STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 403

Computing optimal (s,S) policies by means of Markov - programming

by

H.C. Tijms

november 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 the Advancement of Pure Research (Z.W.O.) and the Central Organization for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

1.1 Introduction.

A firm sells a single product. The stock is reviewed at the start of each period, at which time only an order may be placed. The length of any period is one unit of time. Between delivery and order we assume a lag of T periods, where T is a known nonnegative integer. When delivery is made it takes place at the start of a period. The order quantities are discrete. The firm can store at most M units. Let c(p) be the cost of ordering p units. For any unit the cost of storing it for t units of time is given by $c_1 t$.

Customers, who ask for the product, arrive according to a stationary Poisson process with intensity λ. The demands of the customers are mutually independent and each customer demands $k=(0,1,\ldots)$ units with probability p_k. We assume $p_0<1$, $\sum p_k=1$ and $\sum kp_k<\infty$.

When demand exceeds supply, the excess demand is backlogged until it is subsequently filled by a delivery. For any unit which is delivered subsequently $t>0$ units of time the penalty cost is given by $f(t)=c_2t+c_3$. No discounting will be used.

The stock on hand is the quantity which is stored (negative when a shortage exists). We define the <u>economic stock</u> as the stock on hand plus outstanding orders. When at the start of a period the economic stock is less than or equal to a known integer a, we have to replenish it at least until a known integer B. It is wellknown that both for an infinite planning horizon as a finite one the optimal policy depends only on the economic stock. For this reason we consider hereafter only policies which are functions of the economic stock.

An infinite planning horizon will be considered and the minimization of the expected mean cost per period will be used as criterion for selecting a policy. Under some general conditions it can be proved that an optimal policy exists which is of the (s,S)-type: When the economic stock $i \leq s$ ($s<S$), order S-i units; otherwise, do not order [6,7,10,11]. The wide use of ordering policies of this type in practice may be a reason for determining the best (s,S) policy, even when an optimal policy is not the (s,S)-type.

Markov-programming is a branch of dynamic programming. For an introductory treatment the reader is referred to the excellent book of R.A. Howard [4]. More advanced methods are given by W.S. Jewell [5], G. de Leve [8] and others. The simple properties of a (s,S) policy can be exploited by the methods given in [8]. A formula for the expected mean cost per period of a (s,S) policy can be easily derived. Furthermore an iterationprocedure can be formulated which enables us to determine a policy at least as good as the best (s,S) policy without in any iterationcycle a system of lineair equations is solved. A trick will be necessary to draw up such an iteration-procedure; a special class Z_0 of policies (containing the (s,S) policies) will be considered*).

Before we apply the theory of G. de Leve to our problem, we give some numerical results and we present a review of the methods developed in [8].

---

*) I am indebted to G. de Leve for this trick.

1.2 Numerical results

Let

$c_1 = 1$

$c(p) = cp + K\delta(p).$ *)

For different values of the other parameters an optimal policy and the corresponding average cost per unit of time reduced with $c\lambda \sum_{n=1}^{\infty} np_n$ are given in the tabels.

Tabel 1

$p_1 = 1$, $f(t) = 20t$, $c(p) = cp+4\delta(p)$.

λ \ T	0	1	2	3
3	(2,6),6.799	(6,11),8.144	(10,15),9.216	(13,19),10.114
4	(3,7),7.989	(8,13),9.468	(13,18),10.671	(17,23),11.693
5	(4,7),9.025	(10,15),10.670	(16,21),11.987	(21,26),13.208
6	(4,8),9.825	(12,16),11.703	(19,24),13.173	(25,31),14.406

Tabel 2

$p_1 = 1$, $f(t) = 20t$, $c(p) = cp$.

λ \ T	0	1	2	3
3	(3,4),3.969	(8,9),5.447	(12,13),6.625	(15,16),7.607
4	(5,6),4.717	(10,11),6.325	(15,16),7.645	(20,21),8.772
5	(6,7),5.388	(12,13),7.167	(18,19),8.627	(24,25),9.861
6	(7,8),6.040	(14,15),7.973	(21,22),9.568	(29,30),10.903

Tabel 3

$p_1 = 1$, $f(t) = 20$, $c(p) = cp+4\delta(p)$.

λ \ T	0	1	2	3
3	(4,8),8.499	(7,12),9.809	(11,16),10.740	(15,20),11.572
4	(5,9),9.972	(10,15),11.480	(15,20),12.659	(19,25),13.598
5	(6,10),11.272	(12,17),13.021	(18,24),14.348	(29,30),15.451
6	(7,11),12.393	(14,19),14.415	(21,27),15.923	(28,34),17.137

*) $\delta(p) = 1$ for $p \geq 1$ and $\delta(0) = 0$.

Tabel 4

$p_1 = 1,\ f(t) = 20,\ c(p) = cp$

T / λ	0	1	2	3
3	(5,6),5.522	(9,10),7.058	(13,14),8.166	(17,18),9.087
4	(7,8),6.677	(12,13),8.333	(17,18),9.643	(22,23),10.736
5	(8,9),7.587	(14,15),9.589	(21,22),11.033	(27,28),12.258
6	(9,10),8.556	(17,18),10.657	(24,25),12.206	(31,32),13.636

Tabel 5

$p_0=0.5,\ p_1=0.1,\ p_2=0.3,\ p_3=0.1,\ f(t)=20t,\ c(p)=cp+4\delta p$

T / λ	0	1	2	3
3	(3,7),8.120	(7,12),10.456	(12,17),12.162	(16,21),13.608
4	(4,8),9.408	(10,15),12.003	(15,21),13.950	(20,26),15.601
5	(5,9),10.561	(12,17),13.378	(18,24),15.539	(24,31),17.386
6	(5,10),11.585	(14,19),14.637	(21,28),16.997	(29,36),18.996

Tabel 6

$p_0=0.5,\ p_1=0.1,\ p_2=0.3,\ p_3=0.1,\ f(t)=20t,\ c(p)=cp.$

T / λ	0	1	2	3
3	(5,6),5.752	(10,11),8.185	(14,15),9.982	(18,19),11.509
4	(6,7),6.631	(12,13),9.367	(18,19),11.447	(24,25),13.265
5	(7,8),7.474	(14,15),10.500	(21,22),12.760	(28,29),14.662
6	(8,9),8.280	(17,18),11.491	(25,26),13.961	(33,34),16.068

Tabel 7

$p_0=0.5$, $p_1=0.1$, $p_2=0.3$, $p_3=0.1$, $f(t)=20$, $c(p)=cp+4\delta(p)$

λ \ T	0	1	2	3
3	(5,9),10.337	(9,14),12.367	(13,19),13.811	(17,23),14.974
4	(6,11),12.049	(12,17),14.424	(17,23),16.142	(22,28),17.540
5	(8,12),13.611	(14,20),16.272	(21,27),18.251	(27,33),19.861
6	(9,14),14.987	(17,23),17.962	(24,31),20.163	(31,38),21.969

Tabel 8

$p_0=0.5$, $p_1=0.1$, $p_2=0.3$, $p_3=0.1$, $f(t)=20$, $c(p)=cp$.

λ \ T	0	1	2	3
3	(7,8),7.966	(12,13),10.197	(16,17),11.723	(19,20),12.965
4	(8,9),9.356	(14,15),11.890	(20,21),13.697	(25,26),15.177
5	(10,11),10.533	(17,18),13.412	(24,25),15.498	(30,31),17,200
6	(12,13),11.732	(20,21),14.838	(28,29),17.169	(35.36),19.071

2. Markov-programming.

Problems of the type to which Markov-programming can be applied are always related to some physical system. In our case the system comprises the stock on hand and the quantity on the books.
At each point of time the system is in some state x. In the mathematical model a state of the system is represented by a point in a finite dimensional Cartesian space. The set of all possible states will be called the state space X.
Besides deterministic transformations the state of the system may be subjected to random transitions. Owing to the latter transitions the system performs a random walk through the state space. In case no decisions are made, this evolution is called the natural process.
A condition for application of Markov-programming is that for each initial state of the system the underlying natural process can be described by a stationary strong Markov-process in X.
A family of n-dimensional random vectors*) $\{\underline{x}_t, t \in T\}$ is called a Markov-process, if with probability one,

$$P\{\underline{x}_{t+s} \leq x \mid \underline{x}_u, u \leq t\} = P\{\underline{x}_{t+s} \leq x \mid \underline{x}_t\} \qquad (2.1)$$

for each $x \in R^n$ and every $s, t, \in T$, $s > 0$.
Roughly speaking: If we know the "present" then the additional knowledge of the "past" does not contribute any relevant information about the "future".
The term a stationary Markov-processs will be used, if the probability distribution of (2.1) does not depend on t. If the foregoing also holds when the arbitrary but fixed time t is replaced by a random variable $\underline{\tau}$, which satisfies certain regularity conditions which are given in [8], the process is called a stationary strong Markov-process.

---

*) Measurable functions on a probability space, which assume their values in a n-dimensional Cartesian space. Random vectors, called random variables if n = 1, are underlined.

In addition the following definitions are given with regard to a Markov-process in state space X with some time parameter. A subset S of X is called ergodic if the system remains with probability*) one in S as soon as it has assumed a state of S. A ergodic set is called simple ergodic, if it contains no disjunct ergodic sets. The set T of states, which does not belong to any set from a given system of simple ergodic sets is called the set of transient states, if T does not contain an ergodic set. A decomposition of the state space into simple ergodic sets and a transient set is not always unique. In this paper it is assumed that always a decomposition is given with disjunct simple ergodic sets. We note that if the state space is finite or denumerable a decomposition can be given, such that an ergodic set S is a simple ergodic set, if every state in S can be reached from every other state in S. The simple ergodic sets are in this case always disjunct.

In decision problems losses and gains play important roles. It is no restriction to consider only losses (gains are negative losses). In general the decisionmaker wants to influence the natural process by interventions, basically a finite number in each finite time interval. An intervention causes a transition in the state of the system. A transition is assumed to take no time. The behaviour of the system in each time interval between two successive interventions is described by a natural process. The initial state of that process will be the state into which the system is transferred by the intervention at the beginning of the interval concerned. For that reason for each initial state the natural process has to be defined. It is convenient to assume that at each point of time a decision is made. The decision will be primarily to decide whether to intervene or not and secondly which intervention to choose. The decision not to intervene is called a null-decision. In many situations decisions result in a random transition in the state of the system. For that reason a decision is defined mathematically by means of the probability distribution of the state into which the system may be transferred by the decision. By a null-decision the system is "transferred" with probability one in its present state. Decisions which lead to deterministic transitions are also defined by "concentrated" probability distributions, but now in the new state.

*) Suppose that this probability is defined.

To each state $x \in X$ a set of feasible decisions $D(x)$ is assigned. The solution of a decision problem is given in the form of a strategy. Such a strategy dictates at each point of time a feasible decision on the basis of available information. The result of the natural process and the extra transitions caused by the strategy is called the decisionprocess. Let Z be the class of strategies z, which base their decisions on the present state only and add to each state x a feasible decision $d = z(x)$. Since we have only interventions and null-decisions each strategy $z \in Z$ partitions the state space into two disjunct sets, one denoted by A_z, comprising the states in which always interventions are made, the other consisting of states in which always null-decisions are dictated.

From now on only strategies $z \in Z$ are considered. Under some general conditions it can be shown this is no restriction. Further it can be proved under certain weak conditions the decision process corresponding to a strategy of Z is also a stationary strong Markov-process.

In order to find out which strategy is the best one we need a criterion. As criterion for an optimal strategy we shall adopt the expected mean costs per unit of time, when the system is considered for an infinite period of time.

Suppose the intersection

$$A_0 \stackrel{\text{def}}{=} \bigcap_{z \in Z} A_z \qquad (2.2)$$

is not empty. Assume that in the natural process from each initial state the set A_0 can be reached within a finite time with probability one. Note that each strategy of Z dictates an intervention in any state of A_0.

Choose the sets

$$A_{0,i} \subset A_0 \quad (i = 1,2), \qquad (2.3)$$

such that the sets $A_{0,1}$ and $A_{0,2}$ are not empty and they can be reached in the natural process from each initial state within a finite time with probability one.

For each $i = 1,2$ there corresponds to every state x and decision $d \in D(x)$ two random walks $\underline{w}^{0,i}$ and $\underline{w}^{d,i}$. The walk $\underline{w}^{0,1}$ ($\underline{w}^{0,2}$) has x as initial state and during this walk the system is subjected to the natural process. The walk $\underline{w}^{0,1}$ ($\underline{w}^{0,2}$) ends as soon as the system assumes a state of $A_{0,1}$ ($A_{0,2}$). The walk $\underline{w}^{d,1}$ ($\underline{w}^{d,2}$) has x also as initial state. In state x decision d is made, by which the system is transferred (instantaneously!) into a random state and from this state on the system is subjected to the natural process. The walk $\underline{w}^{d,1}$ ($\underline{w}^{d,2}$) ends as soon as the system assumes a state of $A_{0,1}$ ($A_{0,2}$).
Let the functions $k_0(x)$ and $k_1(x;d)$ represent the expected costs incurred during $\underline{w}^{0,1}$ and $\underline{w}^{d,1}$ respectively. Let the functions $t_0(x)$ and $t_1(x;d)$ be equal to the expected durations of $\underline{w}^{0,2}$ and $\underline{w}^{d,2}$ respectively.
We define,

$$k(x;d) = k_1(x;d) - k_0(x) \qquad (2.4)$$

and

$$t(x;d) = t_1(x;d) - t_0(x). \qquad (2.5)$$

Note that for d = null-decision $\underline{w}^{d,i}$ and $\underline{w}^{0,i}$ are identical, and consequently

$$k(x;d) = t(x;d) = 0, \quad d = \text{null-decision}. \qquad (2.6)$$

It follows from their definitions that k(x;d) and t(x;d) do not depend on any particular strategy. Hence we need only <u>once</u> for all to determine*) the (x;d)-function k(x;d) and t(x;d)

Let $\{\underline{I}_n\}_{n=1}^{\infty}$ be the sequence of future interventionstates if

---

*) Sometimes it can be advantageous to alter the natural process as soon as the system assumes a state of A_0. By doing so and/or by a proper choice of $A_{0,1}$ and $A_{0,2}$ the determination of the k-and t-functions may be simplified greatly. It can be shown that it is allowed to change the natural process as soon as the system assumes a state of A_0. Note that situations caused by such a change does not occur in reality, because in each state of A_0 the decision-maker has to intervene.

strategy $z \in Z$ is applied. The sequence $\{\underline{I}_n, n \geq 1\}$ constitutes a stationary Markov-process in A_z with discrete time parameter. The probability that $\underline{I}_n$ belongs to some Borelset A if the system has $x \in X$ as initial state, will be denoted by

$$p^{(n)}(A;z;x) \ , \ n = 1,2,.. \tag{2.7}$$

Under some general conditions it can be shown the Markov-process $\{\underline{I}_n, n \geq 1\}$ has a stationary probability distribution $\phi(A;z;x)$ (roughly speaking: the distribution of $\underline{I}_n$ for $n \to \infty$), that satisfies*),

$$\phi(A;z;x) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} p^{(k)}(A;z;x). \tag{2.8}$$

If x_1 and x_2 belong to a same simple ergodic set,

$$\phi(A;z;x_1) = \phi(A;z;x_2). \tag{2.9}$$

If $z \in Z$ is the strategy applied and if the decisionprocess has x as initial state, let $\underline{k}_T(z;x)$ be the costs incurred during the period $[0,T)$. Under certain conditions it can be proved that,

$$\lim_{T \to \infty} \frac{\underline{k}_T(z;x)}{T} \tag{2.10}$$

exists with probability one. Note that this limit represents the random mean costs per unit of time.

If the initial state x belongs to a simple ergodic set, it can be proved that with probability one.

$$\lim_{T \to \infty} \frac{\underline{k}_T(z;x)}{T} = \frac{\int_{A_z} k(I;z(I))\phi(dI;z;x)}{\int_{A_z} t(I;z(I))\phi(dI;z;x)} , \tag{2.11}$$

From (2.9) it follows that the right-hand member of (2.11) is constant on a simple ergodic set.

---

*) The Césarolimit $\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} a_k$ exists and is equal to a, if $\lim_{n \to \infty} a_n = a$. The converse is not always true.

Let for ergodic (= non-transient) states x the function r(z;x) be defined by,

$$r(z;x) \overset{\text{def}}{=} \frac{\int_{A_z} k(I;z(I))\phi(dI;z;x)}{\int_{A_z} t(I;z(I))\phi(dI;z;x)} . \qquad (2.12)$$

The domain of definition of r(z;x) is extended to the whole state space X by

$$r(z;x) \overset{\text{def}}{=} Er(z;\underline{y}). \qquad (2.13)$$

where $\underline{y}$ is the first ergodic state taken on in the decision process if x is the initial state.

Note, by (2.11) and (2.9), that the mean costs per unit of time are constant with probability one if the initial state is ergodic, further r(z;x) is constant on a simple ergodic set.

However if the initial state is transient the mean costs depend on the first ergodic state assumed, and therefore, they are random.

Hence, by (2.13), the function r(z;x) determines the expected mean costs per unit of time for all initial states x.

The function r(z;x) (the criterion!) may be determined without calculating the stationary probability distribution of $\underline{I}_n$. A function c(z;x) can be introduced which in a sense enables us to value the initial state x with respect to the total expected costs.

The x-functions c(z;x) and r(z;x) jointly satisfy a system of functional equations and they can be used in an iterationprocedure for obtaining optimal strategies.

Suppose the Markov-process $\{\underline{I}_n, n \geq 1\}$ in A_z has m disjunct simple ergodic sets E. Choose in each set E_j an arbitrary state e_j.

Consider next the following functional equations in r(z;x) and c(z;x):

$$r(z;x) = Er(z;\underline{I}_1) \qquad (2.14)$$

and

$$c(z;x) = k(x;z(x)) - r(z;x)t(x;z(x)) + Ec(z;\underline{I}_1) \qquad (2.15)$$

$$c(z;e_j) = 0, \quad j = 1, \ldots, m, \qquad (2.16)$$

where $\underline{I}_1$ is the first future interventionstate if x is the initial state and strategy z is applied*).

*) $Er(z;\underline{I}_1) = \int_{A_z} r(z;I)p^{(1)}(dI;z;x)$. The same holds for $Ec(z;\underline{I}_1)$.

Note that from (2.6) it follows

$$c(z;x) = Ec(z;\underline{I}_1) \text{ , if } x \notin A_z.$$

The function $r(z;x)$ is constant on a simple ergodic set, hence $r(z;x)$ indicates the most favourable simple ergodic set to start, but not the most profitable initial state in this set.
That state can be determined by means of the function $c(z;x)$.
It can be shown that for two states x_1 and x_2 in the same simple ergodic set the difference in total expected costs is finite and is given by

$$c(z;x_1) - c(z;x_2). \qquad (2.18)$$

By means of the functions $r(z;x)$ and $c(z;x)$ the strategy z can be improved. An iterationprocedure can be given, which yields a sequence of strategies $\{z^{(i)}, i = 1,2,..\}$ of which, under certain conditions, the following interesting properties can be proved:

$$\text{a)} \quad r(z^{(i)};x) \geq r(z^{(i+1)};x) \qquad (2.19)$$

$$\text{b)} \quad \lim_{i \to \infty} r(z^{(i)};x) = \min_{z \in Z} r(z;x), \qquad (2.20)$$

for each $x \in X$. Proofs and conditions are given in [8] and will be omitted here. We shall restrict ourselves to an intuitive explanation of the procedure. First some introductory definitions.

Let the mixed strategy d.z with $z \in Z$ dictate the decision d in the initial state and then decisions in accordance with z. We define the functions $r(d.z;x)$ and $c(d.z;x)$ by

$$r(d.z;x) = Er(z;\underline{u}) \qquad (2.21)$$

and

$$c(d.z;x) = k(x;d) - r(d.z;x)t(x;d) + Ec(z;\underline{u}), \qquad (2.22)$$

where $\underline{u}$ is the random state in which the system is transferred (instantaneously) by the decision d in the initial state x. From the definitions and (2.6) it follows that for both null-decision and $d = z(x)$,

$$r(d.z;x) = r(z;x) \text{ and } c(d.z;x) = c(z;x), \qquad (2.23)$$

Consider now the following problem. Suppose a decisionmaker has to make his decisions in accordance with a strategy z. In the initial state however he is free to choose a feasible decision. The decisionmaker certainly looks for that particular decision, such that the expected mean cost per unit of time is minimized. Each drop in this cost leads to an infinite saving in an infinite period of time. If in the initial state x the feasible decision d is chosen and thereafter strategy z is applied, the expected mean cost per unit of time is given by $r(d.z;x)$. Hence we determine for each state x:

$$\min_{d \in D(x)} r(d.z;x) . \qquad (2.24)$$

Let $D_z(x)$ be the set of minimizing decisions $d \in D(x)$. In order to determine which d has to be chosen if $D_z(x)$ contains more than one decision, we note that it can be shown that the difference in total expected cost of the mixed strategy d.z and the strategy z is given by

$$c(d.z;x) - c(z;x) . \qquad (2.25)$$

The difference will be in general finite. It will now be obvious that in case $D_z(x)$ contains more than one decision, we determine

$$\min_{d \in D_z(x)} c(d.z;x) . \qquad (2.26)$$

We add now to each state x a decision of $D_z(x)$ which minimizes $c(d.z;x)$. (if $d = z(x) \in D_z(x)$ and it minimizes $c(d.z;x)$ we choose $d = z(x)$). By this procedure a (possible new) decision is added to each state. We have then constructed a new strategy z_1. The following important result can now be proved,

$$r(z_1;x) \leq r(z;x) \qquad x \in X. \qquad (2.27)$$

Hence the strategy z_1 is at least as good as the strategy z. However from (2.23) it follows easily that each intervention state of z is an interventionstate of z_1 too, hence A_{z_1} encloses A_z. It will be obvious that we need a mechanism which may cancel an intervention. With the aid of the foregoing we shall now determine a strategy z_2 which is at least as good as z and with the property that A_{z_2} is contained in A_{z_1}. Let strategy $z \in Z$ be given and let strategy z_1 be determined in accordance with (2.24) and (2.26). We now introduce mixed strategies of the following type:

(a) The mixed strategy $(z_1)z$ dictating

1) first an intervention in accordance with z_1
2) then interventions in accordance with z.

Put for abbreviation $\hat{z} = (z_1)z$. We define the x-functions $r(\hat{z};x)$ and $c(\hat{z};x)$ for $x \in A_{z_1}$ by,

$$r(\hat{z};x) = \min_{d \in D(x)} r(d.z;x) \tag{2.28}$$

$$c(\hat{z};x) = \min_{d \in D_z(x)} c(d.z;x) . \tag{2.29}$$

(b) The mixed strategy $A.\hat{z}$, where A is a closed set satisfying

$$A_0 \subset A \subset A_{z_1} . \tag{2.30}$$

This strategy interdicts any intervention up to the moment that the system assumes a state of A for the first time. From that time onwards the mixed strategy $\hat{z}$ is applied.

We define the x-functions $r(A.\hat{z};x)$ and $c(A.\hat{z};x)$ by,

$$r(A.\hat{z};x) = Er(\hat{z};\underline{v}) \tag{2.31}$$

$$c(A.\hat{z};x) = Ec(\hat{z};\underline{v}) , \tag{2.32}$$

where $\underline{v}$ is the first state in A taken on if x is initial state and the mixed strategy $A.\hat{z}$ is applied. We note that the probability distribution of $\underline{v}$ is determined only by the natural process and the set A.

Consider the following problem. Suppose the decisionmaker has to make his decisions in accordance with the mixed strategy $\hat{z}$. But he is allowed to determine the point of time where upon $\hat{z}$ comes into operation. This will be done by choosing a closed set A satisfying (2.30). The mixed strategy $\hat{z}$ comes into operation at the moment the system assumes a state of A. The expected mean cost per unit of time of the mixed strategy $A.\hat{z}$ is given by $r(A.\hat{z};x)$, while $c(A.\hat{z};x) - c(\hat{z};x)$ measures the difference in total expected cost of $A.\hat{z}$ and $\hat{z}$.

Let $X(\hat{z})$ be the class of all closed sets A satisfying (2.32) which have the additional property that for each $x \in A_{z_1}$ either

$$r(A.\hat{z};x) < r(\hat{z};x) \tag{2.33}$$

or

$$r(A.\hat{z};x) = r(\hat{z};x) \quad , \quad c(A.\hat{z};x) \leq c(\hat{z};x) \ . \tag{2.34}$$

It is easily verified that for each state $x \in X$ either (2.33) or (2.34) holds if $A \in X(\hat{z})$. It can be proved that the intersection of any finite number of sets of $X(\hat{z})$ belongs to $X(\hat{z})$.

Let

$$A_{\hat{z}}' = \bigcap_{A \in X(\hat{z})} A \ . \tag{2.35}$$

If $A_{\hat{z}}' \in X(\hat{z})$ it can be shown that the strategy

$$z_2(x) \overset{\text{def}}{=} \begin{cases} z_1(x) & x \in A_{\hat{z}}' \\ \text{null-decision} & \text{otherwise} \end{cases} \tag{2.36}$$

is at least as good as the original strategy z.

From the definitions it follows that relation (2.34) with the equality signs holds for $A = A_{z_1}$, hence $A_{\hat{z}}' \subset A_{z_1}$, in other words A_{z_2} is contained in A_{z_1}.

From the foregoing it can be easily deduced that a strategy $z^* \in Z$ is optimal if it possesses the following properties

$$\min_{d \in D(x)} r(d.z^*;x) = r(z^*;x) \qquad x \in X \tag{2.37}$$

$$\min_{d \in D_{z^*}(x)} c(d.z^*;x) = c(z^*;x) \qquad x \in X \tag{2.38}$$

$$A_{\hat{z}^*}' = A_{z^*}. \tag{2.39}$$

These formulas present us a direct approach, with which an optimal strategy can be determined.

In the most cases an optimal strategy will be determined by means of an iterationprocedure. Several formulations are possible [8], we shall give one.

Prepatory part.

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

Iterative approach

Let $z^{(n-1)}$ be the strategy obtained at the $(n-1)^{st}$ cycle, the i^{th} cycle runs as follows:

1) Determine the functions $r(z^{(n-1)};x)$ and $c(z^{(n-1)};x)$ by solving the functional equations (2.14), (2.15) and (2.16).
2) a) Determine the functions $r(d.z^{(n-1)};x)$ and $c(d.z^{(n-1)};x)$ by using the relations (2.21) and (2.22).
   b) Determine for each $x \in X$ the subset $D_{z^{(n-1)}}(x)$ of decisions $d \in D(x)$, which minimize $r(d.z^{(n-1)};x)$.
   c) Minimize for each $x \in X$ the d-function $c(d.z^{(n-1)};x)$ subject to $d \in D_{z^{(n-1)}}(x)$.
   d) Add to each state x a solution of c). If $z^{(n-1)}(x)$ is a solution of c), this decision will be added to state x.
   [This instruction has been made in order to advance the convergence of the sequence of strategies $\{z^{(i)}, i \geq 1\}$].

   As soon as operation d) has been performed a new strategy $z_1^{(n-1)}$ has been constructed.
3) Determine $r(\hat{z}^{(n-1)};x)$ and $c(\hat{z}^{(n-1)};x)$ for $x \in A_{z_1}$ [c.f (2.28) and (2.29)].
4) Determine the set $A'_{\hat{z}^{(n-1)}}$ [c.f (2.40)]. The new strategy $z^{(n)}$ is given by

$$z^{(n)}(x) = \begin{cases} z_1^{(n-1)}(x) & , \text{ if } x \in A'_{\hat{z}^{(n-1)}} \\ \text{null-decision}, & \text{otherwise}, \end{cases} \qquad (2.40)$$

End n^{th} cycle.

An optimal strategy has been reached if the strategies in two successive cycles are identical.

Notes.

a) The functions $r(z;x)$ and $c(z;x)$ are determined by functional equations. If these equations cannot be solved analytically they often can be solved numerically by Monte Carlo methods.
Sometimes it is much easier to determine the k- and t-functions numerically by simulation, in stead of to determine their values from the analytical formulas.

b) The way in which the set $A'_{\hat{z}}$ can be determined depends heavily on the structure of the decision problem considered.
In the boundary points of $A'_{\hat{z}}$ it will sometimes be indifferent to intervene or not. This property may enable us to construct $A'_{\hat{z}}$.

c) Computations may be reduced considerably, when it is realized that on a simple ergodic set S the function $r(z;x)$ is constant, say $r(z)$, and

$$r(d.z;x) = r(A.\hat{z};x) = r(z;x) = r(z) \;, \quad x \in S. \tag{2.41}$$

3. Preliminaries.

Suppose customers arrive at a firm according to a stationary Poisson process $\{\underline{w}(t), t \geq 0\}$ with intensity λ and they ask for a single product. Each customer demands k units with probability p_k where $\sum_{k=0}^{\infty} p_k = 1$, $\sum_{k=0}^{\infty} kp_k < \infty$ and $p_0 < 1$. Let the demands of the customers be mutually independent.

Three important properties of the Poisson process are

(a) the number of arrivals in any time interval $(t,t+h]$ has a Poisson distribution with mean λh $(h > 0)$. Hence,

$$P\{\underline{w}(t+h)-\underline{w}(t)=n\} = e^{-\lambda h}\frac{(\lambda h)^n}{n!}, \quad n=0,1,\ldots \tag{3.1}$$

(b) the interval from 0 up to the first arrival and thereafter the intervals between two successive arrivals, are independently distributed with common exponential density $\lambda e^{-\lambda t}$.

(c) given an arbitrary but fixed point of time, the waiting time to the first future arrival has an exponential density with parameter λ, irrespective of the past.

Let

$$\underline{v}(t) \overset{\text{def}}{=} \text{number of units demanded in the interval } (0,t], \; t > 0. \tag{3.2}$$

Define

$$\underline{v}(0)=0. \tag{3.3}$$

From the properties of the Poisson process and the assumed independence of the demands of the customers it follows

(1) for any $t,s > 0$ the random variable $\underline{v}(t+s)-\underline{v}(t)$ and $\underline{v}(s)$ are identically distributed.

(2) if $0 \leq t_1 < t_2 < \ldots < t_n$ $(n \geq 3)$ the differences

$$\underline{v}(t_2)-\underline{v}(t_1), \ldots, \underline{v}(t_n)-\underline{v}(t_{n-1}) \tag{3.4}$$

are mutually independent.

(Hence the numbers of units demanded in disjunct intervals are mutually independent).

From the theory of generating functions[3] it follows

$$\sum_{k=0}^{\infty} P\{\underline{v}(t)=k\}\, s^k = e^{-\lambda t(1-\sum_{n=0}^{\infty} p_n s^n)} \tag{3.5}$$

and

$$E\underline{v}(t) = \lambda t \sum_{n=1}^{\infty} np_n. \tag{3.6}$$

If $p_i=0$ for $i > N$ the power-series expansion of (3.5) enables us to write down the coefficients $a_n(t)$, where

$$a_n(t)=P\{\underline{v}(t)=n\} \ , \ n=0,1, \ldots \tag{3.7}$$

Theorem 1

Let $p_i=0$ for $i > N$ then,

$$a_n(t) = e^{-\lambda t(1-p_0)} \frac{\{\lambda t(1-p_0)\}^n}{n!} \qquad \text{for } N=1, \tag{3.8}$$

and for $N > 1$,

$$a_n(t) = e^{-\lambda t(1-p_0)} \sum_{j_1=0}^{\left[\frac{n}{N}\right]} \cdots \sum_{j_{N-1}=0}^{\left[\frac{n-\sum_{k=1}^{N-2}(N-k+1)j_k}{2}\right]}$$

$$\left\{ \prod_{k=1}^{N-1} \frac{(\lambda t p_{N-k+1})^{j_k}}{j_k!} \; \frac{(\lambda t p_1)^{n-\sum_{k=1}^{N-1}(N-k+1)j_k}}{(n-\sum_{k=1}^{N-1}(N-k+1)j_k)!} \right\}. \tag{3.9}$$

Where $[x]$ is the largest integer less than or equal to x.

Let $\underline{t}_k$ be the interval from 0 up to the epoch on which the k^{th} unit is demanded, mathematically

$$\underline{t}_k = \inf\{t \,|\, \underline{v}(t) \geq k\} \ , \ k \geq 1 . \tag{3.10}$$

Obviously,

$$P\{\underline{t}_k < t\} = P\{\underline{t}_k \leq t\} . \tag{3.11}$$

From the definition it follows,

$$P\{\underline{t}_k \leq t\} = P\{\underline{v}(t) \geq k\} = \sum_{j=k}^{\infty} a_j(t). \tag{3.12}$$

Next a number of theorems will be proved and in the proofs the well-known theorem of total expectation,

$$E\underline{x} = \int_{-\infty}^{+\infty} E(\underline{x}|\underline{y}=y)dP\{\underline{y} \leq y\} , \tag{3.13}$$

will be frequently used.

Theorem 2

$$E\underline{t}_k = \sum_{i=0}^{k-1} p_i \, E\underline{t}_{k-i} + \frac{1}{\lambda} , \; k \geq 1. \tag{3.14}$$

Proof Let $\underline{u}$ be equal to the waiting time to the arrival of the first customer. Under the condition that the first customer demands i units, where $0 \leq i \leq k-1$, $\underline{t}_k$ is equal to $\underline{u}+\underline{t}_{k-i}$. $\underline{t}_{k-i}$ and $\underline{u}$ are mutually independent. Given that the first customer demands more than k-1 units $\underline{t}_k$ is equal to $\underline{u}$. Obviously $\underline{u}$ has an exponential distribution with parameter λ. By applying (3.13) the theorem is proved.

Generalisation

$$E\underline{t}_k^{\,m} = \sum_{i=0}^{k-1} p_i \sum_{s=0}^{m} \binom{m}{s} E\underline{u}^{m-s} \, E\underline{t}_{k-i}^{\,s} + \sum_{i=k}^{\infty} p_i \, E\underline{u}^{m}. \tag{3.15}$$

Theorem 3

$$E\underline{v}_k = \sum_{i=0}^{k-1} p_i \, E\underline{v}_{k-i} + \sum_{i=0}^{\infty} i p_i \tag{3.16}$$

$$\text{where } \underline{v}_k \overset{\text{def}}{=} \text{number of units demanded in } (0,\underline{t}_k]. \tag{3.17}$$

Proof

Use the same arguments as in theorem 1.

Generalisation

$$E\underline{v}_k{}^m = \sum_{i=0}^{k-1} p_i \, E(\underline{v}_{k-i} + i)^m + \sum_{i=k}^{\infty} p_i \, i^m. \qquad (3.18)$$

Theorem 4

$$E\underline{v}_k = \lambda E\underline{t}_k \sum_{n=1}^{\infty} np_n. \qquad (3.19)$$

Proof By induction.

Theorem 5

$$E\{(\underline{t}_k - t)\, l\, (\underline{t}_k - t)\} = \sum_{j=0}^{k-1} a_j(t) E\underline{t}_{k-j} \qquad (3.20)$$

where

$$l(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases} \qquad (3.21)$$

Proof Given an arbitrary but fixed point of time the waiting time to the arrival of the first future customer has the same distribution as the interval between two successive arrivals, irrespective of the past. Hence under the condition that in $(0,t]$ i units are demanded the expectation $E\{(\underline{t}_k-t)l(\underline{t}_k-t)\}$ is equal to $E\underline{t}_{k-i}$, where

$$\underline{t}_n = 0 \quad \text{if } n \leq 0 . \qquad (3.22)$$

The theorem is proved by applying the theorem of total expectation.

Theorem 6

$$E\{(t-\underline{t}_k)\, l\, (t-\underline{t}_k)\} = \sum_{j=0}^{k-1} a_j(t) E\underline{t}_{k-j} - E\underline{t}_k + t. \qquad (3.23)$$

Proof

$$E(\underline{t}_k - t) = E\{(\underline{t}_k-t)\, l\, (\underline{t}_k-t)\} + E\{(\underline{t}_k-t)\, l\, (t-\underline{t}_k)\}. \qquad (3.24)$$

Theorem 7

$$\sum_{k=1}^{\infty} E\{(t-\underline{t}_k)\, l\, (t-\underline{t}_k)\} = \frac{\lambda}{2} t^2 \sum_{n=1}^{\infty} np_n. \qquad (3.25)$$

Proof

$$\sum_{k=1}^{\infty} P\{\underline{t}_k \leq u\} = \sum_{k=1}^{\infty} P\{\underline{v}(u) \geq k\} = \sum_{k=0}^{\infty} P\{\underline{v}(u) > k\}. \qquad (3.26)$$

For each integral valued random variable $\underline{a} \geq 0$ holds [3]

$$E\,\underline{a} = \sum_{n=0}^{\infty} P\{\underline{a} > n\}. \qquad (3.27)$$

Hence

$$\sum_{k=1}^{\infty} P\{\underline{t}_k \leq u\} = E\,\underline{v}(u) = \lambda\, u \sum_{n=1}^{\infty} np_n\,. \qquad (3.28)$$

The theorem follows now from (3.28) and the relation

$$E\{(t-\underline{t}_k) \cdot (t-\underline{t}_k)\} = \int_0^t (t-u)\, d\, P\{\underline{t}_k \leq u\}. \qquad (3.29)$$

Let

$$\underline{t}_k^{*} \overset{\text{def}}{=} [\underline{t}_k] + 1 \quad , \quad k \geq 1 \qquad (3.30)$$

and

$$\underline{v}_k^{*} = \text{number of units demanded in } (0, \underline{t}_k^{*}]\,. \qquad (3.31)$$

Theorem 8

$$E\,\underline{t}_k^{*} = \sum_{i=0}^{k-1} a_i(1)\, E\,\underline{t}_{k-i}^{*} + 1 \qquad (3.32)$$

and

$$E\,\underline{v}_k^{*} = \sum_{i=0}^{k-1} a_i(1)\, E\,\underline{v}_{k-i}^{*} + \lambda \sum_{n=1}^{\infty} np_n\,. \qquad (3.33)$$

Proof

The interval $(0, \underline{t}_k^{*}]$ encloses the interval $(0,1]$. Given that in $(0,1]$ there are demanded i units with $0 \leq i < k$, $\underline{t}_k^{*}$ (respectively $\underline{v}_k^{*}$) has the same distribution as $1 + \underline{t}_{k-i}^{*}$ (respectively $i + \underline{v}_{k-i}^{*}$). Under the condition that the number of units demanded in $(0,1]$ exceeds k-1 then $\underline{t}_k^{*}$ (respectively $\underline{v}_k^{*}$) is equal to 1 (respectively i). By applying the theorem of total expectation theorem 8 is proved. (Note $\sum_{i=1}^{\infty} i\, a_i(1) = E\,\underline{v}(1) = \lambda \sum_{n=1}^{\infty} np_n$).

Theorem 9

$$E\,\underline{v}_k^{*} = \lambda\, E\,\underline{t}_k^{*} \sum_{n=1}^{\infty} np_n . \tag{3.34}$$

Proof By induction.

We introduce for $j \geq 0$,

$$u(j) = \sum_{n=0}^{\infty} a_j(n). \tag{3.35}$$

We note that $a_0(0) = 1$ and $a_j(0) = 0$ for $j>0$. Furthermore it follows from (3.5) that $a_0(t) = e^{-\lambda\, t(1-p_0)}$. Hence

$$u(0) = 1/(1-e^{-\lambda\,(1-p_0)}) \tag{3.36}$$

A simple probabilistic argument shows that for $n \geq 1$,

$$a_j(n) = \sum_{k=0}^{j} a_k(n-1)\; a_{j-k}(1). \tag{3.37}$$

Hence

$$u(j) = \sum_{k=0}^{j} u(k)\; a_{j-k}(1) \qquad j>0. \tag{3.38}$$

Theorem 10

$$E\,\underline{t}_k^{*} = \sum_{j=0}^{k-1} u(j). \tag{3.39}$$

Proof

$$\begin{aligned} P\{\underline{t}_k^{*}=n\} &= P\{\underline{v}(n) \geq k \text{ and } \underline{v}(n-1)<k\} = \\ &= P\{\underline{v}(n) \geq k\} - P\{\underline{v}(n-1) \geq k\} = \\ &= \sum_{j=k}^{\infty} a_j(n) - \sum_{j=k}^{\infty} a_j(n-1) = \\ &= \sum_{j=0}^{k-1} \{a_j(n-1) - a_j(n)\}. \end{aligned}$$

The theorem follows now from

$$E\,\underline{t}_k^{*} = \sum_{n=1}^{\infty} n\, P\{\underline{t}_k^{*}=n\}.$$

When the economic stock at the start of a period is equal to j we say that the system is in state j. Suppose that the initial state is i and assume that neither at the start of the initial period nor at the start of any succeeding period an order is placed.
If the initial state is i and A is a given set of integers, define

β_{ijA} = probability that j is the first state in A assumed by the system. (3.40)

Let for $i \geq 1$

$f_{ik} \stackrel{def}{=}$ probability that the system is in state k at the start of the period in which the economic stock falls below 1 for the first time, if the initial state is i. (3.41)

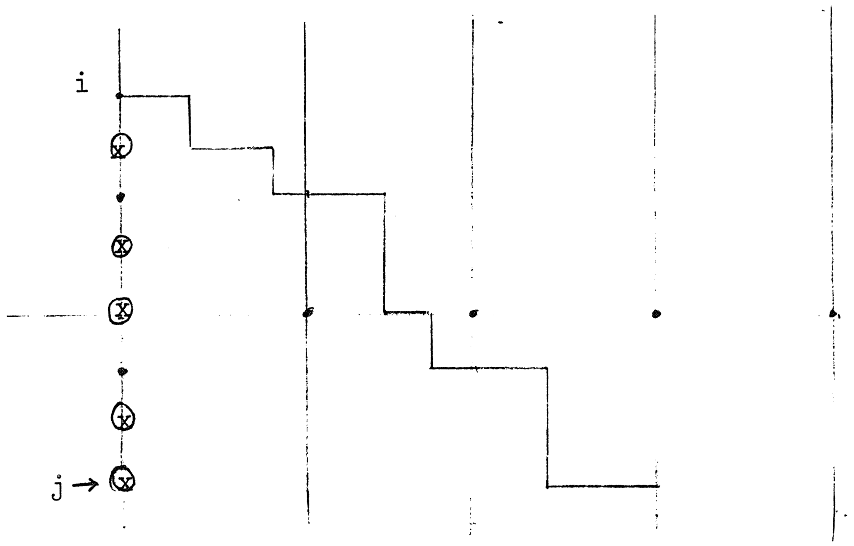

figure 1

The ⓧ marked points belong to A.

From the definition it follows:

1) $i \in A$

$$\beta_{ijA} = \begin{cases} 1 & j = i, \\ 0 & j \neq i. \end{cases} \qquad (3.42)$$

2) $i \notin A$

$$\beta_{ijA} = 0 \qquad j \notin A \text{ or } j > i \qquad (3.43)$$

The recursionformula

$$\beta_{ijA} = \sum_{i \leq k \leq j} a_{i-k}(1)\, \beta_{kjA}. \qquad (3.44)$$

follows by using a simple probabilistic argument

If

$$A = \{j \mid L < j \leq s\} \qquad (3.45)$$

where L may be $-\infty$, an analytical formula for β_{ijA} can be given.

Let $i > s$ and $L < j \leq s$

$$\beta_{ijA} = \sum_{n=1}^{\infty} \sum_{h=0}^{i-s-1} P\{\underline{v}(n-1) = h,\ \underline{v}(n) - \underline{v}(n-1) = i-h-j\} =$$

$$= \sum_{n=1}^{\infty} \sum_{h=0}^{i-s-1} P\{\underline{v}(n-1) = h\}\, P\{\underline{v}(n) - \underline{v}(n-1) = i-h-j\} =$$

$$= \sum_{n=1}^{\infty} \sum_{h=0}^{i-s-1} P\{\underline{v}(n-1) = h\}\, P\{\underline{v}(1) = i-h-j\} =$$

$$= \sum_{h=0}^{i-s-1} u(h)\, a_{i-h-j}(1). \qquad (3.46)$$

Note $$\sum_{j \leq s} \beta_{ijA} = 1 \qquad \text{if} \quad L = -\infty \qquad (3.47)$$

For the probability f_{ik} holds,

$$f_{ik} = \sum_{n=1}^{\infty} P\{\underline{v}(n-1) = i-k,\ \underline{v}(n) - \underline{v}(n-1) \geq k\}$$

$$= \sum_{n=1}^{\infty} P\{\underline{v}(n-1) = i-k\}\, P\{\underline{v}(1) \geq k\} =$$

$$= u(i-k)\left(1 - \sum_{j=0}^{k-1} a_j(1)\right), \qquad i \geq 1 \text{ and } 1 \leq k \leq i. \qquad (3.48)$$

Suppose the econonomic stock is $i \geq 1$ at the start of an initial period, say period 1. Assume no orders are placed if the economic stock is positive. Define the integral-valued $\underline{P}(i)$ by

$$\underline{P}(i) = \min \{ k \mid \underline{v}(k) \leq i \} . \qquad (3.49)$$

Hence in the $\underline{P}(i)^{th}$ period the economic stock falls below 1 for the first time. No order is placed at the start of any of the periods $1, \ldots, \underline{P}(i)$.

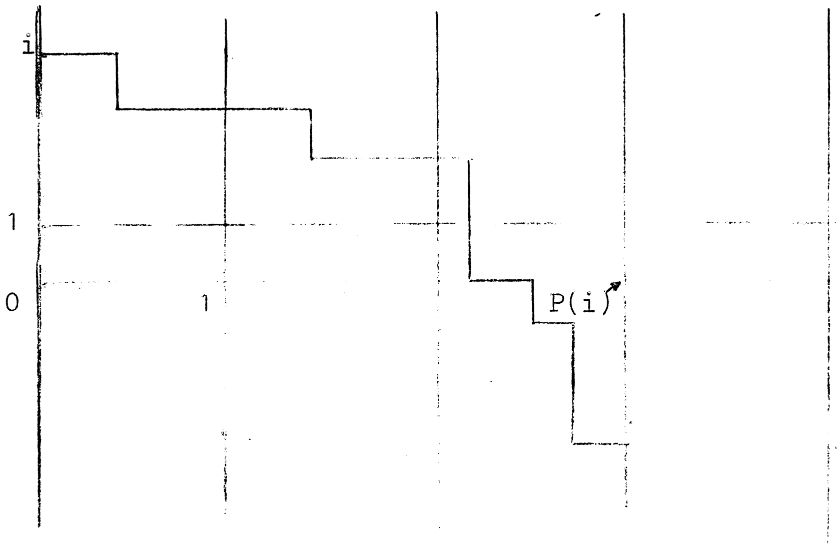

Figure 2

A realisation P(i) of $\underline{P}(i)$.

Let for $i \geq 1$

$$N(i) \stackrel{def}{=} \sum_{k \geq i} E\left[\{c_3 + c_2(\underline{P}(i) - \underline{t}_k + T)\} \tau(\underline{P}(i) - \underline{t}_k)\right] \qquad (3.50)$$

where c_2, c_3 and T are given constants. N(i) can be interpreted as the expected cost of subsequent delivery corresponding to the excess demand in period $\underline{P}(i)$.

It will be proved that

$$N(i) = \sum_{k=1}^{i} u(i-k) \left[c_2 \{ \frac{\lambda}{2} \sum_{n=1}^{\infty} n p_n - \sum_{j=1}^{k} E(1-\underline{t}_k)\ \iota(1-\underline{t}_k) \right] +$$

$$+ (c_2\ T + c_3)\ (\lambda \sum_{n=1}^{\infty} n p_n \sum_{j=0}^{i-1} u(j) - i), \quad i \geq 1. \qquad (3.51)$$

Proof $\qquad N(i) = N_1(i) + N_2(i)$

where

$$N_1(i) = (c_3 + c_2 T) \sum_{k=i+1}^{\infty} E\iota(\underline{P}(i) - \underline{t}_k) \qquad (3.52)$$

and

$$N_2(i) = c_2 \sum_{k=i+1}^{\infty} E\{(\underline{P}(i) - \underline{t}_k)\ \iota\ (\underline{P}(i) - \underline{t}_k)\}\ . \qquad (3.53)$$

Using a simple probabilistic argument it follows

$$N_1(i) = (c_3 + c_2 T)(E\ \underline{v}_i^{*} - i) = \qquad (3.54)$$

$$= (c_3 + c_2 T)(\lambda \sum_{n=1}^{\infty} n p_n \sum_{j=0}^{i-1} u(j) - i).$$

By the theorem of total expectation

$$N_2(i) = c_2 \sum_{k=1}^{i} f_{ik} \sum_{j=k+1}^{\infty} E\{(1-\underline{t}_j)\ \iota(1-\underline{t}_j) \mid \underline{t}_k \leq 1\}\ . \qquad (3.55)$$

It will be obvious that for $j \geq k$ and $0 \leq t \leq 1$,

$$P\{\underline{t}_j \leq t \mid \underline{t}_k \leq 1\} = P\{\underline{t}_j \leq t\} / P\{\underline{t}_k \leq 1\}. \qquad (3.56)$$

Hence

$$E\{(1-\underline{t}_j)\ \varepsilon(1-t_j)\,|\,\underline{t}_k \leq 1\} = E\{(1-\underline{t}_j)\ \varepsilon(1-\underline{t}_j)\}/P\{\underline{t}_k \leq 1\}. \quad (3.57)$$

From (3.48) it follows that,

$$f_{ik} = u(i-k)\ P\{\underline{t}_k \leq 1\}. \quad (3.58)$$

The assertion follows now from (3.25), (3.54), (3.57) and (3.58).

We note that N(i) can also be computed recurively. A simple probabilistic argument shows that,

$$N(i) = c_2 \sum_{k=i+1}^{\infty} E\{(1-\underline{t}_k)\ \varepsilon(1-\underline{t}_k)\} + (c_2T+c_3) \sum_{k=i+1}^{\infty} E\ \varepsilon(1-\underline{t}_k) + \sum_{j=0}^{i-1} a_j(1)\ N(i-j). \quad (3.59)$$

The definition of the ε-function implies that,

$$E\ \varepsilon(1-\underline{t}_k) = P\{\underline{t}_k \leq 1\}\ . \quad (3.60)$$

It is easily verified that,

$$\sum_{k=i+1}^{\infty} E\ \varepsilon(1-\underline{t}_k) = \sum_{j=i+1}^{\infty} (j-i)a_j(1) =$$

$$= \lambda \sum_{n=1}^{\infty} np_n - i + \sum_{j=0}^{i-1} (i-j)a_j(1). \quad (3.61)$$

<u>Appendix</u> In this appendix we shall describe another approach to calculate the probabilities $a_n(t)$, which may be more convenient from a computational point of view. The probabilities $a_n(t)$ are uniquely determined by the power-series development of

$$f(s) = e^{-\lambda t(1-\sum_{k=0}^{\infty} p_k s^k)}\ . \quad (3.62)$$

Obviously,

$$f(s) = e^{-\mu t(1- \sum_{k=1}^{\infty} q_k s^k)}, \qquad (3.63)$$

where

$$\mu = \lambda(1-p_o) \qquad (3.64)$$

$$q_k = p_k/(1-p_o) \qquad k \geq 1 . \qquad (3.65)$$

It follows from (3.63) that it can be equivalently stated that the customers arrive according to a Poisson process with intensity μ and that each customer demands $k \geq 1$ units with probability q_k. Let the behaviour of the customers be described by this process.

A simple argument shows that,

$$a_0(t) = P\{ \text{no arrival in } (0,t] \} = e^{-\mu t} \qquad (3.66)$$

and

$$a_n(t) = \sum_{k=1}^{n} P\{k \text{ customers arrive in } (0,t] \text{, the total demand of the } k \text{ customers is } n\} =$$

$$= \sum_{k=1}^{n} c(n|k)\, e^{-\mu t} \frac{(\mu t)^k}{k!} , \qquad (3.67)$$

where

$c(n|k)$ = probability that the total demand of k customers is n. (3.68)

It will be obvious that for $n \geq k \geq 1$,

$$c(n|k) = \sum_{i=1}^{n-k+1} P\{\text{the first customer demands } i \text{ units and the total demand of the other } k-1 \text{ customers is } n-i\} =$$

$$= \sum_{i=1}^{n-k+1} q_i\, c(n-i|k-1), \qquad (3.69)$$

where

$$q(n|0)=\begin{cases}1 & n = 0\\ 0 & n \neq 0.\end{cases} \qquad (3.70)$$

We note that $c(n|k) = 0$ for $k > n$. Furthermore $c(n|k)=0$ for $1 \leq k \leq \left[\frac{n}{N}\right]$ if $p_i=0$ for $i > N$.

4. Application to the inventory problem.

4.1. As is stated in the introduction the stock level of a single item is reviewed periodically, at which time an order may be placed. An order placed at the start of period $n(=1,2,\dots)$ is delivered at the start of period $n+T$, where T is a known nonnegative integer. Excess demand is backlogged until it is subsequently filled by a delivery. Hence when at the start of a period just after ordering the economic stock (= stock on hand plus orders outstanding) is i and the demand in the period is v, the economic stock at the end of the period is $i-v$. It is wellknown [7] that the optimal policy is only a function of the economic stock. This is the reason we need only to consider policies which base at a review their decisions on the economic stock. A frequently used policy is the (s,S) policy: if, at a review, the economic stock $i\leq s$, order $S-i$ units; if $i>s$, do not order. Using results given in [6,10,11] it can be shown that for our problem generally a (s,S) policy will be optimal if the ordering cost $c(q) = c.q+K\delta(q)$, where $K\geq 0$, $\delta(0) = 1$ and $\delta(q) = 0$ for $q>0$. When $K = 0$ then $s = S-1$ for the optimal (s,S) policy. If $c(q)$ has not the simple form as above, in general a (s,S) policy is not optimal. However it may be interesting to determine the best of the (s,S) policies. For that purpose we consider the class Z_0 of policies, which is rather artificial at a first glance. A policy of Z_0 is the following type: "Once the economic stock has been replenished until S in future each time when replenished it is done until S". In our model the economic stock may never exceed a known integer M. If, at a review, the economic stock is less than or equal to a known integer a it has to be replenished at least until a known integer B.

4.2. Definition of the state space.

The following state space X is considered. It consists of the points:

(a) (i,t). Where $0\leq t\leq 1$ and $i\leq M$, i integer. This state corresponds to the situation that the economic stock is i and a fraction t of the period considered is dapsed.

(b) (i,S,t). Where $0\leq t\leq 1$, $i\leq S$ and $S = B,\dots,M$. This state corresponds to the same situation as the state (i,t) does. The meaning of S is:

"in the past each replenishment was until S". The choice of Z_0 implies states (i,S,t) to describe the behaviour of the system.
We shall later see that the state space X contains all information that we shall need in the sequel if we choose the set $A_{0,1}$ in a proper way. We note that the points $(i,0)$ and $(i,S,0)$ correspond to an economic stock i at the start of a period, at which time an order may be placed. From now on the states $(i,0)$ and $(i,S,0)$ will be denoted by (i) and (i,S). Let

$$X_0 = \{(i)\,|\,i\leq M\}\cup\{(i,S)\,|\,i\leq S,\ S=B,\dots,M\}. \qquad (4.2.1)$$

In the states $x\notin X_0$ only null-decisions can be made. Each decision d will be represented by the economic stock just after the decision. Let $D(x)$ be the set of feasible decisions in state x.

$$D(x) = \begin{cases} \{d\,|\,i\leq d\leq M\} & x=(i),\ i>a \\ \{d\,|\,i<d\leq M\} & x=(i),\ i\leq a \\ \{d=i,\ d=S\} & x=(i,S),\ i>a \\ \{d=S\} & x=(i,S),\ i\leq a. \end{cases} \qquad (4.2.2)$$

It will be obvious that Z_0 is the class of strategies which dictate in each state $x\in X$ a decision $d\in D(x)$. To each strategy $z\in Z_0$ the set A_z corresponds, which consists of the states in which z dictates a non null-decision. Obviously,

$$A_0 = \bigcap_{z\in Z_0} A_z = \{(i)\,|\,i\leq a\}\cup\{(i,S)\,|\,i\leq a,\ S=B,\dots,M\}. \qquad (4.2.3)$$

The natural process results from the passage of time, the demands of customers and the delivery of orders. In each state the natural process can start. However in the natural process no orders are placed by the decisionmaker!
The non null-decision d=S made in state (i) or state (i,S) transforms the system into state $(S,S,0+)$. The system runs then successively through the states $(\underline{i}_t,S,t)$, where $\underline{i}_t$ is the economic stock t $(0<t\leq 1)$ units of time after that decision. At the start of the next period the system assumes a state $(\underline{i},S)$. The null-decision made in state (i) or

(i,S) transforms the system into state (i,0+) respectively (i,S,0+).

4.3. Determination of the k- and t-functions.

Assume from now on that $a \geq 0$. The case $a<0$ can be treated analogeously. To determine the k- and t-functions we may change the natural process when the system assumes a state of A_0. This is allowed because in each state of A_0 the decisionmaker has to intervene. Hence situations which may arise if the natural process is changed in A_0 never occur in any decisionprocess (= in reality). This remark enables us to derive the k- and t-functions in a simple and direct way. Choose

$$A_{0,1} = \{(i)\,|\,i \leq 0\} \cup \{(i,S)\,|\,i \leq 0,\ S=B,\ldots,M\}. \qquad (4.3.1)$$

and

$$A_{0,2} = \{(i)\,|\,i \leq a\} \cup \{(i,S)\,|\,i \leq a,\ S=B,\ldots,M\}. \qquad (4.3.2)$$

When the system assumes a state of A_0 with negative economic stock it is supposed that in the natural process an order (called a natural order) is placed. The order size is equal to the absolute value of the economic stock. The ordering cost is zero and the time of delivery is T periods. Furthermore it is assumed that no customers arrive if a natural order is outstanding. When the walk $\underline{w}^{0,1}$ or $\underline{w}^{d,1}$ assumes a state of A_0 with negative economic stock we may change the course of the walk. The walk considered does not end on that moment but it remains subjected to the natural process and it ends as soon as the natural order is delivered.*) The function k(x;d) is the difference in expected cost of the walks $\underline{w}^{d,1}$ and $\underline{w}^{0,1}$ having x as initial state. The function t(x;d) is the difference in expected duration of $\underline{w}^{d,2}$ and $\underline{w}^{0,2}$.

*) This can be roughly explained as follows: Add a state Q to the state space, where Q corresponds to the moment of a delivery of a natural order. Let each $z \in Z_0$ dictate some null-decision in Q. (Note in any decisionprocess Q is never assumed). Take

$$A_{0,1} = Q \cup \{(0)\} \cup \{(0,S)\,|\,B \leq S \leq M\}.$$

Obviously,

$$k(x;d) = t(x;d) = 0 \qquad d= \text{null-decision}. \qquad (4.3.3)$$

Furthermore it will be obvious that

$$k((i,S);S) = k((i);S) \quad \text{and} \quad t((i);S) = t((i,S);S). \quad (4.3.4)$$

Consider the walks $\underline{w}^{0,1}$ and $\underline{w}^{S,1}$ having (i) as initial state. For any unit which is kept in stock for t units of time the storage cost is c_1t, while for any unit which is delivered subsequently t>0 units of time the penalty cost is c_2t+c_3. Storage cost corresponding to units from the initial economic stock i belong both to the cost of the walk $\underline{w}^{0,1}$ as $\underline{w}^{S,1}$. The same holds for penalty cost which may incurred for these units. In the walk $\underline{w}^{S,1}$ we have storage cost for units of the last order S-i if after the delivery of that order the stock is positive. That expected cost (= difference in expected storage cost of the walks $\underline{w}^{S,1}$ and $\underline{w}^{0,1}$) is given by,

$$c_1 \sum_{k=i^+ +1}^{S} E\{(\underline{t}_k - T)\iota(T-\underline{t}_k)\}, \qquad (4.3.5)$$

where

$$i^+ = \begin{cases} i & i>0 \\ 0 & i\leq 0 \,. \end{cases} \qquad (4.3.6)$$

In the walk $\underline{w}^{S,1}$ we have penalty cost for units of the last order S-i when units of this order are demanded for its delivery. That expected penalty cost is given by

$$\sum_{k=i^+ +1}^{S} E\{(c_2(T-\underline{t}_k)+c_3)\iota(T-\underline{t}_k)\}. \qquad (4.3.7)$$

For any of the walks $\underline{w}^{0,1}$ and $\underline{w}^{S,1}$ in some period the economic stock is less than 1 for the first time. The definition of the natural process for states of A_0 implies that the penalty cost of that excess demand is N(i) respectively N(S). Where for k>0 N(k) is given by (3.52) and N(k) = 0 for $k\leq 0$. Hence the difference in expected penalty cost of the walks $\underline{w}^{S,1}$ and $\underline{w}^{0,1}$ is given by,

$$(4.3.7) + N(S) - N(i). \tag{4.3.8}$$

Consider next the walks $\underline{w}^{0,2}$ and $\underline{w}^{S,2}$ having (i) as initial state. The walk $\underline{w}^{0,2}$ respectively $\underline{w}^{S,2}$ ends as soon as at the start of a period the economic stock is less than or equal to a.
Hence,

$$t((i);S) = E\,\underline{t}^{*}_{S-a} - E\,\underline{t}^{*}_{i-a}\ , \tag{4.3.9}$$

where

$$\underline{t}^{*}_{k} \equiv 0 \qquad \text{for } k\leq 0. \tag{4.3.10}$$

Using (3.20), (3.23), (3.35) it follows that,

$$k((i);S) = \begin{cases} c(S-i) + (c_1+c_2)\sum_{k=i+1}^{S}\sum_{m=0}^{k-1} a_m(T)\, E\,\underline{t}_{k-m} \\ +c_2 \sum_{k=i+1}^{S} (T-E\,\underline{t}_k) + c_3 \sum_{k=i+1}^{S} (1-\sum_{m=0}^{k-1} a_m(T)) \\ +N(S)-N(i) \qquad i\geq 0,\ \max(i+1,B)\leq S\leq M, \\ c(S-i)-c(S)+k((0);S) \qquad i<0,\ B\leq S\leq M. \end{cases} \tag{4.3.11}$$

and

$$t((i);S) = \sum_{k=(i-a)^+}^{S-a-1} u(k) \qquad \max(i+1,B)\leq S\leq M. \tag{4.3.12}$$

We note that,

$$t((i);S) = t((j);S) \qquad \text{for } i,j\leq a. \tag{4.3.13}$$

4.4. <u>Determination of an optimal strategy.</u>
For each strategy $z\in Z_0$ the set

$$E(S) = \{(i,S,t)\,|\,i\leq S,\ 0\leq t\leq 1\} \tag{4.4.1}$$

constitutes a simple ergodic set of the decisionprocess. The transient set is given by

$$T = \{(i,t) \mid i \leq M,\ 0 \leq t \leq 1\}. \qquad (4.4.2)$$

Hence for each decisionprocess the state space can be decomposed in M-B+1 disjunct simple ergodic sets and a transient set. When in a decisionprocess in state (i) the non null-decision d=S is made, the system is transferred into the state $(S,S,0+) \in E(S)$ and the system remains for ever in E(S). For each $x \in E(S)$ the criterionfunction r(z;x) has the same value, say $r_S(z)$. Hence,

$$r(z;x) = r_S(z) \qquad x \in E(S). \qquad (4.4.3)$$

Consider the functional equations

$$r(z;x) = E\, r(z;\underline{I}_1) \qquad (4.4.4)$$

and

$$c(z;x) = k(x;z(x)) - r(z;x)t(x;z(x)) + Ec(z;\underline{I}_1), \qquad (4.4.5)$$

where $\underline{I}_1$ is the first future state in A_z assumed by the system if the initial state is x and strategy z is applied. From (4.3.3) it follows that

$$c(z;x) = Ec(z;\underline{I}_1) \qquad x \notin A_z. \qquad (4.4.6)$$

To solve (4.4.4) and (4.4.5) we choose in each simple ergodic set E(S) an arbitrary state e(S) and we put

$$c(z;e(S)) = 0 \qquad B \leq S \leq M. \qquad (4.4.7)$$

Choose

$$e(S) = (S,S,0+). \qquad (4.4.8)$$

Let $\underline{I}_1^*$ be the first future interventionstate, if the initial state is e(S) and strategy z is applied. We note that $\underline{I}_1^* = (\underline{i}^*,S)$.
From (4.4.3), (4.4.5) and (4.4.6) it follows that

$$c(z;e(S)) = Ec(z;\underline{I}_1^*) =$$

$$=E\{k(\underline{I}_1^*;z(\underline{I}_1^*))-r_S(z)t(\underline{I}_1^*;z(\underline{I}_1^*))\}+ Ec(z;\underline{I}_1^{**}), \tag{4.4.9}$$

where $\underline{I}_1^{**}$ is the first future interventionstate given the initial state $\underline{I}_1^*$. When the system has x as initial state and z(x) = S, the first future interventionstate has the same distribution as $\underline{I}_1^*$, hence

$$0 = Ec(z;\underline{I}_1^*) = Ec(z;\underline{I}_1^{**}). \tag{4.4.10}$$

From (4.4.9) and (4.4.10) it follows that

$$r_S(z) = Ek(\underline{I}_1^*;z(\underline{I}_1^*))/Et(\underline{I}_1^*;z(\underline{I}_1^*)). \tag{4.4.11}$$

When once $r_S(z)$ has been calculated for S=B,...,M the other unknowns can be easily determined [c.f. (2.14), (2.15) and (2.17)].

If $(i,S)\notin A_z$ then,

$$c(z;(i,S) = k((i);S) - r_S(z)t((i);S) + Ec(z;\underline{I}_1) =$$

$$= k((i);S) - r_S(z)t((i);S). \tag{4.4.12}$$

If $(i,S)\notin A_z$ (hence i>a) then,

$$c(z;(i,S)) = Ec(z;(\underline{j},S)), \tag{4.4.13a}$$

where $(\underline{j},S)$ is the first future state in A_z assumed by the system. Using the theorem of total expectation it follows that c(z;(i,S)) for $(i,S)\notin A_z$ can be computed recursively,

$$c(z;(i,S)) = \sum_{k=0}^{\infty} a_k(1)c(z;(i-k,S)). \tag{4.4.13b}$$

If $(i)\in A_z$ and $z((i)) = S$,

$$r(z;(i)) = r_S(z)$$

$$c(z;(i)) = k((i);S) - r_S(z)t((i);S). \tag{4.4.14}$$

If $(i)\notin A_z$,

$$r(z;(i)) = Er(z;(\underline{j})) = \sum_{k=0}^{\infty} a_k(1)r(z;(i-k))$$

$$c(z;(i)) = Ec(z;(\underline{j})) = \sum_{k=0}^{\infty} a_k(1)c(z;(i-k)), \tag{4.4.15}$$

where $(\underline{j})$ is the first future state in A_z.

Consider next the functions $r(d.z;x)$ and $c(d.z;x)$. From the definitions [c.f. (2.21) and (2.22)] it follows that for d=null-decision

$$r(d.z;x) = r(z;x) \ , \ c(d.z;x) = c(z;x) \tag{4.4.16}$$

and for $d = S\text{>}i$,

$$\begin{aligned} r(d.z;(i)) &= r(d.z;(i,S)) = \\ &= r(z;e(S)) = r_S(z) \end{aligned} \tag{4.4.17}$$

and

$$\begin{aligned} c(d.z;(i)) &= c(d.z;(i,S)) = \\ &= k((i);S) - r_S(z)t((i);S) + c(z;e(S)) = \\ &= k((i);S) - r_S(z)t((i);S). \end{aligned} \tag{4.4.18}$$

We note that,

$$\min_{d\in D((i))} r(d.z;(i)) = \min\{r(z;(i));r_S(z)\,|\,S = \max(i+1,B),\ldots,M\}. \tag{4.4.19}$$

Hence if $(i)\in A_{z_1}$ then [c.f. (2.28)],

$$r(\hat{z};(i)) = \min_{M \geq S \geq \max(i+1,B)} r_S(z). \qquad (4.4.20)$$

From (4.4.20) it follows that,

$$r(\hat{z};(i)) \leq r(\hat{z};(j)) \qquad i<j \text{ and } (i),\ (j)\in A_{z_1}. \quad (4.4.21)$$

The relation (4.4.21) simplifies the determination of $A_{\hat{z}}'$.
It will be obvious that for the determination of an optimal strategy we can restrict ourselves to the states of X_0.
Let us define for $x,y\in X_0$,

$p_{xy}(z)$ = probability that y is the first future interventionstate if the initial state is x and strategy z is applied. (4.4.22)

Define for convenience of notation the sets of indices

$$B(z) = \{i \mid (i)\in A_z\} \text{ and } B(z;S) = \{i \mid (i,S)\in A_z\}, \qquad (4.4.23)$$

and the sets of states

$$C(z) = \{(i) \mid (i)\in A_z\} \text{ and } C(z;S) = \{(i,S) \mid (i,S)\in A_z\}. \qquad (4.4.24)$$

A simple argument shows that [c.f. (3.40)],

$$p_{xy}(z) = \begin{cases} \beta_{S,j,B(z;S)} & x = (i,S)\in A_z,\ y = (j,S) \\ \beta_{i,j,B(z;S)} & x = (i,S)\notin A_z,\ y = (j,S) \\ \beta_{S,j,B(z;S)} & x = (i)\in A_z,\ z((i)) = S,\ y = (j,S) \\ \beta_{i,j,B(z)} & x = (i)\notin A_z,\ y = (j). \end{cases} \qquad (4.4.25)$$

Using (4.4.25) we can write (4.4.11) as

$$r_S(z) = \sum_{j\in B(z;S)} \beta_{S,j,B(z;S)} k((j);S) \Big/ \sum_{j\in B(z;S)} \beta_{S,j,B(z;S)} t((j);S). \tag{4.4.26}$$

In particular it follows that the expected mean cost per unit of time of a (s,S) policy (s≥0) is given by,

$$r(s;S) = \sum_{j\leq s} \alpha(s,S,j)k((j);S) \Big/ \sum_{j\leq s} \alpha(s,S,j)t((j);S) \tag{4.4.27}$$

where [c.f. 3.46],

$$\alpha(s,S,j) = \sum_{k=0}^{S-s-1} u(k)a_{S-k-j}(1) \qquad j\leq s. \tag{4.4.28}$$

The range of the summationvariable in the numerator of (4.4.26) can be reduced to a finite one, because $t((i);S) = t((j);S)$ for $i,j,\leq a$ and

$$\sum_{j\leq a} \beta_{S,j,B(z;S)} = 1-\sum_{j>a} \beta_{S,j,B(z;S)}. \tag{4.4.29}$$

We shall later see that the range of the summationvariable in the denumerator can be reduced to a finite one too if the ordering cost is linear from some ordersize,
Next it will be shown how the set $A_{\hat{z}}$ can be determined. We note firstly that each set A_z contains only a finite number of states for which the null-decision is feasible. Hence [c.f. (2.35)] the intersectionset $A_{\hat{z}}$ is only determined by a finite number of sets A, thus

$$A_{\hat{z}}{}'\in X(\hat{z}). \tag{4.4.30}$$

Let $A_0 \subset A \subset A_{z_1}$. When the system is subjected to the natural process with initial state in $C(z_1;S)$ respectively in $C(z_1)$ then the first state in the set A assumed by the system belongs to $C(z_1;S)$ respectively $C(z_1)$.

Hence to determine which states of A_{z_1} belong to $A_{\hat{z}}$ we can consider the sets $C(z_1)$ and $C(z_1;S)$ separately. From the definitions (2.31), (2.32) and (3.40) it follows that,

$$r(A.\hat{z};x) = \sum_{y\in A} \beta_{xyA} r(\hat{z};y)$$

and

$$c(A.\hat{z};x) = \sum_{y\in A} \beta_{xyA} c(\hat{z};y). \tag{4.4.31}$$

In particular for $x\in A$,

$$r(A.\hat{z};x) = r(\hat{z};x) \quad \text{and} \quad c(A.\hat{z};x) = c(\hat{z};x) \tag{4.4.32}$$

From (4.4.20) and (4.4.31) it follows that

$$r(A.\hat{z};(i)) \leq r(\hat{z};(i)) \qquad \text{for } (i)\in A_{z_1}. \tag{4.4.33}$$

Furthermore

$$r(A.\hat{z};(i,S)) = r_S(z). \tag{4.4.34}$$

Conclusion for the determination of $A_{\hat{z}}'$ we need only to compare $c(A.\hat{z};x)$ and $c(\hat{z};x)$ for $x\in A_{z_1}$ [c.f. (2.34)]. Another way than (4.4.31) to calculate $c(A.\hat{z};x)$ for $x\notin A$ follows from the theorem of total expectation and definition (2.32),

$$c(A.\hat{z};x) = \begin{cases} \sum_{k=0}^{\infty} a_k(1)c(A.\hat{z};i-k) & x = (i)\notin A \quad (4.4.35) \\ \sum_{k=0}^{\infty} a_k(1)c(A.\hat{z};(i-k,S)) & x = (i,S)\notin A. \quad (4.4.36) \end{cases}$$

Firstly we shall determine which states of $C(z_1)$ belong to $A_{\hat{z}}'$. The latter set encloses A_0 hence it contains the states (j) with $j\leq a$. Let the initial set H be defined by

$$H = \{(j) | j\leq a\}. \tag{4.4.37}$$

We shall build up H to $A_{\hat{z}}' \cap C(z_1)$.
Calculate successively for $i=a+1,\ldots,M-1$, $i \in B(z_1)$

$$c(H.\hat{z};(i)) - c(\hat{z};(i)). \qquad (4.4.38)$$

State (i) belongs to $A_{\hat{z}}'$ if and only if the difference is positive.
We add (i) to H if the difference is positive, otherwise not. Note that $c(H.\hat{z};(i)) = c(\hat{z};(i))$ if $(i) \in H$. Some reflections about this procedure show that the finally obtained set H consists of the states of $C(z_1)$ which belong to $A_{\hat{z}}'$.
Analogeously we determine the sets $C(z_1;S) \cap A_{\hat{z}}'$ for $S=B,\ldots,M$.
Define for fixed S the initial set

$$H(S) = \{(j,S) \mid j \leq a\}. \qquad (4.4.39)$$

Calculate successively for $i=a+1,\ldots,S-1$; $i \in B(z_1;S)$,

$$c(H(S).\hat{z};(i,S)) - c(\hat{z};(i,S)). \qquad (4.4.40)$$

We add (i,S) to H(S) if and only if the difference is positive. The union of the final sets H(S) and H is the set $A_{\hat{z}}'$,

$$A_{\hat{z}}' = \bigcup_{S=B}^{M} H(S) \cup H. \qquad (4.4.41)$$

We are now in position to give the iterationprocedure for our problem. It will be obvious that in the present formulation of the problem we have to consider a denumerable number of states. However we can restrict ourselves to a finite number of states if for some $N \geq 1$ and some real c,U :

$$c(q) = cq+U \qquad q \geq N. \qquad (4.4.42)$$

Assume from now on that (4.4.42) holds. When the ordering cost c(q) is replaced by

$$c^*(q) = c(q) - c.q, \qquad (4.4.43)$$

the expected mean cost per unit of time of each strategy is reduced with $c\lambda \sum_1^\infty np_n$ (= c times the expected demand per unit of time), because excess demand is backlogged. Hence we reduce our model to an equivalent one when we consider $c^*(q)$ instead of c(q). For this reason we refer to $c^*(q)$ hereafter as the ordering cost. From (4.4.42) and (4.4.43) it follows that,

$$c^*(q) = U \qquad q \geq N \qquad (4.4.44)$$

Hence [c.f. (4.3.11)],

$$k((i);S) = k((j);S) \qquad i,j \leq L(S) \qquad (4.4.45)$$

where

$$L(S) = \min(0,S-N) \ . \qquad (4.4.46)$$

Furthermore we have already [c.f. (4.3.13)],

$$t((i);S) = t((j);S) \qquad i,j \leq a \ . \qquad (4.4.47)$$

If $i \leq a$ then,

$$\min_{d \in D((i))} r(d.z;(i)) = \min_{B \leq d \leq M} r_d(z). \qquad (4.4.48)$$

Let D_z be the set of minimizing decisions d. Let

$$L \overset{def}{=} L(B) = \min(0,B-N) \qquad (4.4.49)$$

From (4.4.45) and (4.4.47) it follows that for each state (i) with $i \leq L$

$$\min_{d \in D_z} c(d.z;(i)) = \min_{d \in D_z} \{k((i);d) - r_d(z)t((i);d)\} \quad (4.4.50)$$

is reached for the same d.
By means of the test quantities r(d.z;x) and c(d.z;x) the strategy z_1 is deduced from z. The formulation of the iterationprocedure (see page 16) implies that when we start with a strategy z satisfying

$$z((i)) = z((j)) \qquad i,j \leq L, \qquad (4.4.51)$$

we meet thereafter only strategies of this type. Hence Z_0 contains an optimal strategy which satisfies (4.4.51). From now on we consider only those strategies z of Z_0 with z((i)) = z((j)) for $i,j \leq L$.
Using (4.4.45) and (4.4.47) it is easily verified that,

$$c(z;(i)) = c(z;(j)) \text{ and } c(\hat{z};(i)) = c(\hat{z};(j)) \qquad i,j \leq L$$

$$c(z;(i,S)) = c(z;(j,S)) \text{ and } c(\hat{z};(i,S)) = c(\hat{z};(j,S)) \qquad i,j \leq L(S). \qquad (4.4.52)$$

It will now be obvious that we need to determine the function k((i);S) only for $L(S) \leq i \leq S-1$ and t((i);S) only for $a \leq i \leq S-1$, where S = B,...,M. Furthermore an optimal strategy is found by applying the iteration-procedure only to the states (i) with $L \leq i \leq M-1$ and the states (i,S) with $L(S) \leq i \leq S-1$.
We shall now formulate the iterationprocedure.

1. Let strategy z be given. Determine for S = B,...,M:

$$r_S(z) = \frac{\sum_{j=L(S)+1}^{S-1} \beta_{S,j,B(z;S)} k((j);S) + \alpha(1;S)k((L(S));S)}{\sum_{j=a+1}^{S-1} \beta_{S,j,B(z;S)} t((j);S) + \alpha(2;S)t((a);S)}, \qquad (4.4.53)$$

where

$$\alpha(1;S) = 1 - \sum_{j>L(S)} \beta_{S,j,B(z;S)}$$

$$\alpha(2;S) = 1 - \sum_{j>a} \beta_{S,j,B(z;S)}. \qquad (4.4.54)$$

For $L(S) \le i \le S-1$,

$$c(z;(i,S)) = \begin{cases} k((i);S) - r_S(z)t((i);S) & (i,S)\in A_z \\ \sum_{k=0}^{i-L(S)-1} a_k(1)c(z;(i-k,S)) + \\ + \sum_{k=i-L(S)}^{\infty} a_k(1)c(z;(L(S),S)). & (i,S)\notin A_z. \end{cases} \qquad (4.4.55)$$

For $L \le i \le M-1$,

$$r(z;(i)) = \begin{cases} r_d(z) & z((i))=d>i \\ \sum_{k=0}^{i-L-1} a_k(1)r(z;(i-k)) + \sum_{k=i-L}^{\infty} a_k(1)r(z;(L)) & (i)\notin A_z. \end{cases} \qquad (4.4.56)$$

$$c(z;((i)) = \begin{cases} k((i);d) - r_d(z)t((i);d) & z((i))=d>i \\ \sum_{k=0}^{i-L-1} a_k(1)c(z;(i-k)) + \sum_{k=i-L}^{\infty} a_k(1)c(z;(L)) & (i)\notin A_z. \end{cases} \qquad (4.4.57)$$

2. Determine for $L \le i \le M-1$,

$$\min_{d\in D((i))} r(d,z;(i)) = \min\{r(z;(i));r_d(z)\,|\,d>i,\ d\in D((i))\}. \qquad (4.4.58)$$

Let $D_z((i))$ be the set of minimizing decisions. Minimize

$$c(d,z;(i)) = \begin{cases} c(z;(i)) & d = \text{null-decision} \\ k(i);d) - r_d(z)t((i);d) & d>i, \end{cases} \qquad (4.4.59)$$

with respect to $d \in D_z((i))$. Let $z_1((i))$ be a minimizing decision [Take $z_1((i)) = z((i))$ if $z((i))$ is a solution].

For each state $(i,S) \notin A_z$ we take $z_1((i,S)) = S$ if

$$k((i);S) - r_S(z)t((i);S) < c(z;(i,S)). \tag{4.4.60}$$

Otherwise $z_1((i,S)) = z((i,S))$. For states $(i,S) \in A_z$ we take $z_1((i,S)) = S$.

3. Determine

$$A_{\hat{z}}' = \bigcup_{S=B}^{M} H(S) \cup H. \tag{4.4.61}$$

[We use the recurrence relation [c.f. (4.4.35) and (4.4.51)],

$$c(H,\hat{z};(i)) = \sum_{k=0}^{i-L-1} a_k(1)c(H.\hat{z};(i-k)) + \sum_{k=i-L}^{\infty} a_k(1)c(\hat{z};(L)). \tag{4.4.62}$$

A similiar relation holds for $c(H(S).\hat{z};(i,S))$].

4. Define

$$z_2(x) = \begin{cases} z_1(x) & x \in A_{\hat{z}}' \\ \text{null-decision} & \text{otherwise.} \end{cases} \tag{4.4.63}$$

<u>Numerical example.</u>

Suppose that the following numerical data are given:

$$T=0 \; ; \; M=8 \; ; \; c(q) = c.q+4 \text{ for } q \geq 1 \; ; \; c_1 = 1 \; ;$$
$$\lambda=4 \; ; \; p_1 = 1 \; ; \; c_2 = 20 \; ; \; c_3 = 0 \; ; \; a = 2 \; ; \; B = 4.$$

Hence [c.f. (4.4.46)]

$$L = L(S) = 0 \qquad\qquad S = 4,\dots,8.$$

To begin with we determine once for all the quantities $a_k(1)$, u(k) and the functions k((i);d) and t((i);d).

Tabel 1.

k	$a_k(1)$	u(k)
0	.0183	1.0187
1	.0733	.0760
2	.1465	.1577
3	.1954	.2258
4	.1954	.2582
5	.1563	.2617
6	.1042	.2546
7	.0595	.2494

The $a_k(1)$ and u(k).

Tabel 2.

i \ d	4	5	6	7	8
0	18.7501	20.6003	22.6042	24.4834	26.4147
1	-6.8730	-5.0228	-3.0190	-1.1397	.7916
2	1.4797	3.3299	5.3337	7.2130	9.1443
3	4.4689	6.3191	8.3230	10.2022	12.1336
4		5.8502	7.8540	9.7333	11.6646
5			6.0038	7.8831	9.8144
6				5.8793	7.8106
7					5.9313

The function k((i);d).

Tabel 3.

i \ d	4	5	6	7	8
2	1.0947	1.2524	1.4782	1.7365	1.9982
3	.0760	.2337	.4596	.7178	.9795
4		.1577	.3836	.6418	.9035
5			.2258	.4841	.7458
6				.2582	.5199
7					.2617

The function t((i);d).

To determine an optimal strategy we need only to consider the states (i) with $0 \leq i \leq 7$ and the states (i,S) with $0 \leq i \leq S-1$, S = 4,...,8.
We shall now apply the iterationprocedure.

1. Let the initial strategy z be given by:

$$z((i)) = \begin{cases} 8 & 0 \le i \le 4 \\ i & 5 \le i \le 7 \end{cases}$$

and for S = 4,...,8

$$z((i,S)) = \begin{cases} S & i = 0,\dots,\min(S-1,4) \\ i & \text{otherwise.} \end{cases}$$

After some calculations:

Tabel 4.

S	$r_S(z)$
4	9.824
5	8.670
6	8.373
7	8.366
8	8.405

The function $r_S(z)$.

Tabel 5.

i \ S	4	5	6	7	8
0	7.996	9.742	10.227	9.956	9.619
1	-17.627	-15.881	-15.396	-15.667	-16.004
2	-9.275	-7.528	-7.043	-7.314	-7.651
3	3.722	4.293	4.475	4.197	3.900
4		4.483	4.642	4.364	4.070
5			.414	.142	-.186
6				-.274	-.596
7					-.318

The function c(z;(i,S)).

Furthermore,

$$r(z;(i)) = r_8(z) \qquad i = 0,\dots,7$$

and

$$c(z;(i)) = c(z;(i,8)) \qquad i = 0,\dots,7.$$

2. Applying step 2 of the iterationprocedure we find that

$$z_1((i)) = \begin{cases} 7 & 0 \le i \le 6 \\ i & i = 7 \end{cases}$$

and

$$z_1((i,S)) = z((i,S)) \qquad \text{for each } (i,S).$$

For states $x \notin A_{z_1}$,

$$c(\hat{z};x) = \min_{d \in D_z(x)} c(d \cdot z;x).$$

Hence,

$$c(\hat{z};(i)) = k((i);7) - r_7(z)t((i);7) \qquad i = 0,\dots,6$$

$$c(\hat{z};(i,S)) = c(z;(i,S)) \qquad i = 0,\dots,\min(S-1,4) \quad S = 4,\dots,8.$$

After some calculations:

$$c(\hat{z};(i)) = c(z;(i)),\ i = 0,\dots,4\ ;\ c(\hat{z};5) = 3.833\ ;\ c(\hat{z};6) = 3.719.$$

3. By applying the cutting procedure given in step 3 we find that,

$$A_{\hat{z}}' = \{0,1,2,3\} \cup \{(0,4),(1,4),(2,4)\} \cup \{(0,5),(1,5),(2,5),(3,5)\} \cup \cup \{(0,6),(1,6),(2,6),(3,6)\} \cup \{(0,7),(1,7),(2,7),(3,7)\} \cup \cup \{(0,8),(1,8),(2,8),(3,8)\}.$$

4.

$$z_2((i)) = \begin{cases} 7 & 0 \le i \le 3 \\ i & 4 \le i \le 7 \end{cases}$$

$$z_2((i,4)) = \begin{cases} 4 & 0 \le i \le 2 \\ i & i = 3 \end{cases}$$

and for S = 5,...,8

$$z_2((i,S)) = \begin{cases} S & 0 \leq i \leq 3 \\ i & 4 \leq i \leq S-1 . \end{cases}$$

With strategy z_2 we repeat the iterationcycle.

1). Let strategy z be equal to strategy z_2 found in step 4 of the previous iterationcycle.

After some calculations:

Tabel 6.

S	$r_S(z)$
4	9.767
5	8.477
6	8.061
7	7.989
8	8.032

The function $r_S(z)$.

Tabel 7.

i \ S	4	5	6	7	8
0	8.059	9.983	10.688	10.611	10.364
1	-17.565	-15.640	-14.935	-15.012	-15.259
2	-9.212	-7.287	-6.582	-6.659	-6.906
3	2.983	4.338	4.618	4.468	4.266
4		1.884	2.557	2.475	2.232
5			.639	.550	.310
6				-.093	-.331
7					-.237

The function c(z;(i,S)).

Furthermore,

$$r(z;(i) = r_7(z) \qquad i = 0,\ldots,7$$

and

$$c(z;(i) = \begin{cases} c(z;(i,7)) & i = 0,\ldots,6 \\ -.001 & i = 7. \end{cases}$$

2. After some calculations we find that strategy z_1 is equal to strategy z.

3. After some calculations we find that,

$$A_{\hat{z}}' = A_{z_1} (=A_z).$$

Hence we find as optimal strategy the strategy (3,7): if, at a review, the economic stock is less than or equal to 3, replenish it until 7. As by-product we find an optimal strategy for each replacement level S.

4.5 Remarks.

1. Often it is assumed that the storage cost incurred during a period is a function h of the stock on hand at the start of that period. The penalty cost for a period is a function p of the shortage at the end of that period. The functions h and p are taken zero for negative arguments.
The difference in expected storage cost of the walks $\underline{w}^{S,1}$ and $\underline{w}^{0,1}$ is then given by,

$$\sum_{k=0}^{S-1} a_k(T)H(S-k) - \sum_{k=0}^{i-1} a_k(T)H(i-k), \tag{4.5.1}$$

where

$$H(i) = h(i) + \sum_{n=1}^{\infty} \sum_{j=0}^{i-1} a_j(n)h(i-j) = \sum_{j=0}^{i-1} u(j)h(i-j). \tag{4.5.2}$$

The difference in expected penalty cost of the walks $\underline{w}^{S,1}$ and $\underline{w}^{0,1}$ is then given by,

$$\sum_{n=1}^{T-1} \sum_{j=0}^{i-1} \sum_{k=i}^{S-1} a_j(n-1)a_{k-j}(1)\Big[\sum_{m=1}^{T-n} \sum_{t=0}^{S-k-1} \{a_t(m)p(k-i+t) + a_t(m-1) \sum_{v=S-k-t}^{\infty} a_v(1)p(k-i+t+v)\}\Big] + \sum_{k=0}^{S-1} a_k(T)\{Ep(\underline{v}^*_{S-k}-S+k) - Ep(\underline{v}^*_{i-k}-i+k)\}. \tag{4.5.3}$$

where $\underline{v}^*_k$ is given by (3.31) for $k>0$ and $\underline{v}^*_k=0$ for $k\leq 0$.

2. Suppose that only a single order can be outstanding. The resulting problem can be solved analogeously. The modification of the state space will be obvious. The states (j) and (j,S) of the new state space correspond to the situation that at the start of a period the stock on hand is j and no order is outstanding.
Only in those states can be intervened. The solution of the new problem runs parallel to that of the original one. Assuming that (4.4.42) holds the iterationprocedure (page 43) can be simply transmitted, provided that we change the formulas for k- and t-functions and for the quantities $r_S(z)$. Those formulas become,

$$
\begin{aligned}
k((i);S) = {} & c^{*}(S-i) + c_1 \sum_{k=i^{+}+1}^{S} E\{(\underline{t}_k - T)\,\epsilon\,(\underline{t}_k - T)\} + \\
& + \sum_{k=i^{+}+1}^{S} E\{(c_2(T-\underline{t}_k) + c_3)\,\epsilon\,(T-\underline{t}_k)\} + \\
& + \sum_{k=S+1}^{\infty} E\{(c_2(2T-\underline{t}_k) + c_3)\,\epsilon\,(T-\underline{t}_k)\} + \\
& + \sum_{k=0}^{S-1} a_k(T)N(S-k) - N(i).
\end{aligned}
\tag{4.5.4}
$$

$$
t((i);S) = T + \sum_{k=0}^{S-a-1} a_k(T)E\underline{t}^{*}_{S-k-a} - E\underline{t}^{*}_{i-a}. \tag{4.5.5}
$$

$$
r_S(z) = A(S)/B(S), \tag{4.5.6}
$$

where

$$
\begin{aligned}
A(S) = {} & \sum_{k=0}^{S-L(S)-1} a_k(T)\{\sum_{j=L(S)+1}^{S-k} \beta_{S-k,j,B(z;S)} k((j);S) + \\
& + (1 - \sum_{j>L(S)} \beta_{S-k,j,B(z;S)}\ k((L(S));S)\} + \\
& + (1 - \sum_{k<S-L(S)} a_k(T)\ k((L(S));S)
\end{aligned}
\tag{4.5.7}
$$

and

$$B(S) = \sum_{k=0}^{S-a-1} a_k(T)\{\sum_{j=a+1}^{S-k} \beta_{S-k,j,B(z;S)} t((j);S) +$$

$$+ (1 - \sum_{j>a} \beta_{S-k,j,B(z;S)}\, t((a);S)\} +$$

$$+ (1 - \sum_{k<S-a} a_k(T)\, t((a);S). \qquad (4.5.8)$$

3. Until now we have supposed that only at the beginning of a period an order can be placed. Suppose now that at each point of time an order may be placed [2]. Between delivery and order we assume a lag of T units of time, where T nonnegative and real. The same assumptions about the various cost and the behaviour of the customers hold as in the periodic review case. Only at the moments a demand occurs we decide whether or not to place an order. This can be done without loss of generality because the customers arrive according to a Poisson process. An arbitrary number of orders may be outstanding. Following the same lines as in the periodic review case it can be shown that the average cost per unit of time of a (s,S) policy ($s\ge 0$) is given by,

$$r(s,S) = \sum_{j\le s} f(j,s,S)k(j;S) / \sum_{j\le s} f(j,s,S)t(j;S). \qquad (4.5.9)$$

Where,

$$k(i;S) = c(S-i) + c_1 \sum_{k=i^{+}+1}^{S} E\{(\underline{t}_k-T)\cdot(\underline{t}_k-T)\} +$$

$$+ \sum_{k=i^{+}+1}^{S} E\{(c_2(T-\underline{t}_k)+c_3)\cdot(T-\underline{t}_k)\} +$$

$$+ (c_2T+c_3)(E v_S^{*} - S - E\underline{v}_i^{*} + i^{+})\gamma(T), \qquad (4.5.10)$$

with

$$\gamma(T) = \begin{cases} 1 & T>0 \\ 0 & T=0. \end{cases} \qquad (4.5.11)$$

$$t(i;S) = E\underline{t}_{S-a} - E\underline{t}_{i-a}. \qquad (4.5.12)$$

$f(j,s,S) =$ P{the first economic stock level $\leq s$ is j if the initial economic stock is S and no orders are placed}. (4.5.13)

Obviously,

$f(j,s,S) =$ P{to reach j from S without the system assumes an intermediate economic stock level} +

$+ \sum_{k=1}^{S-s-1}$ P{to reach S-k from S and to reach j from S-k without the system assumes an intermediate economic stock level} =

$$= \sum_{k=0}^{S-s-1} f_k p_{S-k-j}/(1-p_0), \qquad (4.5.14)$$

where

$$f_k = \begin{cases} 1 & k=0 \\ \text{P\{to reach S-k from S\}} & k \geq 1. \end{cases} \qquad (4.5.15)$$

Using a simple probabilistic argument it follows that,

$$f_k = \sum_{i=0}^{k} p_i f_{k-i} \qquad k \geq 1. \qquad (4.5.16)$$

The range of the summationvariable in the numerator of (4.5.9) can be reduced to a finite one, because $t(i;S) = t(j;S)$ for $i,j \leq a$. When (4.4.42) holds we can reduce the range of the summationvariable in the denumerator to a finite one too by replacing $c(q)$ by $c^*(q)$. Automatically the ranges of the summationvariables are finite if $p_i=0$ for $i>k$ for some k. In this case $f(j,s,S) = 0$ for $j \leq s-k$.

Analogeously like in the periodic review case an iterationprocedure can be formulated to determine the best policy of the class Z_0 of policies (In the non-periodic review case the probabilities p_i play the role of the $a_i(1)$ in the periodic review case). Furthermore the case of a single outstanding order can be treated in the same way.

References.

1. Arrow, K.J., Karlin, S. and Scarf, H., Studies in the Mathematical Theory of Inventory and Production, Stanford University Press, Stanford, 1958.
2. Beckman, M., "An Inventory Model for Arbitrary Interval and Quantity Distributions of Demand", Management Science, 8,35-57 (1961).
3. Feller, W., An Introduction to Probability Theory and its Applications, Volume I and II, John Wiley, New York, 1957 and 1966.
4. Howard, R., Dynamic Programming and Markov Processes, John Wiley, New York, 1960.
5. Jewell, W.S., "Markov-Renewal Programming", Operations Research, 11, 938-971 (1963).
6. Johnson, E.L., "Optimality and Computation of (σ,S) Policies in the Multi-Item Infinite Horizon Inventory Problem", Management Science, 13, 475-492 (1967).
7. Karlin, S. and Scarf, H., "Inventory Models of the Arrow-Harris-Marschalk Type with Time Lag", Chapter 10 in [1].
8. Leve, de G., Generalized Markovian Decision Processes, Mathematical Centre Tracts No. 3 and 4, Amsterdam, 1964.
9. Scarf, H., Gilford, D. and Shelley, M., (eds.), Multistage Inventory Models and Techniques, Stanford University Press, Stanford, 1963.
10. Veinott, A.F., Jr., and Wagner, H.M., "Computing Optimal (s,S) Inventory Policies", Management Science, 11, 525-552 (1965).
11. Veinott, A.F., Jr., "On the Optimality of (s,S) Inventory Policies: New Conditions and a New Proof", Siam Journal on Applied Mathematics, 14, 1067-1083 (1966).