On piecewise deterministic Markov control processes: control of jumps and of risk processes in insurance.

Manfred Schäl
Inst. Angew. Math., Univ. Bonn, Wegelerstr. 6, D–53115 Bonn.

Published in: Insurance: Mathematics and Economics **22** (1998) 75–91.

Abstract. Dynamic programming for piecewise deterministic Markov processes is studied where only the jumps but not the deterministic flow can be controlled. Then one can dispense with relaxed controls. There exists an optimal stationary policy of feedback form.
Further, a piecewise deterministic Markov model for the control of dividend pay–out and reinsurance is introduced. This model can be transformed to a model with uncontrolled flow. It is shown that a classical solution to the Bellman equation exists and that a non–relaxed optimal policy of feedback form can be obtained via the Bellman equation. Lipschitz continuity of the 1–dim. vector field defining the controlled flow will be replaced by strict positivity.

Keywords: piecewise deterministic Markov processes, dynamic programming, non–relaxed controls, reinsurance, dividend pay–out, Bellman equation.

1. Introduction

Piecewise deterministic processes (PDPs) form a class of time–homogenous Markov processes living on a Borel subset E of $\mathbb{R}^d$. (In the insurance model of section 2 we will have d=1.) Davis (1984) calls this class a general class of non–diffusion stochastic models. This name can be justified by the following result of Cinlar & Jacod (1981):

(1.0) Every strong Markov process $\{X_t,\ t\geq 0\}$ living on E with paths that are both continuous and of (locally) bounded variation is deterministic, i.e.,
$X_t = \varphi(t\,|\,X_0)$ where φ is a **deterministic flow**.

A deterministic flow φ just defines a deterministic time–homogenous Markov process, i.e.

(1.1) $\varphi : [0,\infty)\times E \mapsto E$, $\varphi(0\,|\,x) = x$, $\varphi(t+s\,|\,x) = \varphi(s\,|\,\varphi(t,x))$, $s,t\geq 0$, $x\in E$,

where $\varphi(t\,|\,x)$ is continuous in t. Hence, in order to obtain non–trivial Markov processes on $\mathbb{R}^d$, one has to allow for paths having infinite total variation or one has to consider jumps. The first possibility leads to the study of diffusion processes arising as solutions of stochastic differential equations. Davis (1993) says in the preface: It is hard to deny that the stochastic differential equation model has received more than its fair share of attention.

If one wants to stick to paths of (locally) bounded variation, one has to allow for jumps. In view of applications, an appealing assumption is to exclude explosions of jumps, i.e. to assume that the jumps occur at randon times

(1.2) $0 < T_1 < \ldots < T_{n-1} < T_n \to \infty$.

According to Cinlar & Jacod (1981), $\{X_t\}$ is necessarily described by a deterministic flow between two jump times $T_n < T_{n+1}$. Thus, these considerations lead to the class of PDPs in a very natural way. Let us now concentrate on the case where $E \subset R$ (i.e. d=1) as in the insurance model. We will consider a flow φ which is defined as the unique solution to an Initial Value Problem (in the sense of Carathéodory) defined for an autonomous differential equation with a (here one–dimensional) vector field b on E :

(1.3) $\varphi(t) = \varphi(t\,|\,x) = x + {}_0\int^t b(\varphi(s))\, ds.$

Then we can write

(1.4) $dX_t = b(X_t)\, dt + \Delta X_t$ where $\Delta X_t = X_t - X_{t-0}$.

We can compare this approach with that defined by a stochastic differential equation

(1.5) $dX_t = b(X_t)\, dt + \sigma(X_t)\, dW_t$

where $\{W_t\}$ is a Wiener process (standardized Brownian motion). Obviously in (1.4), the 'white noise' perturbations dW_t in (1.5) are replaced by random jumps which however need not be compensated. Hence (1.4) provides not yet a semi–martingale decomposition. For such a decomposition, the compensator of the jumps has to be taken into account.

Davis (1993) assumes that the process will jump if it hits the boundary of the state space. But there are other interesting cases. In the present paper, the topological boundary will not play an exta rôle. In the insurance model, we will have $E = [0,\infty) \cup \{-1\}$ where -1 is an absorbing (cemetary) state and the process leaves the boundary state 0 by means of the flow. Deshmukh & Pliska (1980) consider an optimal consumption model where $S = [0,\infty)$ and the boundary point 0 is sticky, i.e. the process will wait in 0 until the next jump occurs.

When PDPs were introduced, it was soon discovered that the model and the developped techniques are important for risk theory, insurance science [cf. Embrechts (1984)]. Then, there followed a series of papers [cf. Dassios & Embrechts (1989), Embrechts (1990), Davis (1993, (21.12)), Embrechts & Schmidli (1994), Davis & Vellekoop (1995)]. Control in insurance was studied by Martin–Löf (1994) in a discrete–time framework. Recently, diffusions models were developped for the control of dividend pay–out and reinsurance [cf. Asmussens & Taksar (1997), Hϕjgaard & Taksar (1996), (1997), Taksar & Zhou (1997)].

It is known from deterministic control theory that one has to introduce the concept of relaxed controls in order to get optimal controls. Therefore, it was natural to introduce relaxed controls for the control of PDPs [cf. Davis (1993, § 43)]. Yushkevich (1987) derived the Bellman equation for the control of PDPs without use of relaxed controls; however he only studied the case where one can only control the jumps and not the deterministic flow. We will combine the results of Yushkevich with the investigation about continuity and compactness properties by Davis (1986), (1993) in the framework of relaxed controls and will obtain a result for the existence of a nonrelaxed optimal policy for the control of the jumps in §3. The optimal policy can be obtained as a feedback control via a classical solution to the Bellman equation. In terms of Davis (1993, § 42), naive dynamic programming works for the control of jumps.

In general, one has to generalize the concept of a solution to the optimality equation because one does not know whether the value function is sufficiently regular. There, one can use the Clarke generalized gradient and non–smooth analysis. In the presence of strong convexity properties it can be shown that one can dispense with relaxed controls [cf. Davis (1993, § 45), Dempster (1991), Dempster & Ye (1990, 1992, 1995)]. A second direction is the use of the concept of a viscosity solution [cf. Soner (1986), Fleming & Soner (1993)].

In special models, special techniques can be used to show that naive dynamic programming works [cf. Davis et al. (1987)]. In the present paper we will study a one–dimensional insurance control model as another special model. There we will use a property which is typical for a large class of PDPs: there are a deterministic flow in one direction and jumps in the other direction. This property enables us to make a transformation to a model with an uncontrolled flow. A similar transformation is used by Pliska (1978), Deshmukh & Pliska (1980), Soner (1985). For the transformed model, we can then use the results of §3. There are some other properties of the insurance model which will cause some technical problems: e.g. the vector field describing the flow is unbounded. We will get a classical solution to the optimality equation which is absolutely continuous which is the most natural regularity property. The result on the insurance model can be used as a first step to investigate qualitative properties of the value function and of the policy of an insurance company for the optimal reinsurance and distribution of dividends.

2. Control of an insurance model

The real–valued process $\{X_t,\ t\geq 0\}$ describes the **surplus process** (fund of reserves). There is a **premium (income) rate** c which is fixed. The process can be controlled by the choice of the **dividend rate** d and the **premium rate** p paid by the insurer to a reinsurer, hence the **action** consists in the choice of $a = (d,p)$. We assume

2.1 Assumption. $0 \leq d \leq \bar{d},\ 0 \leq p \leq \bar{p}$ for some upper bounds $\bar{d}, \bar{p}$ with $\bar{d} + \bar{p} < c$;

hence the **action space** is $A := [0,\bar{d}]\times[0,\bar{p}]$.

The insurance company gets interest for capital above a certain level $L \geq 0$, the amount of capital the company retains as a liquid reserve (cf. Embrechts & Schmidli (1994)). The corresponding **interest rate** is $\gamma \geq 0$, while we denote the general **discount rate** by $\beta > 0$. There, β and γ may conincide or not. We choose the **state space** E according to

2.2 Assumption. $E = [0,\infty) \cup \{-1\}$ where -1 represents the state of ruin.

[We could also choose $E = [0,\infty)$ with 0 as the state of ruin, but then we had to assume that 0 is an isolated point. As a consequence we would have to work with a topology which does not agree with the usual topology on $\mathbb{R}$ which fact is less convenient.]
The jumps at T_n, $n\geq 1$, are caused by the claims. Between the jumps, there is a deterministic flow which will be described by

2.3 Assumption.
$$b(x,a) := c - d - p + \gamma\cdot(x - L)^+ ,\ x \geq 0,$$
$$b(-1,a) := 0, \qquad \text{for } a = (d,p),$$

where b is now also a function of the action in the controlled case. Obviouly, we have

(2.4) $\quad b(x,a)$ is Lipschitz continuous in $x \geq 0$ uniformly in a.

It is remarkable that b(x,a) is unbounded in x. This property appears here in a natural way, but will cause some technical difficulties. Davis (1993, (41.1)) assumes in the controlled case that b is bounded.

Given the history H_n at T_n, the surplus at T_n+t ($< T_{n+1}$) is deterministic. Therefore one can decide at T_n about the **action** $a_{T_n+t} = u(t) = (d(t),p(t))$ at $T_n+t \leq T_{n+1}$ where

(2.5) $\quad$ the **control function** $u : [0,\infty) \mapsto A$ is measurable

and may depend on H_n. Then the flow $\varphi = \varphi^u(t\,|\,x)$ is a solution to

(2.6) $\varphi(t) = \varphi^u(t|x) = x + {}_0\!\int^t b(\varphi(s),u(s))\,ds\ ,\ x \geq 0,$

$\varphi^u(t|-1) = -1\ ,\ t \geq 0$, i.e.

$\frac{d}{dt}\varphi(t) = c - d(t) - p(t) + \gamma\cdot(\varphi(t) - L)^+$ for almost all t and for $x \geq 0$.

Because of (2.1), $\varphi^u(t|x)$ is strictly increasing in t for $x \geq 0$. We have

(2.7) $X_{T_n+t} = \varphi^u(t|X_{T_n})\ ,\quad T_n+t < T_{n+1}.$

By (2.4) and Carathéodory's theorem in ordinary differential equations [cf. Warga (1972) Theorems II.41, II.4.2] there is a unique solution φ to (2.6). Of special interest are feedback controls $u(t) = \delta(X_{T_n+t})\ ,\ T_n < T_n+t \leq T_{n+1}$ where

(2.8) the **feedback control function** $\delta : E \mapsto A$ is measurable.

This will lead to the autonomous differential equation in the sense of Carathéodory:

(2.9) $\varphi(t) = \varphi^\delta(t|x) = x + {}_0\!\int^t b(\varphi(s),\delta(\varphi(s)))\,ds$.

In the prevailing opinion, one needs at least locally Lipschitz continuity to guarantee existence and uniqueness of the solution φ which would put a severe restriction on the control δ. Here however, we are lucky and can use a special property of the insurance model, namely that b has only one sign and is bounded away from zero. In fact, we have a deterministic flow in one direction and jumps in the other direction. For an arbitrary feedback control function δ, there exists a unique solution to (2.7) according to the following theorem.

2.10 Theorem. Let g be a measurable function $g : [x_o,\infty) \mapsto \mathbb{R}$ such that

$0 < \varepsilon \leq g(x) \leq \alpha + \gamma\cdot x\ ,\ x\geq x_o$, for some $\varepsilon,\alpha,\gamma \geq 0$, $x_o \in \mathbb{R}$.

Then there is a unique solution $\varphi : [0,\infty)\times[x_o,\infty) \mapsto \mathbb{R}$ to

$\varphi(t|x) = x + {}_0\!\int^t g(\varphi(s|x))\,ds\ ,\quad t \geq 0,\ x \geq x_o,$

and φ is continous.

The proof is given in the Appendix. The case where g is negative can be treated by looking at $-\varphi$ and $-g(-x)$. Upon defining $u(t) := \delta(\varphi^\delta(t|x))$, we have a control function in the sense of (2.5) and φ^δ is a solution to (2.6).

As usual, the claim process is described by a compound Poisson process with rate λ and with claims of height Y_n at T_n where $Y_n \sim Q$ and Q is the **claim distribution**.

We will look upon the claims Y_n as disturbances which form an iid sequence of random variables taking values in $D = [0,\infty)$. There is a **system function** f such that

(2.11) $X_{T_{n+1}} = f(X_{T_{n+1}-0}, a_{T_{n+1}}, Y_{n+1})$

where

2.12 Assumption. $f(x,a,y) = f(x,d,p,y) := \begin{cases} x - h(p,y) & \text{for } x - h(p,y) \geq 0 \\ -1 & \phantom{\text{for }} x - h(p,y) < 0 \end{cases}$.

Here, $0 \leq h(p,y) \leq y$ is the part of the claim y paid by the insurer where h(p,y) depends on the premium rate which is is paid when the claim occurs. Hence, $y - h(p,y)$ is the part paid by the reinsurer. In the case of an **excess of loss reinsurance**, with retention level $M(p) \geq 0$ we have: $h(p,y) = M(p) \wedge y$. In the case of a **proportional reinsurance** with retention level $0 \leq \alpha(p) \leq 1$ we have: $h(p,y) = \alpha(p) \cdot y$. In general we only need the

2.13 Assumption. h is continuous in p.

The effect of reinsurance on the probability of ultimate ruin is studied by Waters (1983) and Dickson & Waters (1996). It is a special feature of the insurance model that in general the system function f is not continuous in the action p in spite of (2.13). We define the gain rate $\bar{r}$ by

(2.14) $\bar{r}(x,a) = \bar{r}(x,d,p) = d \cdot \mathbf{1}_{[0,\infty)}(x) - \beta \cdot K \cdot \mathbf{1}_{\{-1\}}(x)$,

where K determines the fixed cost of ruin . If

(2.15) $\tau := \inf \{ t \geq 0, X_t < 0\}$

is the ruin time, then the **total discounted reward** is

(2.16) $\int_0^\infty e^{-\beta t}\, \bar{r}(X_t,a_t)\, dt = \int_0^\tau e^{-\beta t}\, d_t\, dt - e^{-\beta\tau} \cdot K$ where $a_t =: (d_t,p_t)$.

Thus we have modelled a fixed cost of amount K at the ruin time by a cost rate. A general transformation from fixed costs to cost rates is explained by Davis (1993, (31.16)), see also (2.20) below. When maximizing the total discounted reward, we want to minimize the fixed cost incurred by ruin and to maximize the total discounted dividends paid up to the ruin time.
Now we make a simple transformation to nonnegative rewards. Define

(2.17) $r(x,a) := \bar{r}(x,d,p) + \beta \cdot K = [d + \beta \cdot K] \cdot \mathbf{1}_{[0,\infty)}(x) \geq 0$, $a = (d,p) \in A$,

then we get

(2.18) $\int_0^\infty e^{-\beta t}\, r(X_t,a_t)\, dt = K + \int_0^\infty e^{-\beta t}\, \bar{r}(X_t,a_t)\, dt$.

Therefore it is equivalent to maximize the total discounted reward w.r.t. $\bar{r}$ and w.r.t. r. Obviously we have the following properties

(2.19) r(x,a) is bounded, nonnegative and continuous.

A **policy** π is a sequence $\pi = (\pi_n, n \geq 0)$ where $\pi_n(H_n,t)$ is an A–valued measurable function of t and the history H_n which in this model is given through $H_n = (T_1,Y_1,...,T_n,Y_n)$. Then π_n specifies the control function u(t) for the period $(T_k,T_{k+1}]$ according to $u(t) = \pi_n(H_n,t)$ and the action a_t is defined through

$$a_t := \pi_n(H_n,t-T_n) \quad \text{for } T_n < t \leq T_{n+1}.$$

A feedback control function δ defines a **stationary feedback policy** π^δ through

$$a_t = \delta(X_{t-0}) = \delta(\varphi^\delta(t-T_n \mid X_{T_n})) \ \text{ for } \ T_n < t \le T_{n+1} \ \text{ under } \pi^\delta,$$

where X_{T_n} is obviously a measurable function of H_n and φ^δ is defined in (2.9).
There is a probability space $(\Omega,\mathfrak{F},P)$ on which the random variables T_n and Y_n are defined and have the following martingale dynamics [cf. Bremaud (1981, p. 245), Davis (1993, §§26, 31), Last & Brandt (1995, 4.1.14)]:

(2.20) $$E[\,\vartheta(H_{k-1},T_k,Y_k)\,] = E[\,\int_{T_{k-1}}^{T_k}\int \vartheta(H_{k-1},s,y)\,Q[dy]\;\lambda\,ds\,]$$
for any measurable function ϑ bounded from below.

Given $\omega \in \Omega$, an initial capital x, and a policy π, the state and action processes are well–defined and we should write $\{X_t^{\pi.x}, t \ge 0\}$ and $\{a_t^{\pi.x}, t \ge 0\}$. For convenience, we will use the following notation for any function F of the trajectories:

(2.21) $$E_x^\pi[\,F(X_t,a_t,\,t\ge 0)\,] := E[\,F(X_t^{\pi,x},a_t^{\pi,x},\,t\ge 0)\,].$$

There can be given a probabilistic basis to this notation [cf. Davis (1993, §25)].

By use of (2.20) and the chain rule (cf. Appendix A1), one can prove the following identity:

(2.22) $$E_x^\pi[e^{-\beta t}\,v(X_t)] = v(x) + E_x^\pi[\int_0^t e^{-\beta s}\,\{v'(X_s)\cdot b(X_s,a_s) - \beta\cdot v(X_s) + \lambda\cdot\int [v(f(X_s,a_s,y)) - v(X_s)]\;Q[dy]\;\}\,ds\,]$$
for any bounded absolute continuous function v with bounded derivative v'.

The **value function** or **optimal reward function** V* is defined through

(2.23) $$V^*(x) := \sup_\pi V^\pi(x) \quad \text{where} \quad V^\pi(x) := E_x^\pi[\,\int_0^\infty e^{-\beta t}\,r(X_t,a_t)\,dt\,]\,.$$

Now we can state the main result for the insurance model:

2.24 Theorem. V* is bounded, absolutely continuous on $[0,\infty)$ with bounded derivative $\frac{d}{dx}V^*$ and satisfies the (**Hamilton–Jacobi**–) **Bellman equation** :

$$\sup_{0\le d\le\bar{d},\,0\le p\le\bar{p}} \Big\{ r(x,d,p) + \frac{d}{dx}V^*(x)\cdot b(x,d,p) + \lambda\cdot\int [V^*(f(x,d,p,y)) - V^*(x)]\;Q[dy]\; - \beta\cdot V^*(x)\Big\} = 0\,,\ x \ge 0,$$

$V^*(-1) = 0.$

The proof by use of the results of §3 and a transformation defined below is given in the Appendix.

2.25 Corollary. There exists a measurable function $\delta(x) = (d^*(x),p^*(x))$ such that the supremum in the Bellman equation is attained at $(d,p) = \delta(x)$ for $x \ge 0$. Each such function (with $\delta(-1) := 0$ say) defines an optimal stationary feedback policy, i.e.

$$V^{\pi^\delta}(x) = V^*(x) \quad \text{for } x \in E.$$

Proof. The existence of δ follows from the considerations for 3.18. Now, from the Bellman equation we conclude that under any policy π

$$\frac{d}{dx} V^*(X_s)b(X_s,a_s) - \beta \cdot V^*(X_s) + \lambda \cdot \int \left[V^*(f(X_s,a_s,y)) - V^*(X_s)\right] Q[dy] \le - r(X_s,a_s)$$

with equality under the stationary feedback policy π^δ. Then we obtain from (2.22):

$$E_x^\pi\left[e^{-\beta t} V^*(X_t)\right] \le V^*(x) - E_x^\pi\left[{}_0\int^t e^{-\beta s} r(X_s,a_s)\, ds\right] \text{ with equality under } \pi^\delta.$$

Passing to the limit $t\to\infty$, we finally obtain: $V^\pi(x) \le V^*(x)$ and $V^{\pi^\delta}(x) = V^*(x)$, $x \in E$. []

Upon dividing the Bellman equation by $b(x,a)$, which is bounded away from zero by (2.1), it can easily be seen that the Bellman equation is equivalent to

(2.26) $$\frac{d}{dx} V^*(x) + \sup_{a\in A} \left\{ \tilde{r}(x,a) + \tilde{\lambda}(x,a) \cdot \int \left[V^*(f(x,a,y)) - V^*(x)\right] Q[dy] - \tilde{\beta}(x,a) \cdot V^*(x) \right\} = 0 ,\ x \ge 0,$$

$$V^*(-1) = 0$$

where

(2.27) $$\tilde{r}(x,a) := r(x,a)/b(x,a),\ \tilde{\lambda}(x,a) := \lambda/b(x,a),\ \tilde{\beta}(x,a) := \beta/b(x,a),\ x \ge 0,$$

$$\tilde{r}(-1,a) := 0,\ \tilde{\lambda}(-1,a) := 0,\ \tilde{\beta}(-1,a) := \beta.$$

The equation (2.26) is the optimality equation of a controlled PDP where the flow $\tilde{\varphi}$ is

(2.28) $$\tilde{\varphi}(t\,|\,x) = t+x \quad \text{for } t\ge0,\ x\ge0, \text{ and } \tilde{\varphi}(t\,|\,-1) = -1$$

and hence uncontrolled, where the reward rate is $\tilde{r}$, the jump rate is $\tilde{\lambda}$ and the discount rate is $\tilde{\beta}$. Hence, in the transformed model, the jump rate and the discount rate depend both on the state and the action. Even worse, the discount rate $\tilde{\beta}$ is not bounded away from zero. Therefore we have to make sure that discounting in the transformed model is strong enough. We make use of the **bounding function**

(2.29) $$\eta(x) := \frac{1}{c + \gamma \cdot x},\ x \ge 0;\ \eta(-1) := 1,$$

where it is easy to see that η satisfies the following conditions:

(2.30) $${}_0\int^\infty \eta(\tilde{\varphi}(t\,|\,x))\, dt = \infty \quad,\ x \in E\ ;$$

(2.31) For $x_n \to x_o$ $\exists\, n_o$ such that: $\frac{1}{2} \le \eta(\tilde{\varphi}(t\,|\,x_n))/\eta(\tilde{\varphi}(t\,|\,x_o)) \le 2$ for $t \ge 0$, $n \ge n_o$.

Further, we will need the following property:

(2.32) $0 < C_b^{-1} \le b(x,a) \cdot \eta(x) \le 1$, $x \ge 0$, for some constant $C_b \ge 1$,

from which we obtain:

(2.33) $\beta \cdot \eta(x) \le \tilde{\beta}(x,a) \le C_b \cdot \beta \cdot \eta(x)$;

$\tilde{\lambda}(x,a) \le C_b \cdot \lambda \cdot \eta(x)$;

$0 \le \tilde{r}(x,a) \le (\bar{d} + \beta \cdot K) \cdot C_b \cdot \eta(x)$.

The following property will be used as a substitute of the continuity of f.

2.34 Lemma. If w is a nonnegative, bounded, and upper semi–continuous (u.s.c.) mapping on E with $w(-1) = 0$, then $(x,a) \mapsto w(f(x,a,y))$ is u.s.c. $\forall\, y \ge 0$.

Proof. We have $w(f(x,a,y)) = \mathbf{1}_{[0,\infty)}(x - h(p,y))\, w(x - h(p,y))$.
Now $(x,a) \mapsto x - h(p,y)$ is continuous $\forall\, y \ge 0$ and $v(z) := \mathbf{1}_{[0,\infty)}(z) \cdot w(z)$ is u.s.c. ; therefore $(x,a) \mapsto w(f(x,a,y)) = v(x-h(p,y))$ is u.s.c. $\forall\, y \ge 0$. $\Box$

3. Control of jumps

In this section we consider the case that the deterministic flow φ cannot be controlled Then we can rely on results of Yushkevich (1987). In this situation we can allow that the state space E is a Borel subset of any Polish space. Moreover, we will replace continuity of r as in (2.19) by the more general concept of upper semi–continuity. A real–valued function ϑ on some metric space Ξ is called upper semi–continuous (u.s.c.) if

$$\limsup_{\xi_n \to \xi} \vartheta(\xi_n) \le \vartheta(\xi) \quad \forall\, \xi \in \Xi.$$

Any u.s.c. function ϑ attains the supremum on compact sets while continuous functions attain both the supremum and the infimum on compact sets. We will make the following

3.1 Assumptions:

(a) the **flow** $\varphi(t\,|\,x)$ satisfies (1.1) and is a continuous mapping on $[0,\infty)\times E$;

(b) the **action space** A is a compact metric space;

(c) the **bounding function** is a continuous function $\eta : E \mapsto (0.\infty)$ such that

(i) $\int_0^\infty \eta(\varphi(t\,|\,x))\, dt = \infty$, $x \in E$;

(ii) for $x_n \to x_o$ $\exists\, n_o$ such that: $\frac{1}{2} \le \eta(\varphi(t\,|\,x_n))/\eta(\varphi(t\,|\,x_o)) \le 2$ for $t \ge 0$, $n \ge n_o$.

(d) the **jump rate** $\lambda(x,a)$ is a continuous function on $E\times A$ with:

$$\lambda(x,a) \le C_\lambda \cdot \eta(x) \text{ , } a \in A,\ x \in E, \text{ for some constant } C_\lambda;$$

(e) the **gain rate** $r(x,a)$ is an u.s.c. function on $E\times A$ with:

$$|r(x,a)| \le C_r \cdot \eta(x) \text{ , } a \in A,\ x \in E, \text{ for some constant } C_r;$$

(f) the **discount rate** $\beta(x,a)$ is a continuous function on $E\times A$ with:

$$\underline{C}_\beta \cdot \eta(x) \le \beta(x,a) \le \bar{C}_\beta \cdot \eta(x) \text{ , } a \in A,\ x \in E, \text{ for some constants } 0 < \underline{C}_\beta \le \bar{C}_\beta \text{ ;}$$

(g) the **disturbance distribution** Q is a probability on the **disturbance space** $(D,\mathfrak{D})$;

(h) the **system function** $f : E\times A\times D \mapsto E$ is a measurable mapping where $f(x,a,y)$ is continuous in (x,a) or more generally Assumption (3.14) below is satisfied.

We assume that the distribution Q of the disturbances Y_n does not depend on the previous states and actions and hence (Y_n) forms an iid sequence. This assumption is made for convenience and fulfilled in many examples like that of §2. In fact, there is no loss of generality in assuming that $D = [0,1]$ und Q is the uniform distribution on D. This fact was shown by Davis (1993 §§ 23 ,24). In the one–dimensional case $D \subset \mathbb{R}$, the reduction to the present case is achieved by defining the system function f by use of the generalized inverse of the distribution function of the (possibly conditional) distribution Q of Y_n.

The assumption that the discount rate β depends on the present state x and action a is needed for the application to the transformed insurance model. In fact, there would again be no loss of generality to assume that β is one, in particular independent of x,a; compare the transformations by Davis (via killing) (1993, (31.6)) and Yushkevich (1987, § 5). However, we want to work explicitely with the discount rate because this is more convenient for applications. As in Davis (1986), (1993) and Yushevich (1987) we will study an equivalent **discrete–time semi–Markov model**. For reasons of compactness of the corresponding action space, we have to deal with relaxed controls. However, these are only used for the proofs. One of the main results will be that the optimal policy can be chosen by use of a classical non–relaxed feedback control.

We write $\mathbb{P}(A)$ for the set of all probability measures on A which is known to be again a compact metric space when endowed with the topology of weak convergence of probability measures. Set

(3.2) $U := \{ u : [0,\infty) \mapsto A \text{ , u measurable } \}$,

$\hat{U} := \{ \hat{u} : [0,\infty) \mapsto \mathbb{P}(A) \text{ , } \hat{u} \text{ measurable } \}$.

Any $\hat{u} \in \hat{U}$ is called a relaxed control function. The measurability of $\hat{u}$ is equivalent to the fact that u can be considered as a transition probability from $[0,\infty)$ to A [cf. Bertsekas & Shreve (1978, 7.26), Rieder (1975)]. Relaxed controls though well–defined in a mathematical sense have the disadvantage that they cannot be applied not only from a practical but also from a measure theoretical point of view. In fact, in contrast to randomized controls in discrete time models, in general there cannot be constructed an action process (a_t) with values in A (but only in $\mathbb{P}(A)$) under relaxed controls.

The **Young–topology** on $\hat{U}$ is the coarsest topology such that the mappings

$$\hat{u} \mapsto \int_0^\infty \int_A \vartheta(t,a)\, \hat{u}(t;da)\, dt$$

are continuous for all real functions ϑ on $[0,\infty)\times A$, where ϑ is a Carathéodory function, i.e. ϑ is continuous in a and measurable in t , and where ϑ is integrable in the sense that

$$\int_0^\infty \max_{a\in A} |\vartheta(t,a)|\, dt < \infty.$$

The introduction of $\hat{U}$ is justified by the fact [cf. Davis (1993, (43.3)),Warga (1972)] that $\hat{U}$ is a compact space w.r.t the Young topogy, but U is in general not compact.

We have to define quantities, already defined for $a\in A$, also for $\mu\in\mathbb{P}(A)$:

(3.3) $\lambda(x,\mu) := {}_A\int \lambda(x,a)\, \mu[da]$,

$\lambda(x,\mu) \cdot {}_E\int v(z)\, q(dz|x,\mu) := {}_A\int \lambda(x,a) \left[{}_D\int v(f(x,a,y))\, Q[dy]\right] \mu[da]$,

$r(x,\mu) := {}_A\int r(x,a)\, \mu[da]$,

$\beta(x,\mu) := {}_A\int \beta(x,a)\, \mu[da]$,

$\Lambda(x,\hat{u},t) := {}_0\int^t \lambda(\varphi(s|x),\hat{u}(s))\, ds$,

$B(x,\hat{u},t) := {}_0\int^t \beta(\varphi(s|x),\hat{u}(s))\, ds$, $\hat{u} \in \hat{U}$.

From (3.1c(i),f), one immediately obtains:

(3.4) $B(x,\hat{u},\infty) = \infty$, $x \in E$, $\hat{u} \in \hat{U}$,

which fact will ensure that discounting is strong enough. In the equivalent discrete–time model, we will choose $\hat{U}$ as action space. The **discrete–time reward operator** is

(3.5) $Tv(x,\hat{u}) := E\left[{}_0\int^\tau \exp\{-B(x,\hat{u},s)\}\, r(\varphi(s|x),\hat{u}(s))\, ds + \exp\{-B(x,\hat{u},\tau)\} \cdot v(Z)\right]$

where τ is the first jump time and Z is the state immediately after τ. Then $Tv(x,\hat{u})$ describes the expected discounted reward up to the first jump if the process starts in x and the control function $\hat{u}$ is chosen for the first period and if a reward v is obtained immediately after the first jump τ depending on the actual state Z. We can write

$$\begin{aligned} Tv(x,\hat{u}) = {}&{}_0\int^\infty \exp\{-\Lambda(x,\hat{u},t)\}\, \lambda(\varphi(t|x),\hat{u}(t)) \left[{}_0\int^t \exp\{-B(x,\hat{u},s)\}\, r(\varphi(s|x),\hat{u}(s))\, ds\right] dt \\ &+ \exp\{-\Lambda(x,\hat{u},\infty)\} \cdot {}_0\int^\infty \exp\{-B(x,\hat{u},s)\}\, r(\varphi(s|x),\hat{u}(s))\, ds \\ &+ {}_0\int^\infty \exp\{-\Lambda(x,\hat{u},t)\}\, \lambda(\varphi(t|x),\hat{u}(t)) \exp\{-B(x,\hat{u},t)\} \cdot {}_E\int v(z)\, q(dz|\varphi(t|x),\hat{u}(t))\, dt \end{aligned}$$

where (3.4) is used. By an easy computation we get

(3.6) $$Tv(x,\hat{u}) = {}_0\int^\infty \exp\{-(\Lambda+B)(x,\hat{u},t)\} \cdot \left[r(\varphi(t|x),\hat{u}(t)) + \lambda(\varphi(t|x),\hat{u}(t))\, {}_E\int v(z)\, q(dz|\varphi(t|x),\hat{u}(t))\right] dt .$$

The **optimal reward operator** is

(3.7) $T^*v(x) := \inf_{\hat{u}\in\hat{U}} Tv(x,\hat{u})$.

3.8 Lemma Let $w : E\times A \mapsto \mathbb{R}$ be a function with $|w(x,a)| \le C_w \cdot \eta(x)$ for some $C_w<\infty$ and set

$w(x,\mu) := {}_A\int w(x,a)\, \mu[da]$, $\mu\in\mathbb{P}(A)$,

$W(x,\hat{u}) = {}_0\int^\infty \exp\{-(\Lambda+B)(x,\hat{u},t)\} \cdot w(\varphi(t|x),\hat{u}(t))\, dt$, $x \in E$, $\hat{u} \in \hat{U}$.

If $w(x,a)$ is continuous or u.s.c. in (x,a), then $W(x,\hat{u})$ is a bounded and continuous or u.s.c. function in $(x,\hat{u})$, respectively.

The proof is given in the Appendix.

3.9 Remark. Let $\hat{A}$ be any compact metric space and $R(x,\hat{a})$ a bounded function on $E\times\hat{A}$. If

$$R(x,\hat{a}) \text{ is continuous or u.s.c., resp., so is } \quad R^*(x) := \max_{\hat{a}\in\hat{A}} R(x,\hat{a}).$$

This fact is well–known [cf. Bertsekas & Shreve (1978) 7.32]. □

Let B(E) be the set of all bounded and upper semianalytic (u.s.a.) functions on E.
We will not give the definition of an u.s.a. function here. We need the fact that B(E) contains any bounded Borel–measurable function and that each $w \in B(E)$ is μ–integrable w.r.t. any probability $\mu \in \mathbb{P}(E)$ (more exactly w.r.t. the comletion of μ). A problem in dynamic programming is that $T^*(v)$ need not to be Borel–measurable for each bounded Borel–measurable function w, but B(E) is large enough to contain $T^*(w)$ in that case. Moreover, B(E) is small enough such that for any $w \in B(E)$: $T^*(w) \in B(E)$ [cf. Bersekas & Shreve (1978, §§ 7.7, 8.2)].

3.10 Lemma The operator $T^* : B(E) \mapsto B(E)$ is contracting w.r.t. the sup–norm $\|...\|$, in fact:

$$\|T^*w - T^*v\| \le C_\lambda (C_\lambda + \underline{C}_\beta)^{-1} \cdot \|w - v\| \ .$$

Proof. For any x, $\hat{u}$:

$$|Tw(x,\hat{u}) - Tv(x,\hat{u})| \le {}_0\!\int^\infty \exp\{-(\Lambda+B)(x,\hat{u},t)\} \cdot \lambda(\varphi(t|x),\hat{u}(t))\ \|w-v\|\ dt$$

$$= \|w-v\| \cdot {}_0\!\int^\infty \exp\{-(\Lambda+B)(x,\hat{u},t)\} \cdot (\lambda+\beta)(\varphi(t|x),\hat{u}(t)) \cdot$$

$$\lambda(\varphi(t|x),\hat{u}(t)) \cdot (\lambda+\beta)(\varphi(t|x),\hat{u}(t))^{-1}\ dt$$

$$\le C_\lambda (C_\lambda + \underline{C}_\beta)^{-1} \cdot \|w-v\|\ {}_0\!\int^\infty \exp\{-(\Lambda+B)(x,\hat{u},t)\} \cdot (\lambda+\beta)(\varphi(t|x),\hat{u}(t))\ dt$$

$$\le C_\lambda (C_\lambda + \underline{C}_\beta)^{-1} \cdot \|w-v\| \ , \qquad \text{since } \lambda/\beta \le C_\lambda/\underline{C}_\beta \quad \text{by 3.1d,f.} \ \square$$

(3.11) $\hat{C}_b(E)$ stands for the set of all bounded and u.s.c. functions on E,

$\hat{C}_o(E)$ is a non–empty subset of $\hat{C}_b(E)$ which is closed w.r.t. uniform convergence such that

$(x,a) \mapsto v(f(x,a,y))$ is u.s.c. $\forall\, y \in D$ for $v \in \hat{C}_o(E)$.

Obviously we can choose $\hat{C}_o(E) = \hat{C}_b(E)$ if the system function f is continuous.

3.12 Lemma. For $v \in \hat{C}_o(E)$ we have:

(a) $r(x,a) + \lambda(x,a) \cdot \int v(f(x,a,y))\ Q[dy]$ is u.s.c. in (x,a);

(b) $Tv(x,\hat{u})$ is a bounded and u.s.c. function in $(x,\hat{u})$ and $T^*v \in \hat{C}_b(E)$.

Proof. Set $\tilde{w}(x,a) := \int v(f(x,a,y))\ Q[dy]$.
From the assumption and Fatou's Lemma we obtain:

$\limsup_{n\to\infty} \int v(f(x_n,a_n,y))\, Q[dy] \le \int \limsup_{n\to\infty} v(f(x_n,a_n,y))\, Q[dy]$

$\le \int v(f(x_0,a_0,y))\, Q[dy] = \tilde{w}(x_0,a_0).$

Thus, $\tilde{w}(x,a)$ is u.s.c. and bounded. Therefore, we get that $w(x,a) := r(x,a) + \lambda(x,a)\cdot\tilde{w}(x,a)$ is u.s.c. with $|w| \le C_r\cdot\eta(x) + \|\tilde{w}\|\cdot C_\lambda\cdot\eta(x) = (C_r + \|\tilde{w}\|\cdot C_\lambda)\cdot\eta(x)$ and that

$$Tv(x,\hat{u}) = \int_0^\infty \exp\{-(\Lambda+B)(x,\hat{u},t)\}\cdot w(\varphi(t|x),\hat{u}(t)) =: W(x,\hat{u})$$

as in 3.8 from which we conclude that $Tv(x,\hat{u})$ is a bounded and u.s.c. function in $(x,\hat{u})$. Using the compactness of $\hat{U}$ we obtain from 3.9 our result. []

Now we obtain from 3.10 and 3.12:

3.13 Theorem. If the system function f is continuous, then T* is a contraction operator

$$T^* : \hat{C}_b(E) \mapsto \hat{C}_b(E) .$$

In view of the application to the transformed insurance model, we will consider a situation which is more general than that of 3.13.

3.14 Assumption. T* is an operator $T^* : \hat{C}_o(E) \mapsto \hat{C}_o(E)$.

By Banach's fixed point theorem we obtain from 3.10 and 3.14 :

3.15 Corollary There exists a unique $V^* \in \hat{C}_o(E)$ such that $T^*V^* = V^*$.

It is well–known that the value function of the equivalent controlled discrete–time model is also a fixed point of T* [cf. Bertsekas & Shreve (1978, 9.10)] and hence coincides with V*.

There is another argument that $V_n^* := T^{*n}0$ converges to the value function of the equivalent controlled discrete–time model which also works in the case of unbounded costs. From 3.14 and 3.12b we conclude that $V_n^* \in \hat{C}_o(E)$ and $TV_n^*(\cdot,\cdot)$ is bounded and u.s.c. . Thus, $TV_n^*(\cdot,\cdot)$ can be approximated from above by a decreasing sequence of bounded and continuous functions. Now a result by Schäl (1975 Theorem 13.1) applies.

We want to apply the results of Yushkevich (1987) who only considers non–relaxed controls. Therefore we will use the following device. We look on a relaxed control function $\hat{u}\in\hat{U}$ as a classical non–relaxed control function for the action space $\mathbb{P}(A)$ in place of A. Then V* is also the value function for the continuous–time model [cf. Yushkevich (1987): (3.15), Davis (1993): §44]. The boundedness assumptions of Yushkevich (1987, 2.18) are satisfied since V* is bounded. Now we will use the results of Yushkevich (1987) proved for the controlled continuous–time model. We need the **continuous–time reward operators**:

$Lv(x,a) := r(x,a) + \lambda(x,a)\ {}_D\int\ [v(f(x,a,y)) - v(x)]\ Q[dy] - \beta(x,a)\cdot v(x),$
$Lv(x,\mu) := r(x,\mu) + \lambda(x,\mu)\cdot {}_E\int\ [v(z) - v(x)]\ q(dz\,|\,x,\mu) - \beta(x,\mu)\cdot v(x)\ ,$
$L^*v(x) := \sup_{\mu\in\mathbb{P}(A)} Lv(x,\mu)\ .$

From $Lv(x,\mu) = {}_A\int Lv(x,a)\ \mu[da]$ one easily obtains the following important equation:

(3.16) $L^*v(x) = \sup_{a\in A} Lv(x,a).$

3.17 Bellman equation [Yushkevich (1987) 4.10, 4.15, 5.2.6]
$V^*(\varphi(t\,|\,x))$ is an absolutely continuous function in $t\geq 0$ and

$$-\frac{\partial}{\partial t} V^*(\varphi(t\,|\,x)) = L^*V^*(\varphi(t\,|\,x)) \quad \text{a.e. on } [0,\infty) \text{ for all } x\in E.$$

3.18 Lemma. There exists a measurable function $\delta : E \mapsto A$ such that

$$LV^*(x,\delta(x)) = L^*V^*(x),\ x\in E.$$

Proof. By assumption $-\lambda(x,a)\cdot\ v(x) - \beta(x,a)\cdot v(x)$ is continuous in a. Now from 3.15 and 3.12 we obtain that $LV^*(x,a)$ is u.s.c. in a for all $x \in E$. Now a well–known selection theorem applies [cp. Bersekas & Shreve (1978, 7.33)]. []

3.19 Corollary If the system f is continuous or if more generally Assumption 3.14 holds, then there exists an optimal stationary (non–relaxed) feedback policy defined through $a_t = \delta(X_{t-0})$, $t \geq 0$, where δ is chosen as in 3.18 and a_t is the action chosen at time t.

The proof follows as in 2.25 or from Yushkevich (1987, 4.10) or other verification theorems [cf. Yushkevich (1989), Davis (1993, (42.8))].

Appendix.

A.1 Chain Rule (Change–of–Variable Formula). Let V be an absolute continuous function on an interval I , i.e.

$$V(t) - V(t_o) = {}_{t_o}\int^t v(s)\ ds \quad \text{for some locally integrable function v on I.}$$

Let $G : [x_o,\infty) \mapsto I$ be an absolute continuous function, i.e.

$$G(x) - G(x_o) = {}_{x_o}\int^x g(y)\ dy \quad \text{for some locally integrable function g on } [x_o,\infty).$$

If g is strictly positive, then $V\circ G$ is absolute continuous and more exactly

$$V(G(x)) - V(G(x_o)) = {}_{x_o}\int^x v(G(y))\ g(y)\ dy\ \ ,\ x \geq x_o.$$

The present form of the chain rule is probably well known, but a proof for this special situation could not be found in the literature. Therefore we provide the proof.

Proof. The mapping G is continuous and strictly increasing. Let be $G(\infty) := \lim_{t\to\infty} G(t)$ $(\le \infty)$. Then $I' := [G(x_o),G(\infty)) \subset I$. There exists the inverse mapping $G^{-1} : I' \mapsto [x_o,\infty)$ and we have

$$V(G(x)) - V(G(x_o)) = \int_{G(x_o)}^{G(x)} v(u)\,du = \int \mathbf{1}_{(G(x_o),G(x)]}(u)\, v(G\circ G^{-1}(u))\,du$$
$$= \int_{I'} \mathbf{1}_{(x_o,x]}(G^{-1}(u))\, v(G\circ G^{-1}(u))\,du = \int \mathbf{1}_{(x_o,x]}(y)\, v(G(y))\,\mu[dy]$$

where μ is the image measure of the Lebesgue measure λ on $[x_o,\infty)$ under the mapping G^{-1}.
We have for $x > x_o$: $\mu[\,[x_o,x]\,] = \lambda[\{u;\, x_o \le G^{-1}(u) \le x\}] = \lambda[\{u;\, G(x_o) \le u \le G(x)\}]$
$= G(x) - G(x_o) = \int_{x_o}^{x} g(y)\,dy$, hence $\mu[dy] = g(y)\,dy$, and we finally obtain:

$$V(G(x)) - V(G(x_o)) = \int_{x_o}^{x} v(G(y))\, g(y)\,dy. \;[]$$

A2 Proof of Theorem 2.10.

Define $\psi : [x_o,\infty)\times[x_o,\infty) \mapsto [0,\infty)$ by $\psi(x|z) := \int_z^x \frac{1}{g(y)}\,dy$.
Then $\psi(z|z) = 0$, ψ is continuous and $\psi(\cdot\,|z)$ is stricty increasing to ∞ since

$$\lim_{x\to\infty} \psi(x|z) \ge \lim_{x\to\infty} \int_z^x \frac{1}{\alpha + \gamma\cdot y}\,dy = \infty \;.$$

Now let z be fixed and $\psi(\cdot\,|z) =: \psi$. Then, $\psi : [z,\infty) \mapsto [0,\infty)$ is bijective and the inverse $\psi^{-1}(\cdot) : [0,\infty) \mapsto [z,\infty)$ exists. Set

$$\varphi(t) = \varphi(t|z) := z + \int_0^t g\circ\psi^{-1}(s)\,ds,$$

where $0 < \int_0^t g\circ\psi^{-1}(s)\,ds \le \int_0^t [\alpha + \gamma\cdot \psi^{-1}(s)]\,ds < \infty$, $t > 0$, since ψ^{-1} is continuous. Then by the chain rule

$$\varphi(\psi(x)) - z = \varphi(\psi(x)) - \varphi(\psi(z)) = \int_z^x g\circ\psi^{-1}(\psi(y))\,\frac{1}{g(y)}\,dy = x - z\;,\; x\ge z.$$

Thus $\varphi = \psi^{-1}$ and we have $\varphi(t) := z + \int_0^t g\circ\varphi(s)\,ds$. Having proved the existence, we now turn to the uniqueness of the solution. Let $\tilde\varphi$ be any solution which is necessarily increasing, hence $\tilde\varphi(t) \in [z,\infty)$ and $\psi\circ\tilde\varphi$ is defined. From the chain rule we get

$$\psi(\tilde\varphi(t)) - \psi(z) = \psi(\tilde\varphi(t)) = \int_0^t g(\tilde\varphi(s))/g(\tilde\varphi(s))\,ds = t\;,$$

Thus $\tilde\varphi$ is the inverse ψ^{-1} of ψ, i.e. $\tilde\varphi = \varphi$. In order to prove continuity, suppose $t_n\to t_o$, $z_n\to z_o$ and set $x_n := \varphi(t_n|z_n)$, $\underline{x} := \liminf x_n$, $\bar x := \limsup x_n$. Then $\psi(x_n|z_n)\,) = t_n$.
Now choose $\zeta > \underline{x}$, then for infinitely many n: $\zeta > x_n$ and hence $\psi(\zeta|z_n) \ge \psi(x_n|z_n) = t_n$.
This implies $\psi(\zeta|z_o) \ge t_o$ and for $\zeta \downarrow \underline{x}$: $\psi(\underline{x}|z_o) \ge t_o$. By the same sort of argument, we obtain: $t_o \le \psi(\underline{x}|z_o) \le \psi(\bar x|z_o) \le t_o$. This proves $\underline{x} = \bar x = \varphi(t_o|z_o)$. []

A3 Proof of Lemma 3.8.

a) Let w be bounded and continuous. For $(x_n,\hat u_n) \to (x_0,\hat u_0)$ we write

$\Lambda_{mn}(t) + B_{mn}(t) := \Lambda(x_m,\hat u_n,t) + B(x_m,\hat u_n,t)$, $R_{mn}(t) := w(\varphi(t\,|\,x_m),\hat u_n(t))$,
$k(t) := \eta(\varphi(s\,|\,x_0))$, where k is positive and continuous.

Then we have $\quad |\Lambda_{nn}(t) - \Lambda_{00}(t) \le |\Lambda_{nn}(t) - \Lambda_{0n}(t)| + |\Lambda_{0n}(t) - \Lambda_{00}(t)|$.

W.o.l.g. we can assume that $n \ge n_o$ as in 3.1c(ii). By definition of the Young–topology we get:

$$\Lambda_{0n}(t) = {}_0\!\int^\infty {}_A\!\int \mathbf{1}_{(0,t]}(s)\cdot\lambda(\varphi(s\,|\,x_0),a)\,\hat u_n(s,da)\ \ ds \ \to\ \Lambda_{00}(t) ,$$

since $\vartheta(t,a) := \mathbf{1}_{(0,t]}(s)\cdot\lambda(\varphi(s\,|\,x_0),a) \le \mathbf{1}_{(0,t]}(s)\cdot C_\lambda\cdot k(s)$. Further

$$\begin{aligned}|\Lambda_{nn}(t) - \Lambda_{0n}(t)| &\le {}_0\!\int^t \sup_{\mu\in\mathbb{P}(A)} |\lambda(\varphi(s\,|\,x_n),\mu) - \lambda(\varphi(s\,|\,x_0),\mu)|\ ds\\ &\le {}_0\!\int^t \sup_{\mu\in\mathbb{P}(A)}\ {}_A\!\int |\lambda(\varphi(s\,|\,x_n),a) - \lambda(\varphi(s\,|\,x_0),a)|\ \mu[da]\ ds\\ &\le {}_0\!\int^t \sup_{a\in A} |\lambda(\varphi(s\,|\,x_n),a) - \lambda(\varphi(s\,|\,x_0),a)|\ ds .\end{aligned}$$

Now, by 3.9, the last integrand tends to zero and is dominated by $3\cdot C_\lambda\cdot k(s)$ in view of 3.1c,d; therefore the integral tends to zero. The same consideration applies for B. Thus we have by 3.1c,f

(*) $\qquad K_n(t) := \Lambda_{nn}(t) + B_{nn}(t) \to K_0(t)$ and $K_n(t) \ge \frac{1}{2}\cdot \underline{C}_\beta\cdot {}_0\!\int^t k(s)\,ds \quad \forall\ t\ge 0$.

Now we can consider :

$$\begin{aligned}|W(x_0,\hat u_0) - W(x_n,\hat u_n)| &= |{}_0\!\int^\infty \Big[\exp\{-K_0(t)\}\cdot R_{00}(t) - \exp\{-K_n(t)\}\cdot R_{nn}(t)\Big]\ dt|\\ \le {}_0\!\int^\infty &|\exp\{-K_0(t)\} - \exp\{-K_n(t)\}|\cdot|R_{nn}(t)|\ dt\\ &+ |\ {}_0\!\int^\infty \exp\{-K_0(t)\}\cdot[R_{00}(t) - R_{nn}(t)]\ dt\ |\\ \le 2\cdot C_w\cdot\ {}_0\!\int^\infty &|\exp\{-K_0(t)\} - \exp\{-K_n(t)\}|\cdot k(s)\ dt\\ &+ \ |\ {}_0\!\int^\infty \exp\{-K_0(t)\}\cdot R_{00}(t)\ dt\ - {}_0\!\int^\infty \exp\{-K_0(t)\}\cdot R_{0n}(t)\ dt\ |\\ &+ {}_0\!\int^\infty \exp\{-K_0(t)\}\cdot|R_{nn}(t) - R_{0n}(t)|\ dt .\end{aligned}$$

Because of (*) we have dominated convergence to zero of the first integrand and hence of the first integral. By the definition of the Young–topology, the second term converges to zero , since

$$\vartheta(t,a) := \exp\{-K_0(t)\}\cdot w(\varphi(t\,|\,x_o),a) \le C_w\cdot\exp\{-\tfrac{1}{2}\cdot\underline{C}_\beta\cdot{}_0\!\int^t k(s)\,ds\}\cdot k(t) .$$

Finally, the same arguments as for (*) show that the last integral tends to zero.

b) According to a theorem of Baire [cf. Bertsekas & Shreve (1978) 7.14], w is u.s.c. if and only if there exists a decreasing sequence of continuous functions w_n such that $w_n \downarrow w$. Now set $\tilde w_n(x,a)$ $:= (w_n(x,a) \wedge C_w\cdot\eta(x)) \vee (-C_w\cdot\eta(x))$, hence $|\tilde w_n(x,a)| \le C_w\cdot\eta(x)$.

Then $\tilde{w}_n$ is continuous and $\tilde{w}_n \downarrow (w \wedge C_w \cdot \eta) \vee (-C_w \cdot \eta) = w$. From part (a) we know that $W_n(x,\hat{u}) = {}_0\int^\infty \exp\{-\Lambda(x,\hat{u},t)-B(x,\hat{u},t)\} \cdot \tilde{w}_n(\varphi(t\,|\,x),\hat{u}(t))\, dt$ is continuous. Furthermore $\exp\{-\Lambda(x,\hat{u},t)-B(x,\hat{u},t)\} \cdot |\tilde{w}_n(\varphi(t\,|\,x),\hat{u}(t))| \le C_w \cdot \exp\{-\underline{C}_\beta \cdot {}_0\int^t \eta(\varphi(s\,|\,x))\, ds\} \cdot \eta(\varphi(t\,|\,x))$;
therefore we obtain $W_n(x,\hat{u}) \downarrow W(x,\hat{u})$ because of monotone and dominated convergence and $|W| \le C_w/\underline{C}_\beta$. Thus W is u.s.c. and bounded. □

A3 Proof of Theorem 2.24.

In the transformed insurance model, define $\hat{C}_o(E)$ as the set of all function v on E which are nonnegative, u.s.c., and bounded with $v(-1) = 0$. Then (3.11) is satisfied because of 2.34. By 3.12, we further know that T^*v is u.s.c. and bounded for $v \in \hat{C}_o(E)$. It is obvious that T^*v is nonnegative and that $T^*v(-1) = 0$. Hence $T^*v \in \hat{C}_o(E)$ for $v \in \hat{C}_o(E)$ and Assumption 3.14 is satisfied. We further noted in §2 that the Assumptions 3.1 are satisfied. Thus the Bellman equation 3.17 holds. Since $\tilde{\varphi}(t\,|\,x) = x + t$ for $x \ge 0$, absolute continuity of $V^*(\tilde{\varphi}(t\,|\,x)) = V^*(x+t)$ in t implies absolute continuity of V on $[0,\infty)$. Furthermore, $\frac{\partial}{\partial t} V^*(x+t) = (\frac{d}{dx}V^*)(x+t)$.
Now the Bellman equation reads:

$$
\begin{aligned}
&-(\tfrac{d}{dx})V^*(x+t)\\
&\quad= \sup_a \Big\{\tilde{r}(x+t,a) + \tilde{\lambda}(x+t,a) \cdot \int [V^*(f(x+t,a,y)) - V^*(x+t)]\; Q[dy] - \tilde{\beta}(x+t,a) \cdot V^*(x+t)\Big\} .
\end{aligned}
$$

Upon setting $x = 0$, we obtain (2.26); the case $x = -1$ is clear. □

Acknowledgements

The author is grateful to the referee for several useful remarks.

References

Asmussen, S., Taksar, M. (1997): Controlled diffusion models for optimal dividend pay–out. Insurance: Mathematics and Economics 523 (1997) 1 –15.

Bertsekas, D., Shreve, S.E. (1978): Stochastic Optimal Control: the Discrete–Time Case, Academic Press, New York.

Bremaud, P. (1981): Point Processes and Queues: Martingale Dynamics, Springer–Verlag, Berlin.

Cinlar, E., Jacod, J. (1981): Representation of semimartingale Markov processes in terms of Wiener processes and Poisson random measures. Seminar on Stochastic Processes, ed. E. Cinlar, K.L. Chung, R. K. Getoor, Birkhäuser, p. 159 – 242.

Dassios, A. and Embrechts, P. (1989) Martingales and insurance risk. Communications in Statistics – Stochastic Models **5** , 181–217.

Davis, M.H.A. (1984): Piecewise–deterministic Markov processes: a general class of non–diffusion stochastic models. J. Royal Statistical Society B, **46**, 353–388.

Davis, M.H.A. (1986): Control of piecewise–deterministic processes via discrete–time dynamic programming, in Stochastic Differential System (ed. M. Kohlmann). Lecture Notes in Control and Information Sciences **78** , Springer–Verlag, Berlin.

Davis, M.H.A. (1993): Markov Models and Optimization. Chapman & Hall, London.

Davis, M.H.A., Dempster, M.A.H., Sethi, S.P., Vermes, D. (1987): Optimal capacity expansion under uncertainty. Advances in Applied Probability, **19**, 156 – 76

Davis. M.H.A., Vellekoop, M.H. (1995): Permanent health insurance: A case study in piecewise deterministic Markov modelling. Mitteilungen der Schweiz. Vereinigung der Versicherungsmathematiker, Heft 2/1995, 177–211.

Dempster, M.A.H. (1991): Optimal control of piecewise deterministic Markov Processes, in: Applied Stochastic Analysis (ed. M.H.A. Davis and R.J.Elliot), Gordon and Breach, New York, pp. 303–25.

Dempster, M.A.H., Ye, J.J. (1990): A maximum principle for control of piecewise deterministic Markov processes, in: A.G. Law and C.L. Wang (eds.) Approximation, Optimization and Computing: Theory and Applications, North Holland, Amsterdam, pp. 235–40.

Dempster, M.A.H., Ye, J.J. (1992): Necessary and sufficient optimality conditions for control of piecewise deterministic Markov processes. Stochastics and Stochastics Reports, 40, 125–145.

Dempster, M.A.H., Ye, J.J. (1995): Impulse control of piecewise determistic Markov processes. Annals of Applied Probability **5**, 399 – 423.

Deshmukh, S.D., Pliska, S.R. (1980): Optimal consumption and exploration of nonrenewable resources under uncertainty. Econometrica **48**, 177–200.

Dickson, D.C.M., Waters, H.R. (1996): Reinsurance and ruin, Insurance: Mathematics and Economics **19**, 61–80.

Embrechts, P. (1984): Discussion to Davis (1984).

Embrechts, P. (1990): Martingales in Non–Life insurance. Vilnius Conference Proceedings VNU Press, **1**, 314–322.

Embrechts, P., Schmidli, H. (1994): Ruin estimation for a general insurance risk model. Advances in Applied Probability **26**, 404–422.

Fleming, W.H., Soner, H.M. (1993): Controlled Markov Processes and Viscosity Solutions. Springer Verlag, New York.

Hϕjgaard, B., Taksar, M. (1996): Optimal proportional reinsurance policies for diffusion models. Working paper.

Hϕjgaard, B., Taksar, M. (1997): Controlling risk exposure and dividends pay–out schemes: insurance company example. Working paper.

Last, G., Brandt, A. (1995): Marked point processes on the real line. Springer Verlag, New York.

Martin–Löf, Anders (1994): Lectures on the use of control theory in insurance. Scandinavian Actuarial Journal **1**: 1 – 25.

Pliska, S.R. (1978): On a functional differential equation that arises in a Markov control problem. Differential equations **28**, 390–405.

Rieder, U. (1975): Bayesian dynamic programming. Advances in Applied Probability 7 330–348.

Schäl, M. (1975): Conditions for optimality in dynamic programming and for the limit of n–stage optimal policies to be optimal. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **32**, 179–196.

Soner, H.M. (1985): Optimal control of a one–dimensional storage process. Applied Mathematics and Optimization **13**, 175–191.

Soner, H.M. (1986): Optimal control with state space constraint II, SIAM J. Control and Optimization **24**, 1110–1122.

Taksar, M., Zhou, X.Y. (1997): Optimal risk and dividend control for a company with a debt liability. Working paper.

Yushkevich, A.A. (1987): Bellman inequalities in Markov decision deterministic drift processes. Stochastics **23**, 25–77.

Yushkevich, A.A. (1989): Verification theorems for Markov decision processes with controlled deterministic drift and gradual and impulsive controls. Theory of Probability and Its Applications **34**, 474–496 .

Warga, J. (1972): Optimal Control of Differential and Functional Equations, Academic Press, New York.

Waters, H.R. (1983): Some mathematical aspects of reinsurance. Insurance: Mathematics and Economics **2**, 17 – 26.