occuring strategies each set $A_z^{(i,1)}$ or $A_z^{(i,2)}$ can be subdivided into a finite number of subsets, where the decision, dictated by the strategy, is the same for each state belonging to such a subset. The states that separate these subsets together with the decision attached to the states belonging to these subsets specify completely a strategy of the considered class.

The general functional equation for c(z;x) given by (3.1) specializes to the following set of functional equations. For non-intervention states (i,s) we have:

$$c(z;i,s)= \begin{cases} \int_{u\varepsilon A_z^{(i,1)}} c(z;i,u)\; d\, G_i(u;s;b_z^{(1)}) & i=1 \\ \int_{u\varepsilon A_z^{(i,1)}} c(z;i,u)\; d\, G_i(u;s;b_z^{(1)},b_z^{(2)}) & i\varepsilon I_z,\ i\neq 1 \\ \qquad + P_i(b_z^{(2)};s;b_z^{(1)}) .\, c(z;i.b_z^{(2)}) & \\ c(z;i,b_z^{(2)}) & i\notin I_z \end{cases} \qquad (3.3)$$

where:

1) $G_i(u;s;b_z^{(1)})$ with $i \varepsilon I_z$ denotes the probability that the first future interventionstate $(i,\underline{u})$ after starting in the non-intervention state (i,s) is contained in the set $\{(i,\underline{u}) \mid u \leq \underline{u} \leq b_z^{(1)}\} \varepsilon A_z^{(i,1)}$.
2) $P_i(b_z^{(2)};s;b_z^{(1)})$ with $i > 1$ denotes the probability that the first future intervention state, after starting in the non-intervention state (i,s), is given by $(i,b_z^{(2)})$.

For intervention states $(i,u) \in A_z$ the system is transferred into state $z(i,u)$. We have for $c(z;i,u)$:

$$c(z;i,u) = k(i,u;z(i,u)) - r(z)\ t(i,u;z(i,u)) + c(z;z(i,u)) \qquad (3.4)$$

It is easily verified that in (3.1) the function $c(z;x)$ can be determined only relative to an arbitrary constant. If we put $c(z;x)=0$ for one state x then the set of functional equations (3.3) and (3.4) will have a unique solution.

To solve the set of functional equations for $c(z;x)$ given by (3.3) and (3.4) we shall use the special properties of the states $(i,b_z^{(2)})$. We will denote the set of states $(i,b_z^{(2)})$ with $i=2, \ldots, N$ by B_z and the states of this set by y_i. Note that for each strategy there are always N-1 of these states.

Let $\underline{u}_n$ $n=1,2, \ldots$ be the sequence of future intervention states assumed by the decision process for an arbitrary strategy z and starting in an arbitrary state (i,s). If $(i,s) \in A_z$ then $\underline{u}_1 \overset{def}{=} (i,s)$. As proved in [1]) the sequence $\underline{u}_n$ $n=(1,2, \ldots)$ constitutes a stationary Markov-process with discrete time parameter and a non-denumerable state space A_z. In the Markov-process in A_z there is inbedded a stationary Markov-chain with a discrete time parameter and a finite state space $B_z \subset A_z$.

We consider now realisations of the decision process starting in an arbitrary state (i,s) and terminating in $\underline{y}$, the first future state assumed in B_z. These realisations terminate with probability one in a finite time. Before reaching $\underline{y}$, the decision process assumes a stochastic number $\underline{n}$ of intervention states $\underline{n}_l$, $l=1,2, \ldots \underline{n}$ with each $\underline{u}_l \in \bigcup_{i \in I_z} A_z^{(i,1)}$. The functions $ck(z;i,s)$ and $ct(z;i,s)$ are defined being the expected value of the sum of the contributions in each state $\underline{u}_l$ of the functions $k(x;d)$ and $t(x;d)$ respectively.

We have for non-intervention states (i,s):

$$ck(z;i,s) \overset{\mathrm{def}}{=} \mathcal{E} \sum_{l=1}^{n} k(\underline{u}_l; z(\underline{u}_l)) \tag{3.5}$$

$$ct(z;i,s) \overset{\mathrm{def}}{=} \mathcal{E} \sum_{l=1}^{n} t(\underline{u}_l; z(\underline{u}_l)) \tag{3.6}$$

where the expectation is taken with respect to the joint probability distribution of $\underline{u}_l$ ($l=1, \ldots, \underline{n}$) and $\underline{n}$.

For intervention states u_1 we have:

$$ck(z;u_1) \overset{\mathrm{def}}{=} \begin{cases} k(u_1;z(u_1)) + \mathcal{E} \sum_{l=2}^{n} k(\underline{u}_l; z(\underline{u}_l)) & u_1 \notin B_z \\ 0 & u_1 \in B_z \end{cases} \tag{3.7}$$

$$ct(z;u_1) \overset{\mathrm{def}}{=} \begin{cases} t(u_1;z(u_1)) + \mathcal{E} \sum_{l=2}^{n} t(\underline{u}_l; z(\underline{u}_l)) & u_1 \notin B_z \\ 0 & u_1 \in B_z \end{cases} \tag{3.8}$$

where the expectation is taken with respect to the joint probability distribution of $\underline{u}_l$ $l=2, \ldots, \underline{n}$ and $\underline{n}$.

We consider next realisations of the Markov-chain in B_z, starting in state y_i and terminating in $\underline{y}_j$ the first future state assumed in B_z. For $c(z;y_i)$ with $y_i \in B_z$ the following set of N-1 linear equations hold:

$$c(z;y_i) = ck^*(z;y_i) - r(z).\ ct^*(z;y_i) + \sum_{y_j \in B_z} P(y_j;y_i)c(z;y_j) \qquad i=2, \ldots N \tag{3.9}$$

where $P(y_j;y_i)$ denotes the probability of y_j being the first future state in B_z after starting in y_i.

The functions $ck^{*}(z;y_i)$ and $ct^{*}(z;y_i)$ are also related to the walk starting in y_i and terminating in the first future state in B_z:y_j. They follow from the relations

$$ck^{*}(z;y_i)=k(y_i;z(y_i))+ck(z;z(y_i)) \quad (3.10)$$

$$ct^{*}(z;y_i)=t(y_i;z(y_i))+ct(z;z(y_i)) \quad (3.11)$$

where $ck(z;z(y_i))$ and $ct(z;z(y_i))$ follow from (3.5) and (3.6) while $z(y_i)$ denotes the decision dictated by z in y_i. If $ck^{*}(z;y_i)$, $ct^{*}(z;y_i)$ and the transition probabilities $P(y_j;y_i)$ are known then we have N-1 linear equations in the N unknowns $c(z;y_i)$ i=2, ...,N and r(z). To obtain a unique solution we put $c(z;Y_N)=0$.

After having solved this set of linear equations we consider, in order to compute c(z;i,s) for $(i,s) \notin B_z$, realisations of the decision process starting in (i,s) and terminating in the first state y_j assumed in the set B_z. The following relations holds for c(z;i,s) with $(i,s) \notin B_z$:

$$c(z;i,s)=ck(z;i,s)-r(z)ct(z;i,s)$$

$$+ \sum_{y_j \varepsilon B_z} P(y_j;i,s)\ c\ (z;y_j) \quad (3.12)$$

where ck(z;i,s) and ct(z;i,s) are defined by (3.5), (3.7),(3.6) and (3.8). $P(y_j;i,s)$ denotes the probability of reaching $y_j \varepsilon B_z$, starting in $(i,s) \notin B_z$.

Numerically the function c(z;i,s) can be determined by simulation of the stochastic walks on which relations (3.9) and (3.10) are based. Simulation has the advantage that it can be done for every arbitrary strategy, but it is time consuming compared with other numerical methods.

For strategies with the property that in each set $A_z^{(i,1)}$ only interventions are dictated that increase the production rate, the functions ck(z;i,s) and ct(z;i,s) can be computed by numerical integration from the following recursion relations:

$$ck(z;i,s)= \begin{cases} \displaystyle\int_{(i,u)\varepsilon A_z^{(i,1)}} \{k(i,u;z(i,u)) + ck(z;z(i,u)\} \; dG_i\,(u;s;b_z^{(1)},\ b_z^{(2)}) & (i,s) \notin A_z;\ i \ \varepsilon\ I_z \\ k(i,s;z(i,s)) + ck(z;z(i,s)) & (i,s)\ \varepsilon\ \overline{B}_z \cap A_z \\ 0 & (i,s)\ \varepsilon\ B_z \\ 0 & i \notin I_z;\ (i,s) \notin A_z \end{cases} \quad (3.13)$$

The same relations hold for ct(z;i,s) with ck and be replaced respectively by ct and t.

The probabilities $P(y_j;i,s)$ follow from:

$$P(y_j;i,s)= \begin{cases} \displaystyle\int_{(i,u)\varepsilon A_z^{(i,1)}} P(y_j;z(i,u))dG_i(u;s;b_z^{(1)},\ b_z^{(2)}) & (i,s) \notin A_z,\ i < j,\ i\ \varepsilon\ I_z \\ P_i(b_z^{(2)};s;b_z^{(1)}) & (i,s) \notin A_z,\ i=j\neq 1,\ i\ \varepsilon\ I_z \\ 1 & (i,s) \notin A_z,\ i=j;i \notin I_z \\ P(y_j;z(i,s)) & (i,s)\ \varepsilon\ A_z \\ 0 & (i,s) \notin A_z,\ i > j \end{cases} \quad (3.14)$$

For the probabilities $P_i(b_z^{(2)};s;b_z^{(1)})$ functional equations can be derived using the same arguments as before at the derivation of the function $k_i(s)$ and $t_i(s)$. We restrict ourselves to stating the results:

$$\frac{\partial}{\partial s} P_i(b_z^{(2)};s;b_z^{(1)}) = \frac{\lambda}{a_i} P_i(b_z^{(2)};s;b_z^{(1)}) - \frac{\lambda}{a_i}\int_{b_z^{(1)}}^{s} P_i(b_z^{(2)};s-y;b_z^{(1)})dF(y)$$

$$\text{for } b_z^{(1)} < s < b_z^{(2)};\ i \in I_z; i > 1 \tag{3.15}$$

$$P_i(b_z^{(2)};s;b_z^{(1)}) = 1 \qquad \text{for } i \notin I_z; b_z^{(1)} < s < b_z^{(2)} \tag{3.16}$$

$$P_i(b_z^{(2)};b_z^{(2)};b_z^{(1)}) \overset{\text{def}}{=} 1 \qquad i=2, \ldots, N \tag{3.17}$$

For the density functions $g_1(u;s;b_z^{(1)})$ and $g_i(u;s;b_z^{(1)},b_z^{(2)})$ corresponding to $G_1(u;s;b_z^{(1)})$ and $G_i(u;s;b_z^{(1)},b_z^{(2)})$ defined before, we can derive the functional equations:

$$g_i(u;s;b_z^{(1)}) = f(s-u) + \int_{b_z^{(1)}}^{s} g_i(u;s-y;b_z^{(1)})dF(y) \quad i=1 \tag{3.18}$$

$$\frac{\partial}{\partial s} g_i(u;s;b_z^{(1)},b_z^{(2)}) = \frac{\lambda}{a_i} g_i(u;s;b_z^{(1)}, b_z^{(2)}) - \frac{\lambda}{a_i} f(s-u) - \frac{\lambda}{a_i}\int_{b_z^{(1)}}^{s} g_i(u;s-y;b_z^{(1)},b_z^{(2)})dF(y)$$

$$i > 1; i \in I_z. \tag{3.19}$$

$$g_i(u;b_z^{(2)};b_z^{(1)},b_z^{(2)}) \overset{\text{def}}{=} 0 \qquad i > 1; i \in I_z. \tag{3.20}$$

Summarizing for strategies that dictate an increase of production in the sets $A_z^{(i,1)}$ with $i \in I_z$ the functional equations for $P_i(b_z^{(2)};s;b_z^{(1)})$, $g_1(u;s;b_z^{(1)})$ and $g_1(u;s;b_z^{(1)}, b_z^{(2)})$ have to be solved before (3.11) and (3.12) can be used to determine the functions $ck(z;i,s)$, $ct(z;i,s)$ and $P(y_j;i,s)$. All strategies that occured in the iteration cycles of the numerical examples in this paper did have this property and the relations (3.11) ... (3.14) could be solved by numerical integration methods. This has the advantage of obtaining a better accuracy within a shorter computing time, than simulation of (3.9) and (3.10). For strategies not having this property we will have to use simulation.

§ 4 The strategy - improvement routine.

When the function $c(z;x)$ is determined for a given strategy z then based on this function a better strategy can be determined. For that purpose we make use of the following definitions:

a) The function $c(d.z;x)$, given by:

$$c(d.z;x) \stackrel{\text{def}}{=} k(x;d) - r(z)\, t(x;d) + \mathcal{E}\{c(z;\underline{u})\,|\,d\} \tag{4.1}$$

This function results from applying the mixed strategy d.z. The prescription of this mixed strategy is to apply decision d in state x, which transforms the system into the stochastic state $\underline{u}$ and to apply strategy z after this transformation. In this problem u is deterministic.

b) The function $c(A.z;x)$, given by

$$c(A.z;x) = \int_A P_A^{(1)}(du;x)\, c(z;u) \tag{4.2}$$

The strategy A.z prescribes the postponement of decisions according to strategy z until the first future state $\underline{u}$, assumed in the set A.

c) The class K_z of all closed sets A satisfying

$$X = \{x \,|\, c(A.z;x) \leq c(z;x)\} \tag{4.3}$$

d) The set of states A_z' given by:

$$A_z' = {}_{A \in K_z} A \tag{4.4}$$

Suppose that at the n^{th} step of the iteration cycle we have obtained strategy $z^{(n)}$ and the function $c(z^{(n)};x)$. Then the following three steps should be performed.

1. Determine the function $c(d.z^{(n)};x)$ for each x and $d \in D(x)$ by (4.1)
2. Determine the decision $d^* \in D(x)$ for each x, satisfying:

$$c(d^*.z^{(n)};x) = \min_{d \in D(x)} c(d.z^{(n)};x) \tag{4.5}$$

The mixed strategy $d^*.z^{(n)}$ is denoted by $z_1^{(n)}$.

3. Determine the set A_{z_1}' satisfying (4.4)

The new strategy $z^{(n+1)}$ is then given by

$$z^{(n+1)}(x) = \begin{cases} z_1^{(n)}(x) & x \in A'_{z_1} \\ \text{null decision} & x \notin A'_{z_1} \end{cases} \tag{4.6}$$

It is proved in 1) that this iteration cycle leads to the optimal strategy z_o if there is only one ergodic set of states. The extension to more than one ergodic set of states is also given in 1). The optimal strategy z_o has the following properties:

$$\min_{d \in D(x)} c(d.z_o;x) = c(z_o;x) \tag{4.7}$$

$$A'_{z_o} = A_{z_o} \tag{4.8}$$

We shall now consider how in this particular problem the iteration cycle can be performed. Suppose we have obtained strategy $z^{(n)}$ and computed the function $c(z^{(n)};x)$ on a finite grid of states in the state space. Values of $c(z;x)$ between grid points can be determined by interpolation, if this function is continuous in x.

To begin with the first step we determine the function $c(d.z^{(n)};x)$ by means of its definition (4.1) for every $x = (i,s)$ and every $d = (j,s)$ with $i,j \varepsilon \{1,...,N\}$ and $0 \leq s \leq M$. Computationally this operation can also be performed only on a finite grid of states in the state space.

In the second step we determine for each state on the grid the decision d^{*} minimizing $c(d.z^{(n)};x)$.
If the minimizing decisions in two adjacent grid points are different then we determine the point between them where the values of $c(d.z^{(n)};x)$ for both minimizing decisions are equal. This new point seperates two sets of states with different minimizing decisions and will be called a seperation point. The possibility of another minimizing decision occuring on a part of the interval between two adjacent grid points can be investigated by taking a finer grid. The determination of these seperation points can be safely performed when the function $c(d.z^{(n)};x)$ is continuous in x for given d. In practice this condition is fulfilled except for the discontinuities of $c(z;x)$ in the points separating two intervals where different decisions are dictated by z.

If two adjacent grid points have the same minimizing decision then this decision is chosen for all states in the interval between these two grid points. The correctness of this procedure can again be verified by taking a finer grid.

When this operation is performed we have subdivided the state space into a finite number of intervals with the same minimizing decision for each state within such an interval. The separation points of these intervals specify completely the intermediate strategy $z_1^{(n)} = d^{*}.z^{(n)}$.

It should be noted that $A_{z_1}(n) \supset A_z(n)$ because a null-decision in $x \in A_z(n)$ is immediately followed by an intervention according to strategy z. Hence the null-decision can never be better than the decision z(x). For this reason the third step in the iteration cycle has to be performed.

In this third step strategies $A.z_1^{(n)}$ are considered, postponing decisions according to $z_1^{(n)}$, until a closed set $A \supset A_o$ is reached where A satisfies (4.3). The boundary of the smallest intersection of the class K_z of these sets A, denoted by $A'_{z_1(n)}$ being identical to $A_{z(n+1)}$, the intervention set of the new strategy $z^{(n+1)}$, can be determined by the observation that in its boundary points it should be indifferent either to postpone the decision according to $z_1^{(n)}$ or to apply strategy $z_1^{(n)}$ immediately. Because the sets $A_z^{(i,1)}$ are reached from states outside A_z in a different way than the sets $A_z^{(i,2)}$, this property leads to somewhat different criteria in these two cases.

To find the boundary $(i,b^{(1)})$ of each of the sets $A_{z(n+1)}^{(i,1)}$, we define the closed sets $A^{(1)}$ by the states (i,u) with $u \leq b^{(1)}$. In state $(i,b^{(1)})$ the effect of postponing the application of strategy $z_1^{(n)}$ is measured by the amount

$$\int_{(i,u)\in A^{(1)}} c(z_1^{(n)};i,u) \;\; d \;\; G_i\,(u;b^{(1)};b^{(1)},b^{(2)}_{z_1(n)})$$

$$+ \quad P_i\,(b^{(2)}_{z(n)};\, b^{(1)};b^{(1)})\; c(z_1^{(n)};i,b^{(2)}_{z_1}(n)) \qquad (4.8)$$

This amount should be compared for each $(i,b^{(1)}) \in A_{z_1}(n)$ with the result of applying strategy $z_1^{(n)}$ immediately, measured by $c(z_1(n)\;;i,b^{(1)})$.

The state $(i,b^{(1)})$ where both quantities are equal will be the boundary point of the set $A^{(i,1)}_{z(n+1)}$.

It may happen that there is more than one state where both quantities are equal. In that case there will be more than one set of non-intervention states for the considered production rate. An extension of the state space will be convenient in order to solve the functional equations for $c(z;x)$ for the new strategy by the methods described above.

To find the boundary point of each of the sets $A^{(i,2)}_{z(n+1)}$ let $A^{(2)}$ denote the set of states (i,u) with $u \geq b^{(2)}$. In state $(i,b^{(2)})$ the effect of postponing the decision $z_1(i,b^{(2)})$ will be that this decision takes place in state $(i,b^{(2)} + db^{(2)})$. This effect is measured by:

$$P_i(b^{(2)} + db^{(2)};\ b^{(2)};b^{(1)}_{z_1}(n))\ .c(z_1;i,b^{(2)} +db^{(2)})$$

$$+ \int_{(i,u\varepsilon A^{(i,1)}_{z_1}(n)} g_i(u;s;\ b^{(1)}_{z_1}(n)\ ,\ b^{(2)}+db^{(2)})\ c(z_1{}^{(n)};i,u)du \qquad (4.9)$$

The effect of applying strategy $z_1{}^{(n)}$ in $(i,b^{(2)})$ will be measured by $c(z_1{}^{(n)};\ i,b^{(2)})$. The state $(i,b^{(2)}$ where both amounts are equal, will be the extreme point $b^{(2)}_{z(n+1)}$ of the set $A^{(i,2)}_{z(n+1)}$ of the new strategy $z^{(n+1)}$.

This condition can be written as:

$$\left[\frac{\partial\ P_i(b^{(2)};s;b^{(1)}_{z_1})}{\partial\ b^{(2)}}\right]_{s=b^{(2)}} .c(z_1^{(n)};i,b^{(2)}) + \frac{\partial\ c(z_1^{(n)};i,b^{(2)})}{\partial\ b^{(2)}}$$

$$+\int_{(i,u)\varepsilon A^{(i,1)}_{z_1}(n)} \frac{\partial\ g_i(u;s;b^{(1)}_{z_1}\ ,\ b^{(2)})}{\partial\ b^{(2)}}\ c(z_1^{(n)};i,u)du = 0 \qquad (4.10)$$

If there is more than one state where relation (4.10) holds then we should choose the one with the smallest value of the c-function.

§ 5 Numerical example.

In order to determine optimal strategies for this problem numerically a computer program in ALGOL 60 has been developed. The results were obtained on the EL - X8 of the Mathematisch Centrum.

Data:

Order size $\underline{\lambda}$ exponentially distributed, $\mathcal{E}\underline{\lambda} = 5$; 3 production rates, $a_i = 0, 4, 8$.

Arrival rate of orders $\lambda = 1$;

Maximum stock $M = 20$;

Stockholding costs $c_s = 0, 5$;

Production costs per unit $c_p(i) = 0, 8, 16$;

Additional purchases per unit $c_r = 35$;

Transition costs $c_q(i,j)$ given by the matrix: $\begin{bmatrix} 0 & 5 & 10 \\ 5 & 0 & 5 \\ 10 & 5 & 0 \end{bmatrix}$

Strategies, occuring during the iteration, are given for each production rate i by the intervals for s where the same decision (j,s) is dictated.

Strategy:

	i \ j	3	2	1
$z^{(o)}$	1	-	$(-\infty,0]$	$(0,\infty)$
	2	-	$(-\infty,19.5)$	$[19.5,\infty)$
$r(z^{(o)}) = 67,47$	3	$(-\infty,18.15)$	$[18.15,19.5)$	$(19.5,\infty)$
$z_1^{(o)}$		$(-\infty, 14.96)$	$(14.96, 19.03)$	$(19.03,\infty)$
		$(-\infty, 14.96]$	$(14.96, 19.50)$	$(19.50,\infty)$
		$(-\infty, 17.69)$	$[17.69, 19.50)$	$(19.50,\infty)$

Strategy:	i \ j	3	2	1
$z^{(1)}$	1	(- ,2.12]	-	(2.12,∞)
	2	(-∞,2.05|	(2.05,20)	[20,∞)
$r(z^{(1)}) = 42,91$	3	(-∞,17.69)	[17.69,19.50)	(19.50,∞)
$z_1^{(1)}$		(-∞,14.42)	(14.42,18.33]	(18.33,∞)
		(-∞,14.42|	(14.42,20)	[20,∞)
		(-∞,19.54)	(19.54,20)	[20,∞)
$z^{(2)}$		(-∞,11.71]	-	(11.71,∞)
		(-∞,9.86]	(9.86,20)	[20,∞)
$r(z^{(2)}) = 38,77$		(-∞,19.77)	[19.77,20)	[20,∞)
$z_1^{(2)}$		(-∞,12.41)	(12.41,17.56|	(17.56,∞)
		(-∞,12.41]	(12.41,20)	[20,∞)
		(-∞,18.83)	[18.83,20)	[20,∞)
$z^{(3)}$		(-∞,12.41)	(12.41,16.25|	(16.25,∞)
		(-∞,12.06]	(12.06,20)	[20,∞)
$r(z^{(3)}) = 37,97$		(-∞,19.03)	[19.03,20)	[20,∞)
$z_1^{(3)}$		(-∞,12.52)	(12.52,17.00]	(17.00,∞)
		(-∞,12.52]	(12.52,20)	[20,∞)
		(-∞,19.30)	[19.30,20)	[20,∞)
$z^{(4)}$		(-∞,12.52)	(12.52,16.95|	(16.95,∞)
		(-∞,12.51]	(12.51,20)	[20,∞)
$r(z^{(4)}) = 37,93$		(-∞,19.68)	[19.68,20)	[20,∞)

Strategy:

	i \ j	3	2	1
$z_1^{(4)}$	1	(-∞,12.53)	(12.53,16.95[	(16.95,∞)
	2	(-∞,12.51]	(12.51,20)	[20,∞)
	3	(-∞,19.68)	[19.68,20)	[20,∞)
$z^{(5)}$		(-∞,12.53)	(12.53,16.95]	(16.95,∞)
		(-∞,12.51]	(12.51,20)	[20,∞)
$r(z^{(5)}) = 37{,}93$		(-∞,19.68)	[19.68,20)	[20,∞)

The computation time was 20 minutes.

Literature.

1) G. de Leve, Generalized Markovian Decision Processes Mathematical Centre Tract No. 3 and 4, 1964.

2) G. de Leve and P.J. Weeda, Driving with Markovprogramming. The ASTIN bulletin Vol. V, Part I, May 1968.

3) H.C. Tijms. A production problem solved by Markov-programmering Report S 396, Mathematical Centre 1968.