OPTIMAL DYNAMIC CONTROL OF A USEFUL CLASS OF RANDOMLY JUMPING PROCESSES

Domokos Vermes

November 1980
PP-80-15

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

THE AUTHOR

DR. DOMOKOS VERMES is at the Bolyai Institute, University of Szeged and the Operations Research Department, Computer and Automation Institute, Hungarian Academy of Sciences.

ABSTRACT

The purpose of the paper is to present a complete theory of optimal control of piecewise linear and piecewise monotone processes. The theory consists of a description of the processes, necessary and sufficient optimality conditions and existence and uniqueness results, as well as extremal and regularity properties of the optimal strategy. Mathematical proofs are only outlined (they will appear elsewhere), but hints concerning efficient determination of the optimal strategy are included.

Piecewise linear (monotone) processes are discontinuous Markov processes whose state components stay constant or change linearly (monotonically) between two consecutive jumps. All processes of inventory, storage, queuing, reliability and risk theory belong to these classes. The processes will be controlled by feedback (Markov) strategies based on complete state observations. The expected value of a performance functional of integral type with additional terminal costs is to be minimized.

The semigroup theory of Markov processes will be used as the uniform mathematical tool for the whole theory, and the control problem will be reduced to the integration of a system of ordinary differential equations. Special emphasis will be given to the description of the processes by their infinitesimal characteristics which are available explicitly in applied models--no finite dimensional distributions are used.

CONTENTS

INTRODUCTION

The classical theory of risk, reliability, inventory, queueing and storage constitute an almost unified branch of applied (or rather, applicable) probability. They are characterized by a uniform methodology which makes use of limit theorems of probability theory in order to arrive at explicit formulae for one or other asymptotic characteristics of the processes. Using these explicit expressions one can make an appropriate choice of parameters appearing in the formulae to achieve a reasonable limit behaviour. The processes occurring in the above theory are very similar, they are continuous-time random processes with step or saw-tooth shaped trajectories, and in handling them there is a strong trend towards application of Markov process theory. Abstracting the common properties of these processes, Gnedenko and Kovalenko (1966) have defined the very useful class of "piecewise linear processes" which contain almost all processes of applied probability.

Under conditions of growing intensification of economic, technological, etc., competition, current interest is however declining in systems which offer only reasonable performance after the passage of the considerable time represented by asymptotic results. Instead, systems are needed which themselves perform "optimally" in any situation. Moreover, such processes should flexibly follow the changes of their economic, technological, social, etc., environment. In order to be able to design systems meeting these requirements one needs first of all a detailed quantitative description of performance criteria. Further, one needs the freedom to make decisions (to change the values of the free parameters) at any time at which the state of the system demands. One thus arrives at the *optimal dynamical control* problem: in view of the current state of the system determine the necessary decisions which result in optimal future performance. Of course, it is too much to require that the optimal state-decision correspondence be given an explicit analytical expression. But it is a cardinal requirement that any algorithm governing this correspondence can be efficiently processed by the present generation of computers.

In this paper we present a complete theory of optimal control of piecewise linear (and piecewise monotone) processes. *Piecewise linear processes (PLPs)* are discontinuous stochastic processes which have one class of state components changing linearly between consecutive random jumps, while the remaining state components change only by jumps--i.e. they remain constant between two neighbouring discontinuities. *Piecewise monotone processes (PMPs)* differ from PLPs only in the respect that a single state component can change in an arbitrary monotonic way between two consecutive jumps. For examples of PLPs we refer to Gnedenko and Kovalenko (1966) and Kalashnikov (1978). The performance of these processes will be characterized by a cost functional of integral form with an additional term for the termination costs. The aim is to find a strategy which ensures minimal expected costs from *every* initial state. Decisions will be made on the basis of *complete* observation of the current state of the process, in other words one uses *Markov (feedback) control strategies.* As PLPs and PMPs are Markov processes, strategies using information about the past evolution of the processes can be supposed not to give better results than Markov strategies (*cf.* Yushkevich, 1977).

The complete theory of optimal control includes necessary and sufficient optimality conditions (§4), existence of the optimal strategy and unique solvability of the differential equation system appearing in the optimality condition (§5), extremal properties of the optimal strategy and a regularity property of its possible jumps (§6). Special emphasis is given to the description of processes by their infinitesimal characteristics (§2) which are available explicitly in real-life problems, Finite-dimensional distributions--which are extremely hard to compute--are not needed.

The fundamental uniform tool of the whole theory is the linear semigroup theory of Markov processes. The optimality criteria will be given in terms of the infinitesimal operators of the processes as in the author's previous papers (1973, 1974, 1980). As the infinitesimal operators of PLPs and PMPs are first order ordinary differential operators, the determination of the optimal strategy is reduced to the solution of a system of nonlinear ordinary differential equations. The methods of the paper use repeatedly the characteristic property of continuous-time control that the value of the strategy can change at any moment. This freedom makes the world of continuous time processes much richer than that of their discrete-time analogs, so that one is in a position to prove results which have no discrete time counterparts (e.g. the whole of §6). The author is convinced that these are precisely the results which make the optimal strategy efficiently computable and easily realizable. (Contrast this with the poor computational performance of the backward algorithm of discrete-time dynamic programming.)

The paper is addressed to specialists wishing to solve concrete real-life problems using mathematical methods. Therefore it contains a precise, compact, self-contained description of the theoretical results with special emphasis on the determination of the optimal strategy. The results are stated in rigorous form, but mathematical proofs are omitted--they will appear elsewhere. Instead, explanations and applications-oriented hints are included. Potential applications of the theory presented are dynamical problems in the classical fields of risk, reliability, inventory, queuing and storage. But the author hopes that the theory will also give access to such vital modern fields as dynamic capacity expansion, control of natural resource reserves, investment and project phasing, temporal management of financial, manpower and natural resources, pest management and dynamic traffic control problems.

1. PIECEWISE LINEAR PROCESSES

1.1. *Definitions*

We start with a somewhat restricted definition of a *piecewise linear process (PLP)*. Later we shall show that this class of processes may be considered without loss of generality since every PLP can be reduced to this form by a simple state-space transformation.

As *time-axis* for our problems we take the non-negative real line $[0,\infty)$. The *state* x_t of our process at time $t \in [0,\infty)$ will be the two dimensional vector $x_t = (\nu_t, \zeta_t)$ where the *primary* (or *fundamental*) component ν_t takes its values in a finite or infinite set $\mathcal{N}$ of nonnegative integers, while ζ_t is a real variable. The range of the *secondary* component ζ_t depends on the actual value of the fundamental component ν_t. If $\nu_t = n$ then ζ_t lies in the internal $Z^n = [a^n, b^n]$. Consequently, the set $E := \{(n,z) : z \in [a^n, b^n],\ n \in \mathcal{N}\}$ will serve as state space for our processes. Subsequently, we shall denote by E^* the *right-side boundary* $E^* = \{(n, b_n) : n \in \mathcal{N},\ a_n \neq b_n\}$ of the state space, while $E^\circ := E \setminus E^*$. We assume that the state space contains a *terminal subset* Δ with the property that the process is killed if it reaches Δ.

The dynamics of the process can be described as follows. The primary state changes by jumps at random times. The secondary state changes by jumps at jump-moments of ν_t and it increases continuously with unit velocity between jumps. If $(\nu_t, \zeta_t) \notin E^*$, then the *jump intensity* of the fundamental component process is (ν_t, ζ_t), i.e. $P(\nu_{t+h} \neq \nu_t) = h \cdot \lambda(\nu_t, \zeta_t) + 0(h)$. This situation is called *continuous influence of choice*

(Gnedenko, 1966). If (ν_t,ζ_t) reaches the right-boundary E* then a jump in both components occurs necessarily (this situation is termed *discrete influence of choice*). If the process jumps from the state $x = (n,z^n) \in E$, then the probability that after the jump the evolution begins from a point of the set $A \subset E$ is given by the probability measure $\pi_x(A)$. We shall sometimes use the term *jump intensity measure* meaning $Q_x(A) = \lambda(x)\ \pi_x(A)$ for $x \in E^\circ$. In the sequel we shall always assume that our processes are *regular*, i.e. that there are only *finitely* many jumps in finite time intervals. Criteria for regularity can be found in Kalashnikov (1978).

Other authors (Gnedenko and Kovalenko (1966) and Kalashnikov (1978) define PLPs somewhat more generally. They do not explicitely assume that ν is one dimensional and that ζ_t increases with unit velocity. Instead they suppose that for constant ν the secondary component ζ_t moves in a d-dimensional Euclidean space Z^ν with constant velocity vector v^ν. Here Z^ν and v^ν can depend on ν. But it is easy to see that if we rotate Z^ν so that its first coordinate axis lies in the direction of v^ν and rescale this axis appropriately, then in the new coordinate system ζ_t will again move with unit velocity along the first coordinate axis and the remaining dim $Z^\nu-1$ coordinates may be ignored. Notice that if $v^\nu = 0$ (the case of a homogenous Markovian jump process) then Z^ν shrinks to a single point and (ν,Z^ν) does not belong to the boundary E*, i.e. no discrete influence of choice is possible. Thus we have shown that our simpler definition leads to no loss of generality.

By similar--but nonlinear--state space transformations, processes for which the secondary velocity depends not only ν but also on ζ can be reduced to our definition. Such processes occur, for example, in the theory of dams with controlled release rate (*cf.* de Marais (1976) and Pliska (1977)). A different approach to such problems will be given in §7.2.

1.2. *Examples*

We will give three very simple examples of PLPs, which will be used throughout the paper as illustrations of the concrete form of the general results. Our aim in this is not to demonstrate here the broad applicability of the PLP concept, but rather to present the most simple and transparent special cases.

1. *Non-homogenous Markovian jump process.* The process is characterized by the state and time dependent jump intensity $\lambda(n,t)$ and the jump measure $\pi_{n,t}$. If we include time as the secondary component in the state space the process will become a homogenous PLP. Its specific property is that the secondary component makes jumps of 0 saltus, i.e. it *never* jumps.

2. *Semi-Markov process.* It is well known that a homogenous pure jump process is *Markovian* if and only if the times between neighbouring jumps are exponentially distributed. If this is not the case, e.g. if the interjump times have distribution function $B_n(t)$ with density $b_n(t)$, then one can construct a Markovian equivalent of the original process as follows. Adjoin the time s already spent in the current state since the last jump as secondary state component. The resulting PLP with jump intensity $\lambda(n,s) = b_n(s)/[1-B_n(s)]$ will be Markovian, and its primary component coincides with the original process. Notice that since $1-B_n(s)$ can tend to zero faster than $b_n(s)$, it cannot be generally assumed that λ is bounded.

3. *Virtual waiting-time process.* We have a non-negative valued one-dimensional process which has positive jumps at random times. Between the jumps the process decreases with unit velocity if its value is not zero, and does not change until the next jump once level zero is reached. The times between the jumps are exponentially distributed with state dependent intensity $\lambda(x)$, and the independent magnitudes of the jumps have a common distribution function B_x. The process is a PLP with the special property that its state space is $E = \{(0,0)\} \cup \{1\} \times [0,\infty)$, and from state $(1,0)$ there is an immediate jump to $(0,0)$. Otherwise the jump intensity is $\lambda(x)$. Such a PLP describes the virtual waiting time of a customer arriving in an M/G/1 queue. Similar processes occur in the theory of water-reservoirs and in risk theory.

2. MARKOV PROCESSES AND INFINITESIMAL GENERATORS

2.1 *Definitions*

It is not difficult to show that any PLP $\{x_t\}$ is a *Markov* process (Dynkin, 1963; Gihman & Skorohod, 1973) and that if x_t is regular then the *transition function* $P(x,t,\Gamma)$ is uniquely determined by the intensities $\lambda(x)$ for $x \in E \setminus E^*$ and the jump measures π_x for $x \in E$. On the other hand, it is not easy to find an explicit expression for $P(x,t,\Gamma)$ in terms of λ and π (*cf.* Gihman & Skorohod, 1973). Therefore, we shall describe the transition behaviour of our processes by their infinitesimal operators, which are closely connected to the transition probabilities but can be expressed in terms of λ and π.

Denote by $\mathscr{B}(E)$ the space of bounded measureable real-valued functions on E. The norm in $\mathscr{B}$ is defined by $\|f\| = \sup_{x \in E} |f(x)|$. We say that the sequence f_k in $\mathscr{B}$ converges *strongly* to $f \in \mathscr{B}$ if $\|f_k - f\| \to 0$, while f_k converges *weakly* to $f \in \mathscr{B}$ if $f_k(x) \to f(x)$ for any $x \in E$ and the sequence is uniformly bounded, i.e. $\sup_n \|f_n\| \leq K$. We call two functions f and g *a.e. equal* (or *equivalent*) on E if for every n, $f(n,z) = g(n,z)$ for almost every $z \in z^n$ (with respect to Lebesque measure). To the Markov process x_t with the transition function $P(x,t,\Gamma)$ we adjoin the *semigroup* T_t of linear operators mapping $\mathscr{B}$ into itself, with

$$T_t f(x) := \int_E f(y) P(x,t,dy) \quad . \tag{2.1}$$

Clearly T_0 is the identity, $T_t T_s = T_{t+s}$ and $\|T_t f\| \leq \|f\|$.

Corresponding to each of the different types of convergence in $\mathscr{B}$ we can define different *"infinitesimal operators"* of the semigroup T_t by the formula

$$Af := \lim_{t \downarrow 0} \frac{1}{t}(T_t f - f) \quad . \tag{2.2}$$

In this paper we shall say that a function $f \in \mathscr{B}$ belongs to the *domain* $\mathscr{D}(A)$ of the *infinitesimal operator* (or *generator*) A if the limit in (2.2) exists in the sense of weak convergence in $\mathcal{B}$. This notion of generator is somewhat weaker than the weak infinitesimal operators (Dynkin,1963) but it is the appropriate definition for our purposes. For the later development it will be of fundamental importance that Dynkin's formula holds true in the following form.

Let τ be a Markov time and $f \in \mathcal{D}(A)$ such that $E_x \int_0^\tau Af(x_t)dt$ is bounded. (Here E_x denotes the expectation with respect to the measure P_x.) Then

$$E_x f(x_\tau) - f(x) = E_x \int_0^\tau Af(x_t)\,dt \quad . \tag{2.3}$$

Notice that the infinitesimal generator is an unbounded linear operator, it is determined by formula (2.2) *and* its domain. It can be seen from subsequent considerations, that besides the actual expression of Af the relation $f \in \mathcal{D}(A)$ also contains essential information about the functions f and Af.

2.2. *Generators of PLPs*

The fundamental method for the determination of the infinitesimal operator of a PLP was developed by Vermes (1974). Although only PLPs which arose from semi-Markovian processes were considered, the method of elaboration of the generator can be applied without essential change to general PLPs. Taking into account that--contrary to the situation for the usual definition of the weak infinitesimal operator--for our generators no weak continuity is needed, we have the following.

THEOREM 1. *The domain of the infinitesimal operator of the PLP defined in §1 consists of those functions {f} which are uniformly Lipschitzian and right differentiable with respect to the secondary variable and for which, for small enough t and some constant K,*

$$\sup_{\nu,\eta} |T_t f(\nu,\zeta) - f(\nu,\zeta)| \leq Kt \quad . \tag{2.4}$$

For these functions the generator is given by the formula

$$Af(\nu,\zeta) = \frac{d^+}{d\zeta} f(\nu,\zeta) + \lambda(\nu,\zeta) \int_E [f(n,z) - f(\nu,\zeta)]\pi_{\nu,\zeta}(dn,dz) \quad (2.5)$$

$$\text{if} \qquad (\nu,\zeta) \in E\setminus E^* .$$

From (2.4) it follows that for $(\nu,\zeta) \in E^*$, the "boundary condition"

$$f(\nu,\zeta) = \int_E f(n,z)\ \pi_{\nu,\zeta}(dn,dz) \qquad (2.6)$$

holds true for any $f \in \mathscr{D}(A)$.

If λ is bounded on all of E, then $\mathscr{D}(A)$ consists of all uniformly Lipschitzian right differentiable functions which satisfy the boundary condition (2.6), and Af is given by (2.5).

If $\lambda(\nu,\zeta) \to \infty$ as $\zeta \to b^\nu$ then (for regular processes) (2.4) is satisfied for those functions for which the second term on the right hand side of (2.5) remains bounded. In this case, in addition to the boundary condition (2.6), the asymptotic equality

$$|f(\nu,\zeta) - f(\nu,b^\nu)| \sim 1/\lambda(\nu,\zeta) \qquad \text{as } \zeta \to b^\nu \qquad (2.7)$$

also characterizes the relation $f \in \mathscr{D}(A)$.

We would like to call the attention to a property of the generators of PLPs. Namely that $g = Af$ always has the regularity property

$$g(n,z) = \lim_{h \downarrow 0} h^{-1} \int_0^h g(n,z+t)dt \qquad (2.8)$$

if $(n,z) \in E\setminus E^*$. Sometimes it will be convenient to make use of formulae by which Af is determined only up to equivalence. In such cases equality and inequality relations will be understood as relations holding in every point provided that every equivalence class of functions is represented by its regular element, for which (2.8) holds. A typical example is formula (2.10).

2.3. *Examples*

It is easy to see that if $\lambda(n,t)$ is bounded, then for the *first* example of §1.2 the generator is defined for any function having bounded right derivatives and AF is given by

$$Af(n,t) = f^{+}(n,t) + \lambda(n,t) \sum_k [f(k,t) - f(n,t)]\, \pi_{n,t}(k) . \qquad (2.9)$$

For the *second* example $\lambda(n,s)$ remains bounded if, for large s, $b_n(s) \geq e^{-ks}$, for some $k > 0$. In this case the generator is defined for every function having bounded right-derivative and is given by the formula

$$Af(n,s) = f^{+}(n,s) + (b_n(s)/[1-B_n(s)]) \sum_k [f(k,0)-f(n,s)]\, \pi_{n,s}(k) . \qquad (2.10)$$

If the duration time distribution is supported by an infinite interval, i.e. if $b^n = +\infty$ for any n, but $b_n(s)$ is not minorized by an exponential function, then only functions satisfying condition (2.7) are in the domain of the generator, but as no discrete influence of choice takes place, the boundary condition of type (2.6) is not necessary. However, if the duration time distribution is supported by the finite interval $[a^n, b^n]$ with $b^n < \infty$, then all functions in the domain of the generator must satisfy conditions (2.6) and (2.7) at b^n.

For the *third* example, if λ is bounded, Af is given by

$$Af(x) = f^{+}(x) + \lambda(x) \int_0^{\infty} [f(y)-f(x)]\, B_x(dy) \qquad (2.11)$$

for any bounded right-differentiable function $f \in \mathscr{B}$.

3. OPTIMAL CONTROL OF PLPs

3.1 *Controllable processes*

Up to now we have been dealing with single processes only; now we shall consider families of processes which contain a free parameter in their characteristics. By appropriate choice of this free parameter we can single out that process from the family which has good dynamics from some point of view. If the choice of this parameter depends on the evolution of the process then we term the family of PLPs a *controlled process*. Since PLPs are uniquely determined by intensities λ and jump measures π we can control such a process by these two characteristics.

Problems in which control acts only on the jump measures at the boundary point $x \in E^*$ are essentially equivalent to the control problem for discrete time Markov chains. The fundamental tool for solving such problems is dynamic programming. Even if there are some serious computational difficulties connected with the application of dynamic programming procedures, the theory can be regarded as closed. The interested reader should consult the extensive literature (Dynkin, 1963 and Howard, 1964) and references contained therein.

In the present paper we shall treat the polar case, namely when the terms connected with the continuous influence of the choice can be controlled. In other words, we shall deal with processes for which the intensity measure Ω depends not only on x but also on a free parameter y. (Recall that Ω_x is defined only for $x \in E^0$ by $\Omega_x = \lambda(x)\pi_x$.) More precisely, let Y be a closed bounded subset of R^n, the so-called *action space*, and let x_x^y be a continuous measure-valued function on $E^0 x Y$. Here the space $\mathcal{M}$ of finite measures is endowed with its usual weak (more precisely w*) topology. In order to set off the characteristic features of this "continuously acting" control, we do *not* allow the discrete jump measures π_x $(x \in E^*)$ to depend on the control parameter y.

Since a regular PLP is uniquely defined by its intensity measures and discrete jump measures, for any fixed $y \in Y$ we get a unique PLP on the same state space E. But we shall in fact study a much broader class of processes, where the value of the control parameter y can be chosen to depend on the actual state of the process. We denote by U the set of all measurable mappings u of E into Y, and call it the set of all *feedback* (or *Markov*) *strategies*. To every $u \in U$ there belongs an intensity measure function $Q_x^{u(x)}$, which together with the discrete jump measures determine a new PLP. We say this process is *governed* by the strategy u. The transition functions, probability measures, expectations and generators belonging to the Markov process governed by u will be denoted by $P^u(x,t,\Gamma)$, P_x^u, E_x^u, A^u respectively. For random variables, as x_y^u, τ^u, the upper index u will be used only if confusion would otherwise arise.

In many problems it is not reasonable to consider *all* measurable strategies, since a general measurable feedback law can be very difficult to realize. Sometimes we wish to consider only piecewise constant or piecewise continuous strategies, or feedback laws taking values only in a subset of Y (*cf.§6).* In order to treat this situation generally, we define a subset $U \subset U_0$ to be the set of all *admissable* strategies if it satisfies the following conditions:-

a) All constant strategies are in U.

b) If u and v are in U, ν is an arbitrary primary state and $\mathcal{I}$ an interval from Z^ν, then the strategy defined by

$$w(n,z) := \begin{cases} u(n,z) & \text{if } n = \nu \text{ and } z \in \mathcal{I} \\ v(n,z) & \text{otherwise} \end{cases}$$

is also admissable, i.e. $w \in U$.

3.2. *Cost functionals and optimality*

In order to estimate the effectiveness of different control strategies we have to specify the costs which arise in the course of the evolution and termination of the process. We assume that the whole cost is the sum of a *terminal cost* component p(x), $x \in \Delta$, which depends on the state where the process is killed, and an *evolutionary* component-- the *expense rate*--which increases at a state and control dependent rate q(x,y) along the trajectories of the process. Our aim is thus to minimize the expected cost

$$J_x(u) = E_x^u\{p(x_\tau) + \int_0^\tau q(x_t, u(x_t))dt\} . \tag{3.1}$$

Here p and q are assumed to be bounded continuous functions of their arguments. (For problems with unbounded q see §7.1.)

Clearly $J_x(u)$ depends also on the *initial* state x. We shall call a strategy u* *optimal*, if for every initial state $x \in E$ it minimizes the expected cost, i.e. if $J_x(u^*) = \inf J_x(u)$ for every $x \in E$. Of course the existence of a sole universal strategy which minimizes J_x for all initial states is by no means a trivial matter. But in Section 5 we shall see that in fact there exists such a uniformly optimal strategy.

In many practical problems--especially in those arising in economics-it is adequate to consider a cost structure with a discounting factor. The *discount rate* $\alpha(x)$ can even depend on the states x visited by the process, i.e. we have a more general cost functional

$$\tilde{J}_x(u) = E_x^u \; p(x_\tau)e^{-\int_0^\tau \alpha(x_s)ds} + \int_0^\tau q(x_t, u(x_t))e^{-\int_0^t \alpha(x_s)ds} dt . \tag{3.2}$$

If x is a constant, we have the usual *constant* rate discounting.

Observe that such a discounted problem can be reduced to the original undiscounted problem, we have only to kill our process with the (variable) rate $\alpha(x)$. In other words the new terminal time will be $\tau := \min(\tilde{\tau}, \sigma)$, where σ and x_t have the common distribution

$$P_x^u(\{\sigma > t\} \cap \{x_t \in \Gamma\}) = E_x^u\{I_\Gamma(x_t) \cdot e^{-\int_0^t \alpha(x_s)ds}\} . \tag{3.3}$$

It is easy to see that by killing in this manner from a process with infinitesimal generator A we obtain a process with generator $A - \alpha$, and the domains of both operators coincide.

Consequently the following two problems are equivalent:

a) To solve the discounted problem for processes with generators A^u.

b) To solve the undiscounted problem for processes with generators $A^u-\alpha$.

For simplicity, we shall assume in the sequel that the cost functional is uniformly bounded, i.e. $\sup_{x,u} J_x^u \tau < \infty$. This is always the case if p and q are bounded and the expected lifetimes of the processes are bounded, i.e. $\sup_{x,u} E_x^u \tau < \infty$. This is al-discount rate α is strictly positive. The assumption is not seriously restricting since if there exists one strategy with bounded cost, then clearly the optimal cost must remain below this bound. Those strategies which lead to costs exceeding the bound cannot be candidates for optimality, i.e. they can be deleted from the set of admissible strategies. Problems with unbounded cost rate and consequently with unbounded cost functional, will be investigated in §7.1.

4. OPTIMALITY CRITERIA

In this section we state the fundamental result of the paper, a necessary and sufficient optimality condition. In order to include some explanations and refinements we treat sufficiency and necessity separately.

4.1. *Sufficient optimality condition*

THEOREM 2. *Suppose we have a function Ψ belonging to the intersection of the domains of the infinitesimal generators A^u for all $u \in U$ and a strategy $u^* \in U$ which satisfy the equation*

$$A^{u^*}\Psi(x) + q(x,u^*(x)) = \min_{y \in Y} A^y\Psi(x) + q(x,y) = 0 \qquad \text{if } x \in E^0 \tag{4.1}$$

with boundary condition

$$\Psi(x) = p(x) \qquad \text{if} \quad x \in \Delta \quad . \tag{4.2}$$

Then u^* is optimal in U and Ψ is the corresponding optimal cost function $\Psi(x) = J_x(u^*) = \min_u J_x(u)$.

Notice that (4.1) can be rewritten in the form

$$A^{u^*}\Psi(x) + q(x,u^*(x)) = 0 \qquad (4.3)$$

and $$A^y\Psi(x) + q(x,y) \geq 0 \qquad \text{for any } y \in Y. \qquad (4.4)$$

Theorem 2 concerns the analog of the *Hamilton-Jacobi* sufficient condition of the calculus of variations for the case of stochastic PLPs. Being a sufficient condition it is adequate to decide whether a strategy (found by some other method) is optimal. Taking into account the actual form of the generator for PLPs, we have only to check whether u^* and Ψ satisfy the differential equation (4.3), inequality (4.4) and boundary condition (4.2).

The proof of Theorem 2 is contained in Vermes (1973) and is based on the application of Dynkin's formula. We remark that the proof does not use the property (a) and (b) of the set of admissable strategies (see §3.1), hence the theorem is valid for any class of strategies.

4.2. *Necessary optimality condition*

THEOREM 3. *Suppose that the domains of the infinitesimal generators corresponding to all admissible strategies $u \in U$ coincide. If u^* is an optimal strategy and $\Psi(x) = J_x(u^*)$ is the corresponding optimal cost function, then Ψ is in the common domain of the generators and u^* and Ψ together satisfy (4.1)-(4.2).*

Moreover Ψ is continuously differentiable with respect to the secondary variable and $\frac{d}{d\eta}\Psi(\nu,\eta)$ is uniformly Lipschitzian in η on E.

The proof of Theorem 3 is accomplished along the following lines. First we prove that $\Psi \in \mathscr{D}(A^{u^*})$ and that it satisfies (4.2), (4.3). From $\Psi \in \mathscr{D}(A^{u^*})$ it follows immediately that $\Psi \in \cap \mathscr{D}(A^{u^*})$ and that Ψ is uniformly Lipschitzian and right-differentiable. Next we prove the validity of (4.4); this critical step is based on the following lemma.

Lemma. *Let $G \subset E$ denote on arbitrary open set and let $\sigma = \sigma_G$ be the first exit time from G. If $u(x) \equiv v(x)$ for $X \in G$, and for any $x \in E$*

$$E_x^u [\int_0^\sigma q^u(x_t)dt + J_{x\hat{\sigma}}(v)] \leq E_x^v \int_0^\sigma q^v dt + J_{x\hat{\sigma}}(v) \qquad (4.5)$$

holds, then $J_x(u) \leq J_x(v)$ for all $x \in E$. If strict inequality holds in (4.5) for some x_0, then also $J_{x0}(v)$.

In the third and final step of the proof of Theorem 3 we use (4.1) to prove the additional regularity property of Ψ.

Theorem 3 uses only the condition that the domains of the generators coincide. It follows from Theorem 1 that this condition is automatically satisfied if λ is bounded. Notice that the discrete jump measures η_x describing the discrete influence of choice ($x \in E^*$) do not by assumption depend on the control; hence the *same* boundary condition(2.6) holds for any strategy u.

If λ is *singular*, i.e. if $\lambda(\nu,\zeta,u) \to \infty$ as $\zeta \to b^\nu$, then the validity of the asymptotic equality (2.7) is also necessary for $f \in \mathscr{D}(A^u)$. Consequently, in order to have $\mathscr{D}(A^u) = \bigcap_{u \in U} \mathscr{D}(A^v)$, we must additionally demand that for every $y \in Y$ the intensities $\lambda(\nu,\zeta,y)$ tend to the infinity with the same order, i.e. for any pairs y_1, $y_2 \in Y$ and for any ν

$$\lambda(v,\zeta,y_1) \sim \lambda(\nu,\zeta,y_2) \qquad \text{as} \qquad \zeta \to b^\nu \quad . \qquad (4.6)$$

From the theoretical point of view, Theorems 2 and 3 leave something to be desired. Namely, since in Theorem 3 the existence of the optimal strategy u* is explicitly assumed, the theorems cannot be used to prove existence of the optimal strategy. At the expense of a laborous proof, this defect can be remedied as the following theorem shows.

THEOREM 4. *Suppose the domains of the infinitesimal generators* A^u *coincide for any* $u \in U$. *Denote by* $\Psi(x) = \inf_{u \in U} J_x(u)$, *then*

$$\inf_{y \in Y} [A^y \Psi(x) + q^y(x)] = 0 \qquad (4.7)$$

is valid for any $x \in E^0$. *Moreover* Ψ *is continuously differentiable in* η *with uniformly Lipschitzian derivative, and* Ψ *satisfies (4.2) and (4.4).*

The first two steps of the proof of Theorem 3 *cannot* be repeated for the present context since the fundamental relation $\Psi \in \mathcal{D}(A^{u*})$ is now meaningless. The bulk of the work needed to prove Theorem 4 is included in the proof of the rigorous version of the *invariant imbedding theorem* for our problem. As it is interesting in itself, we formulate this result as an independent theorem.

THEOREM 5. (Invariant imbedding). *Let* $G \subset E$ *be an arbitary open set and let* $\sigma = \sigma_G$ *be the first exit time from* G. *Then we have for every* $x \in E$

$$\Psi(x) = \inf_{u \in U} E_x^u \{\int_0^\sigma q^u(x_t)dt + \Psi(x_\sigma) \quad . \qquad (4.8)$$

4.3. *Computation of the optimal strategy.*

Besides their theoretical significance (see also §§5 and 6) the practical value of Theorems 2 and 3 is that they give a constructive method for the determination of the optimal strategy Piecing together the results of Section 2.2, 4.1, and 4.2, we see that the optional strategy u* and the optimal cost Ψ together satisfy the difference-differential equation system

$$\frac{d}{dz}\Psi(n,z) = \lambda(n,z,u^*(n,z)) \int [\Psi(n,z) - \Psi(\nu,\zeta)] \pi_{n,z}^{u^*(n,z)}(dn,dz) + q(n,z,u^*(n,z))$$

$$= \inf_{y \in Y} \{\lambda(n,z,y) \int [\Psi(n,z) - \Psi(\nu,\zeta)] \pi_{n,z}^{Y}(dn,dz) + q(u,z,y)\} \qquad (4.9)$$

for $z \in (a^n, b^n)$ and each $n \in \mathcal{N}$, with boundary conditions.

$$\Psi(n,b^n) = \int \Psi(\nu,z)\pi_{n,b^n}(d\nu,dz) \tag{4.10}$$

and $$\Psi(n,z) = p(n,z) \qquad \text{if } (n,z) \in \Delta \quad . \tag{4.11}$$

Equation 4.9 is nothing but an ordinary (nonlinear) differential equation. Notice that though the optimal strategy u* can have jumps, the right-hand side of (4.9) remains not only continuous but even Lipschitzian in z for u* and fixed n ; consequently it can be solved by any method of discretization and convergence is ensured.

The value u*(x) of the optimal strategy is that control y* which minimizes the right-hand side of (4.9) for the actual state (n,z). Consequently, simultaneously with the solution of the differential equation we have to carry out a minimization in y at every step.

In many important cases, especially if the intensity measures $Q_x^y = \lambda(x,y)\pi_x^y$ depend linearly (or affine linearly) on y, the minimizing control can be expressed explicitly in terms of Ψ. In such cases, we have only to determine the solution ψ of the ordinary nonlinear differential equation, which does not contain any step by step minimization, and after that we may elaborate u* from ψ.

If no explicit formula is available for u* in terms of Ψ then the minimization at every step over the whole action space can considerably increase the computational effort. There arises the danger that plagues discrete time dynamic programming; namely, that we have a general method--theoretically applicable for all complicated processes arising in applications--but the necessary computational effort increases so rapidly with the complexity of the system that it restricts the applicability of the method to the most simple problems only. Further theoretical work is necessary to show that this is not the case. In Section 6 we shall show that one need not carry out the minimization at every step over the whole state space. We shall see that only a very few distinguished actions determine the optimal strategy, and even switching between these distinguished actions cannot be arbitrary.

5. EXISTENCE AND UNIQUENESS

Recall that according to §3.2 we call a strategy u^* optimal if it minimizes the expected cost $J_x(u)$ for any initital state $x \in E$. It could occur that for different starting points x different strategies are optimal. This would mean that our declared task, to find one universal strategy which is optimal for any initial state, is impossible. In this case all results of Section 4 would remain formally correct, but they would be practically useless(i.e. involve empty conditions). Even if there were a universal optimal strategy, it is not sure that it could be chosen from the relatively simple class U_0 of Markovian feedback strategies. It is possible that it might also depend on the past of the process and not only on the current state. The following theorem shows that the problem formulated in Section 3 is solvable, and the class U_0 of Markovian feedback strategies is broad enough to contain an optimum.

THEOREM 6. *There exists an optimal strategy in the class* U_0.

This result can be deduced from Theorem 4 by the measurable choice theorem.

In the last section we have seen that an optimal cost function Ψ satisfies the differential system (4.9) with boundary conditions (4.10)-(4.11). But there arises the question of whether this is the only solution. Is it possible that we might find another solution of (4.9)-(4.11) which differs from J_x ? This case would be dangerous--even from the point of view of numerical methods, since different approximating sequences might converge to different solutions depending on the choice of discretizing points. Notice that because of the nonstandard boundary conditions and the possible discontinuities of the right hand side, the classical uniqueness theorems from the theory of differential equations cannot be used. However, from Theorem 3 we can deduce the following.

THEOREM 7. *The differential equation system (4.9) has only one bounded solution with boundary conditions (4.10), (4.11).*

Notice that Theorem 7 states only the uniqueness of the optimal cost function, but does not exclude the possibility that several *alternative strategies* might result in the same cost.

Recall that according to §3 we seek the optimum in some class U of admissible strategies satisfying condition (a) and (b). One might expect that in different classes of admissible strategies there are different optima, especially since a strategy which is optimal in a small strategy class is not necessarily the best which can be found in a broad class. But this is not true. Our uniqueness result shows that if we find a strategy optimal in some class satisfying (a) and (b), then the same strategy is optimal even in the broadest class U_0.

6. EXTREMAL PROPERTIES OF THE OPTIMAL STRATEGY

6.1. *Bang-bang principle.*

As we saw in §4.3, in order to determine the optimal strategy we must solve a differential equation system and simultaneously in any state carry out a minimization over the whole space. The integration of a differential equation is a standard numerical procedure, which even for large systems can be accomplished in reasonable time. But the necessary minimizations are extremely time-consuming, since generally the action space consists of infinitely many points.

One of the fundamental results of the control theory of linear deterministic systems is the bang-bang principle. It states that any admissible strategy can be substituted by another which takes values only from the extremal points of the action space and which is equivalent to the original strategy from the point of view of time optimality. This means, seeking for the time-optimal strategy one needs to take into account only the extremal points of the action space. For example, if Y is the unit square, then one has to minimize only amongst its four vertices instead of the infinitely many points of the square. The result is an essential saving of computational effort. Unfortunately, the bang-bang principle is not valid for nonlinear systems or for discrete time deterministic processes in general.

In the present section we shall see that the analog of the bang-bang principle holds true for PLPs. First we state the result for PLPs depending in some sense linearly on the control and for a time optimality criterion, then we state the best result for general PLPs and performance criteria.

Suppose that the intensities are of the following form: $\lambda(x,y) = \lambda_0(x) + y\,\lambda_1(x)$ and that the jump measures π_x do not depend on y. Further let the criterion functions be $q \equiv 1$ and $p \equiv 0$, i.e. we look for the strategy which controls the process to reach the target set Δ in the shortest possible time. Then we have the following result.

THEOREM 8. (Linear bang-bang principle). *The values of the optimal strategy can always be chosen from the extremal points of the action set.*

The benefits arising from Theorem 8 are obvious. In equation (4.9) we can write $\inf_{y \in \text{ex}\, Y}$ instead of $\inf_{y \in Y}$, and since generally there are only a few extremal points, we can save a large amount of computational effort. Also the realization of the control strategy will be much simpler. We need not memorize a complicated function, only the points where we have to switch over from one extremal point to another.

Contrary to the deterministic case, this result can be generalized to nonlinear systems and to general performance functionals as well. For this we have to introduce an auxiliary notion. The set of all possible pairs of intensity measures and cost rates $\mathscr{I}(x) := \{(Q_x^y, q(x,y)) : y \in Y\}$ is called the *indicatrix* of the problem at the point $x \in E$. $\mathscr{I}(x)$ is a compact subset of the cartesian product $\mathscr{M} \times R^1$, where $\mathscr{M}$ denotes the set of all bounded measures with the weak* topology induced from C(E). It is easy to see that in the linear case, the indicatrix is isomorphic with the action space (since $q \equiv 1$)--that is why in that case this extra notion is superfluous. We can formulate the general bang-bang principle for controlled PLPs in terms of the indicatrix.

THEOREM 9. (Nonlinear bang-bang principle). *The value $u^*(x)$ of the optimal strategy at the state x can always be chosen from the extremal points of the indicatrix $\mathcal{I}(x)$ at x.*

Thus a strategy cannot be optimal if its values lie in the interior of the convex hull of the indicatrix on a subset $E_1 \subset E$ of nonzero measure.

The proof of Theorem 8 and 9 use a balayage technique combined with a sharpened version of the measurable choice theorem (see Vermes 1980).

Our results show that in Markovian continuous time stochastic control problems the *optimal* strategy is much *simpler* than a general nonoptimal strategy. Therefore, we cannot share the views of those authors who suggest finding a nearly optimal solution in *lieu* of the true optimum in order to simplify algorithms for construction and realization of the strategy. In our view, an optimal bang-bang strategy is much simpler than a nearly optimal continuous one--even if the continuity assumption is comfortable for the theory.

We would like to emphasize that the bang-bang principle is an essentially *continuous-time* result. It is closely connected with the notion of intensity and intensity measures which cannot even be defined in discrete time. For the validity of the bang-bang principle it is essential that the controller is in the position to switch at the correct instant (*cf.* Property (b) of 3.1). If this freedom is restricted--e.g. by requiring that switchings are allowed only at some fixed moments (e.g. points of discrete time scale)--then the optimal strategy looses its bang-bang property.

Translating this property to the language of the realization of the strategy, we could say that a continuous time process model is justified if the technological or organizational structure of the process enables its controller to intervene operationally in its evolution at any moment when the necessity arises. For example, if a manager can change the production structure only at the beginning of a year, then the continuous time process model is inadequate even if production itself runs continuously. According to the bang-bang principle, the optimal process runs in some sense with *extremal velocity*; consequently relatively small errors in the choice of the points when we change its direction can have disasterous consequences.

6.2. *Randomized Strategies.*

Up to now we have treated only pure feedback strategies. We controlled the process on the basis of its (current) states, and our aim was to find the best such strategy. One might expect that in a larger strategy class one can find a better optimum.

A pure feedback strategy means that if we observe that our system is in state x then we necessarily apply control u(x). One could control in such a way that if we are in state x then with some probability we apply action y_1, with some other probability y_2, etc. In other words, to every state x there corresponds a probability measure μ_x on the action space and we control the process by an action chosen randomly according to the measure μ_x.

These form the class of so called *randomized strategies*.

Of course, this class is much larger than the class of pure strategies and the question arises whether randomized strategies are more effective. The answer is negative and follows from the bang-bang principle.

THEOREM 10. *To every randomized strategy corresponds a pure one which yields a criterion value not worse than the original one.*

This means that randomized strategies have no advantage over pure feedback strategies; hence it is enough to deal with the latter simpler class.

6.3. *Jump conditions*

The optimal strategy and optimal cost function together satisfy the fundamental differential equation (4.9). We already know that it is enough to take into consideration only the extremal points of the indicatrix. But there arises the question whether or not the optimal strategy can jump arbitrarily between the extremal points or if there is some further regularity in its evolution.

Suppose first that the intensity measures depend linearly on the control and that the action space is an n-dimensional cube. We call two *vertices* (extreme points) *neighbouring* if they lie on a common edge. Then we have the following result.

> THEOREM 11. *The optimal strategy can always be chosen so that depending on the secondary variable, with the primary variable held fixed, its value jumps between neighbouring extremal points of the action space.*

This result further simplifies the solution algorithm of (4.9)-(4.11). Equation (4.9) is a differential equation in the secondary variable with the primary variable held fixed. Numerically, we must solve it forward or backward along the z axis. Theorem 11 says that in order to determine the value of the optimal strategy $u^*(n,z_0+\Delta z)$ we need not minimize over all extremal points of Y, but only over those which are neighbouring to $u^*(n,z_0)$. Since the discretizing points $a^n+k\Delta z$ are much more dense along the z axis than the jumps of u^*, in fact no stepwise minimization is necessary. We solve the differential equation (4.9) without "inf", with the value of u^* from the last step. Simultaneously we also compute the value of the right-hand side for the neighbouring extremal y values and check whether they are still larger than the right-hand side for $u^*(n,z_0+\Delta z)$. We have to carry out the minimization only in the case when the latter condition is not satisfied. Otherwise we go on with the integration without any change.

Theorem 11 can be generalized to nonlinear problems as well, we have only to define what we mean by neighbouring points. If the indicatrix is finite dimensional, then two extremal points are called *neighbouring* if they have a *common supporting* *hyperplane* . That means there exists a hyperplane such that the whole indicatrix lies on one side of it and both extremal points are on this plane. If the indicatrix does not vary with z, more precisely if all $\mathscr{J}(n,z)$ are isomorphic for $z \in (a^n, b^n)$, then Theorem 11 remains valid; one has only to write neighbouring points of the indicatrix instead those of the action space.

In the general case the indicatrix is not finite dimensional, consequently we have to use linear functionals instead of hyperplanes. Moreover, if the indicatrix varies with z, then it can occur that points which are neighbouring for one z are not neighbours for another z. To be precise, we can say in this case that in any state (n,z_0) there is a hyperplane in $\mathscr{M} \times R^1$ such that all limit points of the sequence

$$(Q_{n,z}^{u^*(n,z)}, q(n,z,u^*(n,z)))$$

lie on it as $z \to z_0$, and in a neighbourhood of z_0 all indicatrices lie on one side of the hyperplane. In other words we have the following result.

THEOREM 12. *For any pair (n,z_0) there exists a continuous linear functional on $\mathscr{M}$ and a constant c such that for some $\varepsilon > 0$*

$$\ell(Q_{n,z}^{y}) + q(n,z,y) > c$$

for any $y \in Y$ and $|z-z_0| < \varepsilon$. Moreover if (z_k) is an arbitrary sequence tending to z_0, then

$$\ell(Q^{u^*(n,z_k)}) + q(n,z_k,u^*(n,z_k)) \to c \quad .$$

Theorems 11 and 12 have important implications regarding the continuity properties of the optimal strategy. If the optimal strategy is unique, then it necessarily has the bang-bang property and satisfies the jump condition. Suppose that we have a linear system, then if the action space is a cube or a polyhedron, the optimal strategy is a pure jump function. If on the other hand the action space has *no* neighbouring extremal points, then the optimal strategy *cannot* have jumps and must be continuous. This, for example, is the case if the indicatrix is a disc. Then all points of the boundary circle are extremal, but there are no two points with common tangent. Consequently the value of the optimal strategy can only vary continuously along the circle. Analogous results can be formulated for nonlinear systems in terms of the indicatrix.

The information that the optimal strategy is continuous also simplifies the computations, since in this case it is enough to seek for the minimum in a small neighbourhood.

7. EXTENSIONS

7.1. *Unbounded expense rates*

In §3.2 the expense rate q and the terminal cost p were assumed to be bounded. There are several real-life applications in which these functions are finite but not bounded. In this subsection we investigate the changes in the theory which are necessary if we want to include discounted problems with finite but unbounded expense rate.

For simplicity, we assume a constant discount rate and zero terminal costs. In this case our cost function will be of the form

$$J_x(u) = E_x \int_0^\infty e^{-\alpha t} q(x_t)dt \tag{7.1}$$

The inclusion of an arbitrary stopping time τ and a bounded terminal cost component would make no difference.

The fundamental difficulty is cased by the fact that for unbounded q the equation (*cf.* §4.3)

$$Af - \alpha f + q = 0 \tag{7.2}$$

has in general no bounded solution although it has several different solutions in the class of finite but not necessarily bounded functions. Even for bounded q such extra finite but unbounded solutions exist, but J_x is the unique bounded solution of (7.2). Consequently for unbounded q one must determine a class of unbounded functions in which the cost J_x is the only solution of (7.2).

Let A_α and A_0 denote the generators of the discounted (killed) and of the undiscounted (permanent) processes respectively. They are defined for all finite functions such that the respective limits $\frac{1}{t}(T_t^\alpha f-f)$ exist as $t\downarrow 0$. Then we have the obvious relation $A_\alpha f = A_0 f-\alpha f$. We are ready to state the fundamental result of this subsection.

THEOREM 13.

a) *Equation* $A_\alpha f + q = 0$ *has at most one solution* f *such that* $A_0 f$ *remains bounded.*

b) *If* $\|T_t q-q\|$ *remains bounded on some finite interval* $t\in[0,t_0]$, *then there exists a function* f *such that* $A_0 f$ *is bounded and* $A_\alpha f + q = 0$.

The condition in part (b) means that $\sup_x|E_x q(x_t)-q(x)|$ may not be infinite for arbitrary small t. It is interesting to note that this condition is equivalent to the seemingly more stringent assumption

$$\|T_t q-q\|\leq c_0 + c_1 t \qquad (7.3)$$

holding for all t where c_0,c_1 are constants.

By Theorem 13 we can extend the whole control theory of PLPs developed in §§3-6 to unbounded expense rates q with the property (7.3), provided that the discount rate is bounded away from zero. The only necessary changes are that in the conditions of Theorem 2-12 we must use the domains of operators A_0^u while in Equation (4.1) one has to consider A_α^u and A_α^y instead of A^u and A^y respectively. Here we give only the reformulation of Theorem 7 in which the largest number of changes are necessary.

THEOREM 7'. *Suppose the domains of the operators* A_0^u *coincide and that*

$$\sup_x E_x^u|q(x_t,u(x_t))| - q(x,u(x))| \leq c_0^u + c_1^u t$$

for some constants c_0^u, c_1^u, *and every* t. *Then the equation*

$$\inf_{y \in Y} [A_0^y \Psi(x) - \alpha\Psi(x) + q(x,y)] = 0 \quad x \in E \setminus E^* \qquad (7.4)$$

has exactly one solution $\Psi \in \bigcap_u \mathcal{D}(A_0^u)$. *This solution is*

$$\Psi(x) = \inf_{u \in U_0} E_x^u \int_0^\infty e^{-\alpha t} q(x_t, u(x_t))dt \quad .$$

7.2. *Piecewise monotone processes*

As we saw in §1.1 the assumption that the secondary component increases with unit velocity results in no loss of generality compared with the original definitions of Gnedenko and Kovalenko (1966) and Kalashnikov (1978), where the velocity may depend on the primary component.

But there are other problems where the velocity of the secondary component depends not only on the primary component but on the secondary component as well. A typical example is the control of a storage system with content dependent release rate, where the secondary component is the dam-content itself (*cf.* de Marais,1976). If the secondary velocity is bounded and has constant sign, then we arrive at *piecewise monotone processes (PMPs)* to which our theory has a straightforward extension.

Suppose we are given the control problem defined in §§1-3 with only the change that the secondary component increases with velocity $v(n,z,y) \geq 0$ depending continuously on both state components as well as on the control variable y.

If v is not zero then Theorem 1 remains valid with the only change that in the expression of the generator (2.5) the first term is $v(n,z)df^+(n,z)/dz$ instead of $df^+(n,z)/dz$. All further expressions can be divided by $v \neq 0$ and the entire theory remains valid. In other words the control problem for a PMP with velocity $v > 0$ is equivalent to the control of a PLP with expense rate $q(x,u(x)/v(x,u(x))$.

If $v(x,y) = 0$ for some $x \in E$, $y \in Y$, then the situation is complicated by the following two facts:-

1. The domain of the generators A^u do not coincide for all $u \in U_0$, as at points where $x(x,u(x)) = 0$ the functions $f \in \mathcal{D}(A^u)$ need not be differentiable.

2. The velocity function $v(x,u(x))$ does not uniquely determine the path of the secondary component between two jumps, as $v^u(x) = v(x,u(x))$ is in general not then Lipschitzian.

Observe that the coincidence of the domains was needed only to ensure that $\Psi = \inf_u J(u)$ as in $\mathcal{D}(A^u)$ for all u, and that Ψ is absolutely continuous. Thus the first difficulty can be resolved by substituting for $d^+\Psi/dz$ a regularized version of the Radon-Nikodym derivative Ψ'_{RN}.

The second difficulty is caused by the fact that while PLPs can be uniquely defined by infinitesmal characteristics, when zero velocities are allowed, the infinitesmal generator does not determine a unique PMP. We have to specify which points (n,z) with $v^u(n,z) = v(n,z,u(n,z))$ are *instantaneous*, i.e. from which the process X^u exits continuously, and which ones are *stable* in the sense that the process exits from them only by jumps. We define

$$V_S(u) := \{(n,z) \in E \setminus E^* : \lim_{h \downarrow 0} h^{-1} \int_0^h v^u(n,z+t)\,dt = 0\}$$

and $V_I(u) := E \setminus V_S(u)$, and supplement the definition of the process X^u by the requirement that the points $(n,z) \in V_I(u)$ are exactly those where the state of the process remains unchanged until the next jump.

In order to promote an easier formulation of the conditions and results in the sequel we assume that there exists a $y_0 \in Y$ such that $v(n,z,y_0) \equiv 0$ and for all other, i.e. $y \in Y \setminus \{y_0\}$, we have that $v(n,z,y) > 0$ for all $(n,z) \in E$.

THEOREM 14. *The statements of Theorems 2-13 remain valid with the following changes.*

1. *We require only the coincidence of the domains* $\mathcal{D}(A^y)$ *for* $y \neq y_0$.

2. Ψ *does not necessarily belong to the common domain of the generators, it is generally not continuously right differentiable.*

3. *In each formula* $d^+\Psi/dz$ *is to be substituted by*

$$\widetilde{\frac{d^+\Psi}{dz}}\,(n,z) = \lim_{h \downarrow 0} h^{-1} \int_0^h \Psi'_{RN}(n,z+t)\,dt \quad . \qquad (7.5)$$

   (At those points where $d^+\Psi/dz$ *exists, in coincides with* $\widetilde{d^+\Psi}/dz$.*)*

4. *In Theorem 2 the condition that* Ψ *belongs to the joint domain is meaningless. Instead we have to require that* Ψ *is uniformly Lipschitzian with respect to the secondary variable and that it satisfies such boundary conditions (2.6)-(2.7) as arise from the domains* $\mathcal{D}(A^y)$, $y \neq y_0$.

5. *The statements of Theorems 11-12 do not hold in so far as jumps to or from* y_0 *are concerned.*

8. ACKNOWLEDGEMENT

I wish to thank M.A.H. Dempster for a careful reading of an earlier draft of this paper.

REFERENCES

Dynkin, E.B. (1963). *Markov Processes.* Moscow, 1963 (in Russian) Berlin-Heidelberg-New York, 1965 (in English).

Dynkin, E.B. and A.A. Yushkevitch (1975). *Controlled Markov Processes.* Moscow, 1975 (in Russian); Berlin-Heidelberg-New York, 1979 (in English).

Gihman, I.I. and A.V. Skorohod (1973). *Theory of Stochastic Processes.* Vol.II. Moscow, 1973 (in Russian); Berlin-Heidelberg-New York, 1977 (in English).

Gnedneko, B.V. and I.I. Kovalenko (1966). *Introduction to the Theory of Mass Service.* Moscow (in Russian).

Howard, R.A. (1960). *Dynamic Programming and Markov Processes.* New York, 1960 (in English; Moscow, 1964 (in Russian).

Kalashnikov, V.V. (1978). *Qualitative Analysis of the Behavior of Complicated Systems using Liapunov Functions.* Moscow (in Russian).

de Marais, P.R. (1976). Optimal Control of a Storage System. (Ph.D. dissertation). Evanston.

Pliska, S.R. (1977). On a functional differential equation that arises in a Markov control problem. *J. Differential Equations.*

Richel, R.W. (1975). A minimum principle for controlled jump processes. In: Lecture Notes in Economics and Mathematical Systems, Vol.107. *Control Theory, Numerical Methods and Computer Systems Modelling.* Berlin-Heidelberg-New York.

Snyder, D.L. (1975). *Random Point Processes.* New York.

Vermes, D. (1973). A necessary and sufficient condition of optimality for Markovian control problems. *Acta.Sci.Math.* (Szeged) 34:401-413.

Vermes, D. (1974). On optimal control of semi-Markov processes *Acta.Sci.Math.* (Szeged) 36:345;356.

Vermes, D. (1980). Extremality properties of the optimal strategy in Markovian control problems. In: *Analysis and Optimization of Stochastic Systems.* London.

Yushkevich, A.A. Markov controlled models with countable state space and continuous time. *Teor. Veroiatnosteĭ. Primen* XXII:222-241 (in Russian).