SIAM J. CONTROL AND OPTIMIZATION
Vol. 25, No. 6, November 1987

008

AN OPTIMAL STOCHASTIC PRODUCTION PLANNING PROBLEM WITH RANDOMLY FLUCTUATING DEMAND*

W. H. FLEMING†, S. P. SETHI‡ AND H. M. SONER§

Abstract. This paper considers an infinite horizon stochastic production planning problem with demand assumed to be a continuous-time Markov chain. The problems with control (production) and state (inventory) constraints are treated. It is shown that a unique optimal feedback solution exists, after first showing that convex viscosity solutions to the associated dynamic programming equation are continuously differentiable.

Key words. production planning, stochastic optimal control, viscosity solutions

AMS(MOS) subject classifications. 93E20, 35J65, 35K60, 60J60

Introduction. Thompson and Sethi [13] consider a production-inventory model which determines production rates over time to minimize an integral representing a discounted quadratic loss function. The model is solved both with and without nonnegative production constraints. It is shown that there exists a turnpike level of inventory, to which the optimal inventory levels approach monotonically over time. The model was generalized by Sethi and Thompson [11] and Bensoussan et al. [1] by incorporating an additive white noise term in the dynamics of the inventory process.

In this paper we consider an analogue of the Thompson–Sethi model, in which the demand rate $z(t)$ is a finite state Markov chain. A similar, but technically more complicated, analysis applies if $z(t)$ is a jump Markov process or a reflected diffusion subject to bounds $0 < z_0 \leqq z(t) \leqq z_1 < \infty$ (see [7]). We denote by $y(t)$, $p(t)$ the inventory level and the production rate. Production is the control variable, subject to the constraint $p(t) \geqq 0$. In § 5 we impose the state constraint $y(t) \geqq y_{\min}$ on the inventory level.

The control objective is to minimize an expected discounted cost of the form (4.1), which involves convex holding or shortage costs $h(y)$ and productions costs $c(p)$. The value (or minimum cost) $v(y, z)$ defined in (1.4) for initial data $y(0) = y$, $z(0) = z$ obeys the dynamic programming equation (1.6). Special features of the model allow us to show that $v(\cdot, z)$ is convex and that the quantity $\partial v/\partial y$ which appears in the dynamic programming equation exists and is continuous. The optimal feedback production law $p^*(y, z)$ is expressed as a function of $\partial v/\partial y$ by formula (4.3). We do not know that $p^*(\cdot, z)$ is Lipschitz continuous. However, since $p^*(\cdot, z)$ is a nonincreasing function of y, the differential equation

$$\frac{dy^*}{dt} = p^*(y^*(t), z(t)) - z(t), \qquad y^*(0) = y,$$

has a unique solution for the optimal inventory level $y^*(t)$.

---

* Received by the editors October 28, 1985; accepted for publication (in revised form) December 12, 1986.

† Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912. The work of this author was supported by the National Science Foundation through grant MCS-812940, by the Office of Naval Research under grant N00014-83-K-0542, and by the Air Force Office of Scientific Research under grant AFOSR-810116-C.

‡ Faculty of Management Studies, University of Toronto, Toronto, Ontario, Canada M5S IV4. The work of this author was supported by the Connaught Senior Research Fellowship and by the Natural Sciences and Engineering Research Council of Canada through grant A4619.

§ Department of Mathematics, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213. The work of this author was partly supported by the Air Force Office of Scientific Research under grant AFOSR-810116-C, and partly by the Institute for Mathematics and Its Applications with funds provided by the National Science Foundation and the Office of Naval Research.

We begin in § 1 by formulating a more general class of discounted optimal control problems with Markov chain parameter $z(t)$. It is elementary that the value function is convex in the state y, provided the state dynamics are linear in state y and control p and the cost criterion is convex jointly in (y, p); see Lemma 1.1. In § 2 we find that the value $v(y, z)$ is the unique solution to the dynamic programming equation, and that the gradient $\nabla_y v$ is continuous. For the continuity of $\nabla_y v$, an additional assumption (2.2) on the Hamiltonian appearing in the dynamic programming equation is needed.

In § 3 we discuss optimal controls, both from the viewpoint of dynamic programming and the theory of controlled piecewise deterministic processes. Under a strict convexity condition (3.1) on the cost criterion, there is a continuous optimal control policy $p^*(y, z)$ and the corresponding optimal control process $p^*(t)$ is unique.

These results are applied to the production planning model in § 4. Finally, in § 5 the analysis is modified to deal with a state-space constraint $y(t) \geqq y_{\min}$. Such a constraint imposes an inequality (5.2) on $\partial v/\partial y$ at $y_{\min}$.

The production planning model considered here does not impose any upper bound on the production rate. A more interesting extension of the problem involves production processes, which are bounded from above by a stochastic process representing the capacity of the production system. The capacity process over time may be modelled as a jump process or a piecewise deterministic process [5], [14]. Moreover, there may be several different products competing for a variety of scarce capacities. This is a problem faced by flexible manufacturing systems [10], upon which the methods developed in this paper have some bearing.

1. Discounted optimal control problems with Markov chain parameters. Let us begin with a model of the following rather general form, and then specialize. Let $y(t), p(t), z(t)$ denote, respectively, state, control and parameter processes for $t \geqq 0$. We assume that $y(t) \in R^n$, $p(t) \in K$, $z(t) \in Z$ for each $t \geqq 0$, where R^n is n-dimensional Euclidean space, K is a closed convex subset of some Euclidean space, $0 \in K$ and Z is a finite set. The parameter process $z(t)$ is a finite state continuous time Markov chain, defined on some underlying probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with jumping rate $q_{zz'}$ from state z to state z'. The associated generator L of the Markov chain $z(t)$ has the form

$$Lg(z) = \sum_{z' \neq z} q_{zz'}[g(z') - g(z)]. \tag{1.1}$$

In the general formulation, the state dynamics are $dy(t) = f(y(t), p(t), z(t))$, $t \geqq 0$. Actually we shall consider only f of the special form (1.5) below. A control process $P = \{p(t, \omega), t \geqq 0, \omega \in \Omega\}$ will be called *admissible* if: (i) P is adapted to $\mathcal{F}_t = \sigma(z(s): 0 \leqq s \leqq t)$; (ii) $\sup\{|p(t, \omega)|: t \geqq 0, \omega \in \Omega\} < \infty$; (iii) $p(t, \omega) \in K$ for all $t \geqq 0$ and $\omega \in \Omega$ (in whatever follows the ω-dependence will be supressed). Let $\mathcal{A}$ denote the set of admissible control processes.

We consider a cost criterion $l(y, p)$ about which the following assumptions are made:

$$\begin{aligned}
&\text{(a)} \quad l(\cdot, \cdot) \text{ is convex on } R^n \times K, \\
&\text{(b)} \quad -C \leqq l(y, p) \leqq C_N(1 + |y|^m) \text{ whenever } |p| \leqq N, \\
&\text{(c)} \quad \lim_{|p| \to \infty} l(y, p)|p|^{-1} = +\infty \text{ if } K \text{ is unbounded}
\end{aligned} \tag{1.2}$$

where m, C are fixed constants and C_N may depend on N. Let $\alpha > 0$ denote the discount rate. For every $P \in \mathscr{A}$ and $y = y(0)$, $z = z(0)$, let

$$J(y, z, P) = E \int_0^\infty e^{-\alpha t} l(y(t), p(t))\, dt. \tag{1.3}$$

The value function is

$$v(y, z) = \inf \{J(y, z, P) : P \in \mathscr{A}\}. \tag{1.4}$$

From now on, let us assume that the dynamics have the following special form:

$$dy(t) = [B(z(t))p(t) + C(z(t))]dt, \qquad t > 0. \tag{1.5}$$

(In the simple production planning model to be considered in § 4, both $y(t)$ and $p(t)$ are scalar valued and $B(z) = 1$, $c(z) = -z$.) Instead of (1.5) we could take equally well $dy(t) = [A(z(t))y(t) + B(z(t))p(t) + C(z(t))]dt$ provided that the eigenvalues of $A(z)$ have strictly negative real parts.

LEMMA 1.1. *For each $z \in Z$, $v(\cdot, z)$ is convex on R^n and $-C \leqq v(y, z) \leqq C(1 + |y|^m)$ for some $C > 0$.*

This lemma is easily proved, after observing that $J(\cdot, z, \cdot)$ is convex jointly in (y, P) for each $z \in Z$ and $p(t) \equiv 0$ is an admissible control process.

The dynamic programming equation associated with this optimal stochastic control problem is as follows:

$$\alpha v(y, z) = H(y, z, \nabla v(y, z)) + Lv(y, z), \quad y \in R^n, \quad z \in Z \tag{1.6}$$

where ∇v is the gradient in y

$$Lv(y, z) = [Lv(y, \cdot)](z) = \sum_{z' \neq z} q_{zz'}[v(y, z') - v(y, z)]$$

and for $y, z, r \in R^n \times Z \times R^n$

$$H(y, z, r) = \inf_{P \in K} [l(y, p) + (B(z)p + C(z)) \cdot r]. \tag{1.7}$$

Since Z is a finite set, (1.6) is a system of nonlinear first order PDE's in y, coupled through the zeroth order term Lv.

We are concerned with solutions to (1.6) belonging to the following space D_0.

DEFINITION 1.2. We say that a real-valued function v with domain $R^n \times Z$ is in D_0 if

(i) $v(\cdot, z)$ is convex on R^n for each $z \in Z$,

(ii) $-C \leqq v(y, z) \leqq C(1 + |y|^\beta)$, for suitable C and β (depending on v),

(iii) The gradient $\nabla v(y, z)$ is continuous.

The following "verification theorem" is standard, but for completeness we indicate the proof.

THEOREM 1.3. *Let $v \in D_0$ satisfy the dynamic programming equation* (1.6). *Then*

(a) *$v(y, z) \leqq J(y, z, P)$ for all $P \in \mathscr{A}$,*

(b) *Suppose that there are $P^* \in \mathscr{A}$, $y^*(t)$ that satisfies* (1.5) *with $y^*(0) = y$, $r^*(t) = \nabla v(y^*(t), z(t))$, and*

$$H(y^*(t), z(t), r^*(t)) = l(y^*(t), p^*(t)) + (B(z(t))p^*(t) + C(z(t)) \cdot r^*(t))$$

a.e. in t with probability 1. *Then*

$$v(y, z) = J(y, z, P^*).$$

Proof. For $T < \infty$, we have the usual dynamic programming relation

$$v(y, z) \leqq E \int_0^T e^{-\alpha t} l(y(t), p(t))\, dt + e^{-\alpha T} v(y(T), z(T)). \tag{1.8}$$

Since any admissible P is bounded, $|y(t)| \leqq c_1 t + c_2$ for suitable constants c_1, c_2. We obtain (a) as $T \uparrow \infty$, using the polynomial growth condition (ii) in Definition 1.2. In part (b), inequality (1.8) becomes an equality. □

In the next section we shall show that, under an additional condition (2.2) on H, the value function in fact belongs to D_0.

2. Viscosity solutions to the dynamic programming equation. The definition of viscosity solution used here is a straightforward generalization of the original definition given by M. G. Crandall and P.-L. Lions [4]. See also [3], [9] for more information.

Let v be a continuous function on $R^n \times Z$. For each (y, z) we define convex subsets $D_y^{\pm} v(y, z)$ of R^n, as follows:

$$D_y^+ v(y, z) = \{r \in R^n : \limsup_{h \to 0} (v(y+h, z) - v(y, z) - r \cdot h)|h|^{-1} \leqq 0\},$$

$$D_y^- v(y, z) = \{r \in R^n : \liminf_{h \to 0} (v(y+h, z) - v(y, z) - r \cdot h)|h|^{-1} \geqq 0\}.$$

We say that any continuous function v is a *viscosity solution* of (1.6), if for each y, z: (i) $\alpha v(y, z) \leqq H(y, z, r) + Lv(y, z)$ for all $r \in D_y^+ v(y, z)$, and (ii) $\alpha v(y, z) \geqq H(y, z, r) + Lv(y, z)$ for all $r \in D_y^- v(y, z)$.

Remark 2.1. v is differentiable in the y-direction at (y, z) if and only if $D_y^+ v(y, z)$ and $D_y^- v(y, z)$ are both singletons. In this case, the singleton is the gradient $\nabla v(y, z)$. Moreover, if v is convex in y, then $D_y^+ v(y, z)$ is empty unless v is differentiable there and $D_y^- v(y, z)$ coincides with the set of subdifferentials in the sense of convex analysis,

$$D_y^- v(y, z) = \overline{\mathrm{co}}\, \Gamma(y, z) \tag{2.1}$$

where

$$\Gamma(y, z) = \{r = \lim_{n \to \infty} \nabla v(y_n, z) : y_n \to y \text{ as } n \to \infty \text{ and } v(\cdot, z) \text{ is differentiable at } y_n\}$$

and where $\overline{\mathrm{co}}\, \Gamma$ denotes the convex closure of Γ. (See [2, Thm. 251, pp. 63].)

We now make the additional assumption that the Hamiltonian $H(y, z, \cdot)$ is constant on no nontrivial convex set:

$$\text{If } H(y, z, \lambda r_1 + (1-\lambda) r_2) = \text{constant in } \lambda \text{ for } 0 \leqq \lambda \leqq 1, \quad \text{then } r_1 = r_2. \tag{2.2}$$

THEOREM 2.2. *Let* (2.2) *hold and let v be a viscosity solution to the dynamic programming equation* (1.6). *If, in addition $v(\cdot, z)$ is convex for each z, then $\nabla v(y, z)$ exists for all (y, z) and $\nabla v(\cdot, z)$ is continuous on R^n.*

Proof. By Remark 2.1 and formula (2.1) it suffices to show that $D_y^- v(y, z)$ is a singleton. If $v(\cdot, z)$ is differentiable at y_n, then (1.7) holds at (y_n, z):

$$\alpha v(y_n, z) - H(y_n, z, \nabla v(y_n, z)) - Lv(y_n, z) = 0.$$

We then obtain, taking $y_n \to y$ as $n \to \infty$,

$$\alpha v(v, z) - H(y, z, r) - Lv(y, z) = 0 \quad \text{for } r \in \Gamma(y, z).$$

Moreover, $H(y, z, \cdot)$ is concave, and hence by (2.1)

$$\alpha v(y, z) - H(y, z, r) - Lv(y, z) \leqq 0 \quad \text{for } r \in D_y^- v(y, z).$$

However, the viscosity property implies the opposite inequality, and hence

$$\alpha v(y, z) - H(y, z, r) - Lv(y, z) = 0 \quad \text{for } r \in D_y^- v(y, z).$$

Thus, for fixed (y, z), $H(y, z, \cdot)$ is constant on the convex set $D_y^- v(y, z)$. By (2.2), $D_y^- v(y, z)$ is a singleton, which proves Theorem 2.2. □

The remainder of this section consists of a proof that the value function $v(y,z)$ defined by (1.4) is a viscosity solution to (1.6): the argument is rather standard.

LEMMA 2.3. *Suppose K is bounded. Then v is a viscosity solution to* (1.6).

Proof. It is a direct modification of Theorem 1.1 of [12]. □

THEOREM 2.4. *v is a viscosity solution to* (1.6).

Proof. Let $K_m=\{p\in K:|p|\leqq m\}$ and v_m be the optimal value function of the corresponding control problem. Then, v_m is a viscosity solution to (1.7) with $H(y,z,r)$ replaced with $H_m(y,z,r)=\min_{p\in K_m}\{r\cdot[B(z)p+c(z)]+l(y,p)\}$. Also, v_m converges to v uniformly on bounded subsets of R^n as m tends to infinity.

Take $r\in D_y^- v(y_0,z_0)$. For each $\varepsilon>0$, let $\varphi^\varepsilon(y,z)=\varphi(y,z)-\varepsilon(y-y_0)^2$ where $\varphi(y,z_0)=v(y_0,z_0)+r(y-z_0)$, and $\varphi(y,z)=v(y,z)$ if $z\neq z_0$. Since v is convex, the map $y\mapsto v(y,z_0)-\varphi^\varepsilon(y,z)$ has a strict maximum at y_0. Therefore, $y\mapsto v_m(y,z_0)-\varphi^\varepsilon(y,z_0)$ has a maximum at y_m, and y_m converges to y_0 as m tends to infinity. But this implies that $\nabla\varphi^\varepsilon(y_m,z_0)\in D_y^- v_m(y_m,z_0)$ and the viscosity property of v_m implies that

$$\alpha v_m(y_m,z_0)\geqq H_m(y_m,z_0,r-2\varepsilon(y_m-y_0))+Lv_m(y_m,z_0).$$

Now send m to infinity, and then ε to zero in the above inequality, to obtain

$$\alpha v(y_0,z_0)\geqq H(y_0,z_0,r)+Lv(y_0,z_0)\quad \text{for all } r\in D_y^- v(y_0,z_0).$$

The reversed inequality for $r\in D_y^+ v(y,z)$ is proved by a similar argument. □

3. Optimal controls. Let us now assume the following stricter form of convexity for the cost criterion l than what was assumed in (1.2)(a)

$$\begin{aligned}(3.1)\qquad & l(\lambda y_1+(1-\lambda)y_2,\lambda p_1+(1-\lambda)p_2)\\ &=\lambda l(y_1,p_1)+(1-\lambda)l(y_2,p_2)\quad \text{for some } 0<\lambda<1 \text{ implies } p_1=p_2.\end{aligned}$$

For example, for the production planning problem that will be considered in § 4, $l(y,p)=c(p)+h(y)$ and (3.1) holds if h is convex on R^n and l is strictly convex on K. Assumption (3.1) also holds, if $l(\cdot,\cdot)$ is convex and the second derivative of l in p exists and is positive at each (y,p).

Condition (3.1) implies, in particular, strict convexity of $l(y,\cdot)$, by taking $y=y_1=y_2$. This fact together with the superlinear *growth* condition (1.2)(c) imply that the minimum in (1.7) is attained at a unique $P=\Phi(y,z,r)$. Moreover, Φ is continuous on $R^n\times Z\times R^n$. Consider the control policy

$$(3.2)\qquad p^*(y,z)=\Phi(y,z,\nabla v(y,z)),$$

where v is the value function. By Theorems 2.2 and 2.4, p^* is continuous. The differential equation,

$$(3.3)\qquad dy(t)=[B(z(t))p^*(y(t),z(t))+C(z(t))]\,dt,$$

has locally a solution $y^*(t)$. Let us assume the following:

(3.4) There exists a bounded solution $y^*(t)$ to (3.3) for $t\geqq 0$.

In § 4 we shall verify (3.4) in the production planning example. The control process $P^*=\{p^*(t);\ t\geqq 0\}$, where

$$(3.5)\qquad p^*(t)=p^*(t^*(t),z(t))$$

is admissible, by the superlinear growth condition (1.2)(c) and is optimal, by the verification Theorem 1.3(b). Also, a straightforward application of (3.1) yields that P^* is unique. This implies, in particular, uniqueness of $y^*(t)$. We sum these results into the following proposition.

PROPOSITION 3.1. *Let* (3.1) *and* (3.4) *hold. Suppose* $P \in \mathscr{A}$ *and* $J(y, z, P) = v(y, z)$. *Then with probability* 1, $P(t) = P^*(t)$ *for almost all* $t > 0$. *In particular, there is a unique solution to* (3.3).

Proof. For $0 < \lambda \leqq 1$, let $P^\lambda = \lambda P + (1-\lambda)P^*$, $y^\lambda = \lambda y + (1-\lambda)y^*$, where $y(t)$ is a solution to (1.5) corresponding to P, with $y(0) = y$. By convexity of l and J, $J(y, z, P^\lambda) = v(y, z)$ which implies with probability one

$$l(y^\lambda(t), p^\lambda(t)) = \lambda l(y(t), p(t)) + (1-\lambda) l(y^*(t), p^*(t))$$

for almost all $t \geqq 0$. Assumption (3.1) then implies, with probability one, $p(t) = p^*(t)$ for almost all $t \geqq 0$. $\square$

Remark 3.2. The optimal policy P^* was obtained by the method of dynamic programming. The theory of piecewise deterministic processes [5], [14] provides an alternate approach. In the present context, the piecewise deterministic theory considers bounded, Borel measurable functions $\pi = [0, \infty) \times R^n \times Z \to K$. Given initial data $y(0) = y$, $z(0) = z$, each such π determines an admissible control process P as follows. Let $\tau_0 = 0$ and $\tau_1 < \tau_2 < \cdots$ denote the successive jump times of the Markov chain $z(t)$ and let

$$p(t) = \pi(t - \tau_i, y(\tau_i), z(\tau_i^+)), \qquad \tau_i < t \leqq \tau_{i+1}$$

where $y(t)$ is determined by solving (1.5) successively on each interval $[\tau_i, \tau_{i+1}]$. An optimal π^* is found as follows. For fixed z, as in (3.4), assume that there is a solution to

$$d\tilde{y}(t) = [B(z)p^*(\tilde{y}(t), z) + C(z)]\, dt, \tag{3.6}$$

with $\tilde{y}(0) = y$. Let

$$\pi^*(t, y, z) = p^*(\tilde{y}(t), z). \tag{3.7}$$

We claim that $\pi^*(\cdot, y, z)$ is unique, almost everywhere on $[0, \infty)$, for each y, z. This can be seen by slightly modifying the uniqueness proof above. We write the dynamic programming equation (1.6) as follows. Let

$$q_z = \sum_{z' \neq z} q_{zz'}, \qquad v_1(y, z) = \sum_{z' \neq z} q_{zz'} v(y, z').$$

Then (1.6) becomes

$$(\alpha + q_z) v(y, z) = H(y, z, \nabla v) + v_1(y, z). \tag{3.8}$$

For fixed z, (3.8) is the dynamic programming equation for a discounted deterministic control problem, with dynamics (3.6), discount factor $\alpha + q_z$ and cost criterion $l(y, p) + v_1(y, z)$. As before, $p^*(\cdot, z)$ is an optimal feedback control and $\pi^*(\cdot, y, z)$ determined by (3.7) is the unique optimal (open loop) control.

4. Production planning problem. Let us return to the model mentioned in the Introduction. We now have the following:

$$\begin{aligned} y(t) &= \text{inventory level at time } t \quad (y(t) \in R), \\ p(t) &= \text{production rate at time } t \quad (p(t) \geqq 0), \\ z(t) &= \text{demand rate at time } t \quad (z(t) \in Z). \end{aligned}$$

In the notation of § 1, we now have $n = 1$, $K = [0, \infty)$. The demand process is a finite state Markov chain, with state space $Z = \{z_1, \cdots, z_M\}$. The dynamics are as follows:

$$dy(t) = [p(t) - z(t)]\, dt.$$

Thus, in (1.5), $B(z)=1$ and $C(z)=-z$. It is assumed that $z_i>0$ for all $i=1,\cdots,M$. We assume that the cost criterion has the form

$$l(y,p)=h(y)+c(p)$$

and we seek to minimize

$$J(y,z,P)=E\int_0^\infty e^{-\alpha t}[h(y(t))+c(p(t))]\,dt. \tag{4.1}$$

The following assumptions are made about the holding cost h and production cost c:

(A1) h is convex, nonnegative on $(-\infty,\infty)$ with $h(0)=0$.

(A2) c is twice continuously differentiable, nonnegative on

$$[0,\infty) \text{ with } c'(0)=c(0)=0 \text{ and } c''(p)>0 \text{ for } p>0.$$

(A3) $$C(|y|^\beta-1)\leqq h(y)\leqq C(|y|^\gamma+1)\quad \text{for all } y\in R.$$

(A4) $$C(|p|^\nu-1)\leqq c(p)\quad \text{for all } p\geqq 0,$$

where $C>0$ and $\gamma,\beta,\nu>1$ are fixed constants.

The Hamiltonian H in (1.7) now takes the form $H(y,z,r)=F(r)-zr+h(y)$ where

$$F(r)=\min_{p\geqq 0}[pr+c(p)]. \tag{4.2}$$

The assumption (2.2) is satisfied since $z>0$ for all $z\in Z$, $F(r)$ is strictly concave for $r<0$ and $F(r)=0$ for $r\geqq 0$. Theorems 2.2 and 2.4 imply that the value function $v(y,z)$ belongs to the class D_0 and is the unique viscosity solution to the dynamic programming equation.

The optimal feedback production policy is now given by

$$P^*(y,z)=\begin{cases}(c')^{-1}\left(-\dfrac{\partial}{\partial y}v(y,z)\right) & \text{if } \dfrac{\partial}{\partial y}v(y,z)>0,\\ 0 & \text{if } \dfrac{\partial}{\partial y}v(y,z)\leqq 0.\end{cases} \tag{4.3}$$

Since v is convex in y, and $(c')^{-1}$ is an increasing function, p^* is nonincreasing in y. Therefore, the differential equation

$$dy(t)=[p^*(y(t),z(t))-z(t)]\,dt \tag{4.4}$$

has a unique solution $y^*(t)$ (see [8, Thm. 6.2].)

In the rest of the section, we shall show that y^* satisfies (3.4).

LEMMA 4.1. *There is a constant C, depending only on the initial condition $y^*(0)=y$, such that $|y^*(t)|\leqq C$ for all $t\geqq 0$.*

Proof. Let $\bar{y}=\sup\{y\in(-\infty,\infty): p^*(y,z)\geqq z \text{ for some } z\in Z\}$. Since v is convex in y and is nonnegative, (4.3) implies that $\bar{y}$ is finite. Similarly, let $\tilde{y}=\inf\{y\in(-\infty,\infty): p^*(y,z)\leqq z \text{ for some } z\in Z\}$. Suppose that $\tilde{y}$ is not finite. Then, there is $z\in Z$ such that $(\partial/\partial y)v(y,z)\geqq -c'(z)$ for all $y\in R$. But this contradicts with Lemma 4.2, which follows. Now one completes the proof of the lemma, by observing that $[\tilde{y},\bar{y}]$ is an attracting set for the differential equation (4.4). We refer to this set as the *turnpike set* in [7]. □

Let $\bar{z}=\max\{z\in Z\}$.

LEMMA 4.2. *For each y, z, $v(y,z)\geqq C(|y|-1)$ for a suitable constant $C>c'(\bar{z})$.*

Proof. Let $\bar{v}(y)$ be the value function of the following variational problem:

$$\bar{v}(y)=\inf\left\{\int_0^\infty e^{-\alpha t}\left[h(y(t))+C\left|\frac{d}{dt}y(t)\right|^\nu\right]dt;\ y(0)=y \text{ and } y(\cdot)\in W^{1,\infty}\right\}$$

with $\nu > 1$. In view of (A4), $\bar{v}(y) \leqq v(y, z)$ for a suitable $C > 0$. It suffices to show that $\bar{v}(y)|y|^{-1}$ converges to $M > \sup\{c'(z): z \in Z\}$ as y tends to $-\infty$. Since $\bar{v}(\cdot)$ is convex, $M = -\lim_n (d/dy)\bar{v}(y_n)$ where $\{y_n\}$ is any sequence which converges to $-\infty$ and $\bar{v}(\cdot)$ is differentiable at y_n, invoke Theorem 2.4 to conclude

$$\alpha \bar{v}(y_n) = \bar{F}\left(\frac{d}{dy}\bar{v}(y_n)\right) + h(y_n) \tag{4.5}$$

where $\bar{F}(r) = \sup\{rp + c|p|^\nu: -\infty < p < \infty\} = -\bar{c}|r|^{\nu/(\nu-1)}$. The positivity of $\bar{v}$ yields $M \in [0, \infty]$. Suppose that $M < \infty$. Then divide (4.5) by $|y_n|$ and pass to the limit to obtain

$$\alpha M = \alpha \lim_n \bar{v}(y_n)|y_n|^{-1} = \lim_n h(y_n)|y_n|^{-1} = \infty.$$

Hence $M = \infty$ and the proof of the lemma is complete. □

5. Inventory constraints. In this section, in addition to the nonnegative production constraint earlier, we impose the constraint that the inventory level cannot fall below a certain prescribed level $y_{\min}$. For each $y, z \in [y_{\min}, \infty) \times Z$, the set of admissible production processes $\mathscr{A}(y, z)$ is given by

$$\mathscr{A}(y, z) = \left\{P \in \mathscr{A}: y + \int_0^t [p(s) - z(s)]\, ds \geqq y_{\min} \text{ for all } t \geqq 0\right\}.$$

Then the corresponding value function is

$$v(y, z) = \inf\{J(y, z, P): P \in \mathscr{A}(y, z)\}.$$

The following characterization of v is a straightforward analogue of Theorem 1.1 of [12].

THEOREM 5.1. *The value function v for the constrained problem is in D_0 and is the only solution to the following equation*:

$$\alpha v(y, z) = H\left(y, z, \frac{\partial}{\partial y} v(y, z)\right) + Lv(y, z), \qquad y, z \in [y_{\min}, \infty) \times Z, \tag{5.1}$$

$$\frac{\partial}{\partial y} v(y_{\min}, z) \leqq -c'(z), \qquad z \in Z. \tag{5.2}$$

Proof. The first two conditions in the definition of D_0 are easily verified after observing that if $P \in \mathscr{A}(y, z)$ and $\hat{P} \in \mathscr{A}(\hat{y}, z)$, then $\frac{1}{2}P + \frac{1}{2}\hat{P} \in \mathscr{A}(\frac{1}{2}y + \frac{1}{2}\hat{y}, z)$ and $J(\cdot, z, \cdot)$ is convex for each $z \in Z$.

Repeating the proofs of Theorems 2.2 and 2.4 we show that v is continuously differentiable in the y-variable on $(y_{\min}, \infty)$ and satisfies (5.1). Define $(\partial/\partial y)v(y_{\min}, z)$ as the limit of $(\partial/\partial y)v(y, z)$ as y approaches to $y_{\min}$ from above (this limit exists due to the convexity of v in y). Now proceed as in Lemma 2.3 and use the fact that for any $P \in \mathscr{A}(y, z)$ the corresponding inventory level $y(t)$ is no less than $y_{\min}$, to obtain:

$$\alpha v(y_{\min}, z) \geqq H(y_{\min}, z, r) + Lv(y_{\min}, z) \quad \text{for } r \leqq \frac{\partial}{\partial y} v(y_{\min}, z) \tag{5.3}$$

(also, see Theorem 1.1 of [12]). Equations (5.1) and (5.3) yield

$$H\left(y_{\min}, z, \frac{\partial}{\partial y} v(y_{\min}, z)\right) \geqq H(y_{\min}, z, r) \quad \text{for } r \leqq \frac{\partial}{\partial y} v(y_{\min}, z). \tag{5.4}$$

The inequality (5.2) follows from (5.4), after observing that the map $r \to H(y_{\min}, z, r)$ achieves its maximum only at $r = -c'(z)$.

Uniqueness follows from the verification theorem, by observing that the optimal feedback policy P^* constructed in (3.5) is admissible on account of (5.2). □

Remark 5.2. For each $\varepsilon > 0$, define $h^\varepsilon(y) = h(y) + [(1/\varepsilon) - 1] \max\{y_{\min} - y, 0\}$. Let v^ε be the value function of the unconstrained problem with inventory cost h^ε. Then, the following estimate is proved in [7]:

$$0 \leqq v(y, z) - v^\varepsilon(y, z) \leqq \sqrt{\varepsilon} K_R \quad \text{for } y, z \in [y_{\min}, R] \times Z$$

where K_R is a suitable constant.

REFERENCES

[1] A. BENSOUSSAN, S. P. SETHI, R. VICKSON AND N. DERZKO, *Stochastic production planning with production constraints*, this Journal, 22 (1984), pp. 920-935.

[2] F. CLARKE, *Optimization and Non-smooth Analysis*, Wiley-Interscience, New York, 1983.

[3] M. G. CRANDALL, L. C. EVANS AND P.-L. LIONS, *Some properties of viscosity solutions of Hamilton-Jacobi equations*, Trans. AMS, 282 (1984), pp. 487-502.

[4] M. G. CRANDALL AND P.-L. LIONS, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. AMS, 277 (1983), pp. 1-42.

[5] M. H. A. DAVIS, *Piecewise deterministic Markov processes: A general class of non-diffusion stochastic model*, J. Royal Statist. Soc. Ser. B, 46 (1984), pp. 353-388.

[6] W. H. FLEMING AND R. RISHEL, *Deterministic and Stochastic Optimal Control*, Springer, New York, 1975.

[7] W. H. FLEMING, S. P. SETHI AND H. M. SONER, *Turnpike sets in optimal stochastic production planning problems*, LCDS Report No. 86-2, Brown University, Providence, RI, 1986.

[8] P. HARTMAN, *Ordinary Differential Equations*, Birkhauser-Verlag, Basel, Switzerland, 1982.

[9] P.-L. LIONS, *Generalized Solutions of Hamilton-Jacobi Equations*, Research Notes in Mathematics, 69, Pitman, Boston, 1982.

[10] J. G. KIMENIA AND S. B. GERSHWIN, *An algorithm for computer control of production in flexible manufacturing systems*, IIE Transactions, 15 (1983), pp. 353-362.

[11] S. P. SETHI AND G. L. THOMPSON, *Applied Optimal Control: Applications to Management Science*, Nijhoff, Boston, 1981.

[12] H. M. SONER, *Optimal stochastic control with state-space constraint*, II, this Journal, 24 (1986), pp. 1110-1123.

[13] G. L. THOMPSON AND S. P. SETHI, *Turnpike horizons for production planning*, Management Sci., 26 (1980).

[14] D. VERMES, *Optimal control of piecewise deterministic processes*, Stochastics, 14 (1985), pp. 165-208.