This is to certify that the
dissertation entitled

OPTIMAL CONTROL OF DYNAMICAL SYSTEMS
WITH JUMP MARKOV PERTURBATIONS

presented by

Alexey G Stepanov

has been accepted towards fulfillment
of the requirements for the

Ph. D. degree in Statistics

Professor Anatoli V. Skorokhod (Major Professor)

Professor Habib Salehi (Major Professor)

August 22, 2003
Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

OPTIMAL CONTROL OF DYNAMICAL SYSTEMS WITH JUMP MARKOV PERTURBATIONS

By

Alexey G Stepanov

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHYLOSOPHY

Department of Statistics & Probability

2003

ABSTRACT

OPTIMAL CONTROL OF DYNAMICAL SYSTEMS
WITH JUMP MARKOV PERTURBATIONS

By

Alexey G Stepanov

The control problem of a dynamical system with jump Markov perturbations is considered. The partial differential equation of dynamic programming for the value function $V(t, x, y)$ is derived. Also an integral equation for the value function $V(t, x, y)$ is obtained and is used to construct a sequence of functions that converge uniformly to the value function $V(t, x, y)$ from above.

TABLE OF CONTENTS

INTRODUCTION

Optimal control theory has seen tremendous growth over the past four decades, with important applications in finance, networks, manufacturing, medicine, operations research, and other areas of science and engineering. Optimal control theory is crucial to the design and operation of complicated modern systems since it ensures that vital variables are kept in check, regardless of the disturbance the system undergoes. In a variety of naturally occurring problems, we wish to control the system governed by a set of differential equations in order to minimize (or maximize) a given performance criterion.

In this dissertation, we discuss an optimal control problem of a system with a jump Markov disturbance. The evolution of the system is described by a system of differential equations, and there is a running (instantaneous) cost and a terminal cost associated with the process. A control is chosen in order to minimize the total cost of the process. However, due to a random jump Markov disturbance in the system, the position of the system at any given time is also random, and so is the total cost associated with the process. Thus, the control is chosen in order to minimize the expected value of the total cost of the process, and is also random, depending only on the present state of the process, and maybe also its past states.

The two main approaches used in control problems are the Dynamic Programming approach and Pontryagin's maximum principle.

In Dynamic Programming approach, the minimum (infimum) value of the performance criterion is considered as a function of the initial (starting) point. This function is called the value function, and in many ways it holds the key to solving the optimal control problem. R. Bellman [1] applied dynamic programming to the optimal control of discrete-time systems, demonstrating that the natural direction for solving optimal control problems is backwards in time. That is, we start solving an optimal control problem by first finding the optimal control policy on the very last step. Then, armed with that information, we search for the optimal control on the second to last step and so on, each time using the fact that we already have the optimal control policy for the steps that follow the one that is being currently considered. The dynamic programming approach decomposes the optimal control problem into a sequence of minimization problems that are carried out over the space of controls, and are far simpler than the original problem.

To illustrate the idea of dynamic programming, suppose the state of a system at time k is described by a process $x_k \in R^n$, satisfying

$$x_{k+1} = f(x_k, u_k), \qquad 0 \le k \le N-1,$$

where u_k is the control at time k whose value is chosen from the set of admissible controls $U_k \subset R^m$. Suppose that the performance criterion is given by $J_0(x, u_\bullet)$, where

$$J_i(x, u_\bullet) = \sum_{k=i}^{N} g_k(x_k, u_k), \qquad 0 \le i \le N, x \in R^n.$$

is the performance of the process between time i and time N initiating from $x = x_i$ at time i, when the control policy $u_\bullet = (u_0, u_1, \dots, u_N)$ is used. Let

$$V_k(x) = \inf_{u_\bullet}\{J_k(x,u_\bullet)\}, \qquad 0 \le k \le N, x \in R^n.$$

Then

$$V_N(x) = \inf_{u_N \in U_N}\{g_N(x_N,u_N)\}, \qquad x \in R^n,$$

and

$$V_k(x) = \inf_{u_k \in U_k}\{g_k(x_k,u_k) + V_{k+1}(f(x_k,u_k))\}, \qquad 0 \le k \le N-1, x \in R^n.$$

The problem of minimizing the performance criterion $J_0(x,u_\bullet)$ over the choice of the entire control policy $u_\bullet = (u_0, u_1, \dots, u_N)$ splits up into a sequence of smaller and simpler problems of finding the optimal choice of each u_k separately, working backwards in time. Furthermore, the values of u_k where the infimum is achieved on each step, u_k^*, $0 \le k \le N$, form the optimal (overall) control policy $u_\bullet^* = (u_0^*, u_1^*, \dots, u_N^*)$.

For continuous-time systems, the dynamic programming approach uses the same idea of working backwards in time, and uses the value function as a tool in the analysis of the optimal control problem. At time s, we choose the control $u(s)$ for our process $x(s)$, based on the assumption that the optimal control would be used after time s. Here, whenever the value function is differentiable, it satisfies a first order partial differential equation called the partial differential equation of dynamic programming. Suppose that the state of a system at time s is described by a process $x(s) \in R^n$, satisfying the system of differential equations

$$\frac{dx(s)}{ds} = a(s,x(s),u(s)), \qquad 0 \le t \le s \le T,$$

with initial condition

$$x(t) = x \in R^n,$$

where $u(s)$ is the control process whose value at time s can be chosen from the set of admissible controls $U(s, x) \subset R^m$. Suppose that the performance criterion is given by

$$J(t,x,u(\cdot)) = \int_t^T \Phi(s,x(s),u(s))ds + \Psi(x(T)).$$

Then if the value function $V(t,x) = \inf_{u(\cdot)} \{J(t,x,u(\cdot))\}$ is differentiable, it satisfies the (Hamilton-Jacobi-)Bellman equation (or dynamic programming equation)

$$V_t(t,x) + \inf_{v \in U(t,x)} \{V_x(t,x) \cdot a(t,x,v) + \Phi(t,x,v)\} = 0, \qquad 0 \le t < T, x \in R^n.$$

with initial condition

$$V(T,x) = \Psi(x), \qquad x \in R^n.$$

The dynamic programming equation is often rewritten as

$$-V_t(t,x) + \sup_{v \in U(t,x)} H(t,x,V_x(t,x),v) = 0, \qquad 0 \le t < T, x \in R^n,$$

where

$$H(t,x,p,v) = -p \cdot a(t,x,v) - \Phi(t,x,v).$$

$H(t,x,p,v)$ is generally called the Hamiltonian in analogy with a corresponding quantity occurring in classical mechanics.

Solving the (Hamilton-Jacobi-)Bellman equation gives us the value function. Moreover, the dynamic programming equation can be used to find the optimal control policy. Similarly to the discrete-time case, the values of $v \in U(t, x)$ where the infimum is achieved are the values of the optimal control policy $u^*(t)$.

If the value function fails to be differentiable at some points (t, x), it does not satisfy the dynamic programming equation everywhere. Thus, we need to consider a generalized solution to the dynamic programming equation if we wish to use the dynamic programming approach. However, we may encounter a serious lack of uniqueness when dealing with generalized solution. Crandall and Lions [2] introduced the concept of the viscosity solution. For a large class of optimal control problems, the value function is the unique viscosity solution of the related dynamic programming equation in the case when the value function is not smooth enough to be a classical solution. For more on viscosity solutions see Fleming, Soner [8].

The optimal control policy may not exist in some optimal control problems. Then for some points (t, x), the imfimum will not be achieved in the dynamic programming equation. In this case, the dynamic programming approach can be used to obtain an ε-optimal (almost optimal) control policy.

In general, the dynamic programming equation is hard to solve. In some cases one has to resort to numerical solution of the dynamic programming equations. Typically, the state space and the control space are discretized, and the minimization is carried out for the final number of states. However, the computational difficulties may be too restrictive for complex multidimensional problems.

Pontryagin's maximum principle developed by L.S. Pontryagin [15] gives a necessary condition that must hold on an optimal trajectory. Simply stated, if $u^*(s)$ and

$x^*(s)$ represent the optimal control and the state trajectory, then there exists an R^n-valued function $P(s)$ called an adjoint variable, such that together $u^*(s)$, $x^*(s)$ and $P(s)$ satisfy

$$\frac{dx^*(s)}{ds} = -H_p(s, x^*(s), P(s), u^*(s))$$

$$\frac{dP(s)}{ds} = H_x(s, x^*(s), P(s), u^*(s))$$

and for all s, $0 \le s \le T$, the optimal control $u^*(s)$ is the value of v maximizing $H(s, x^*(s), P(s), v)$, i.e., for all $v \in U(s, x)$,

$$H(s, x^*(s), P(s), v) \le H(s, x^*(s), P(s), u^*(s)).$$

If V is differentiable at each point $(s, x^*(s))$ of the optimal trajectory, then a candidate for an adjoint variable is $P(s) = V_x(s, x^*(s))$. Under certain conditions, Pontryagin's maximum principle can also be a sufficient condition for optimality.

Let $y(s)$ be a jump Markov process, and suppose that the state of a system at time s is described by a stochastic process $x(s) \in R^n$, satisfying the system of differential equations

(1) $$\frac{dx(s)}{ds} = a(s, x(s), y(s), u(s)), \qquad 0 \le t \le s \le T,$$

with initial condition

(2) $$x(t) = x \in R^n.$$

In (1), $u(s)$ is a parameter whose value we can choose at any instant in a set $U \subset R^m$ in order to control the process $x(s)$. Thus the control $u(s) = u(s, \omega)$ is a stochastic process. Since our decision at time s must be based upon what happened up to time s, the function $\omega \to u(s, \omega)$ must be measurable with respect to

$$\Im_s = \sigma\{(x(r), y(r)),\ r \le s\}.$$

The Markovian nature of the problem suggests that it might suffice to consider control processes of the form

$$u(s) = \underline{u}(s, x(s), y(s)).$$

Suppose that the performance criterion is given by

$$J(t,x,y,u(\cdot)) = E_{t,x,y}\left\{\int_t^T \Phi(s,x(s),y(s),u(s))ds + \Psi(x(T),y(T))\right\}$$

(3)

$$= E\left\{\int_t^T \Phi(s,x(s),y(s),u(s))ds + \Psi(x(T),y(T)) \,\middle|\, x(t)=x, y(t)=y\right\}$$

The problem is - for each $(t,x,y)\in[0,T]\times R^n\times Y$ - to find the number $V(t, x, y)$ and the pair $(x^*(s), u^*(s))$, $t \le s \le T$, satisfying (1) and (2), such that

(4) $$V(t,x,y) = \inf_{u(\cdot)}\{J(t,x,y,u(\cdot))\} = J(t,x,y,u^*(\cdot)),$$

where the infimum is taken over a given suitable class $\mathcal{U}$ of controls. $V(t, x, y)$ is called the value function.

In similar deterministic optimization problems, examples show that optimal controls may be discontinuous and that the partial differential equation of Dynamic Programming (Hamilton-Jacobi-Bellman equation) may not have a smooth solution. A similar situation is to be expected here, so class $\mathcal{U}$ of controls should include discontinuous controls.

Problems similar to the one described above with the disturbance being a finite state continuous time Markov chain and terminal cost were investigated by Rishel [16].

For earlier work see Krassovskii and Lidskii [11], [14], where the control problem is on an infinite interval and terminal conditions are not imposed. Florentin [9] discussed the optimal control of systems with disturbance inputs being generalized Poisson processes. Wonham [20] considered the special case of the linear regulator version of this problem. He explicitly computes the optimal control law for this case. In a slightly different setting Kushner [13] defined a stochastic maximal principle. Sworder [18] gave a maximal principle for fixed terminal time linear problems without terminal conditions. Rishel [16] discussed the relationship between dynamic programming optimality conditions and stochastic minimum principle optimality conditions different from those in [13] and [18]. Goor [10] proved the existence of the optimal control in some special cases.

The systems with finite state jump Markov disturbances are the most general special class of piecewise deterministic processes (PDPs), first introduced explicitly by Davis [3]. Roughly speaking, such processes are continuous time Markov processes consisting of a mixture of deterministic motion and random jumps. PDPs, with stochastic pure jump processes and deterministic dynamical systems as special cases, include nearly all non-diffusion continuous time processes in applied probability. The optimal control theory of PDPs was developed by Vermes [19], Davis [4], Soner [17] and Dempster and Ye [5], [6]. PDPs, with stochastic pure jump processes and deterministic dynamical systems as special cases, include virtually all of the stochastic models of applied probability except those involving diffusions. For the optimal control problems involving diffusions see Krylov [12].

In this paper the partial differential equation of dynamic programming for the value function $V(t, x, y)$ is derived. Also an integral equation for the value function $V(t, x, y)$ is obtained and is used to construct a sequence of functions that converge uniformly to the value function $V(t, x, y)$ from above. This sequence can be used to obtain the value function without solving the partial differential equation of dynamic programming, and the value function, in turn, can be used to obtain the optimal control, if it exists. Moreover, the convergence to the value function $V(t, x, y)$ is uniform in $(t,x,y) \in [0,T] \times R^n \times Y$, this sequence of functions can be used to obtain ε- optimal control functions, (that is, controls $u_\varepsilon(\cdot)$ for which $J(t,x,y,u_\varepsilon(\cdot)) \le V(t,x,y) + \varepsilon$ for all $(t,x,y) \in [0,T] \times R^n \times Y$ for some $\varepsilon > 0$). In fact, each function in the sequence is a performance function of a control that is selected from a class of controls $\mathcal{M}$ that is more general than $\mathcal{U}$. We will also show that if there is an optimal control in the class $\mathcal{U}$, then that control is also optimal in class $\mathcal{M}$. Therefore, while constructing the sequence of functions that converge uniformly to the value function $V(t, x, y)$ from above, we would also construct a sequence of controls that are ε- optimal in class $\mathcal{M}$.

CHAPTER 1

PRELIMINARIES

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Let Y be a complete separable metric space, and let $\mathcal{B}_Y$ denote the Borel σ-algebra on Y.

Let $y(s)$ be a jump Markov process with values in Y defined on $[0, T]$ with transition probabilities $P(t, y, s, A)$, $0 \le t \le s$, $y \in Y$, $A \in \mathcal{B}_Y$.

Suppose that for all $t \in [0, \mathrm{T})$, $y \in Y$, $A \in \mathcal{B}_Y$.

$$\frac{1}{s-t}\left[P(t,y,s,A) - 1_A(y)\right] \to \Pi(t,y,A)$$

uniformly in (t, y, A) as $s \to t$, $s > t$.

Suppose that for fixed (y, A) $(y \in Y,\ A \in \mathcal{B}_Y)$ $\Pi(t, y, A)$ is continuous in t uniformly in (y, A).

Let

$$\lambda(t,y) = \Pi(t,y,Y \setminus \{y\}) = -\Pi(t,y,\{y\}),$$

and suppose that for some K

$$\lambda(t,y) \le K \qquad \text{for all } t \in [0, \mathrm{T}),\ y \in Y.$$

Suppose that the values of the controls are restricted to a closed subset U of R^m.

Let Q denote $[0,T) \times R^n \times Y$ and $\overline{Q}$ denote $[0,T] \times R^n \times Y$.

Let the vector function $a(t,x,y,u): \overline{Q} \times U \to R^n$ and the running (instantaneous) cost $\Phi(t,x,y,u): \overline{Q} \times U \to R$ be continuous in (t, x, u) uniformly in y and differentiable in (t, x, u) for each $y \in Y$. Suppose also that $a_t(t,x,y,u)$, $a_x(t,x,y,u)$, $a_u(t,x,y,u)$, $\Phi_t(t,x,y,u)$, $\Phi_x(t,x,y,u)$, $\Phi_u(t,x,y,u)$ are all continuous in (t, x, u) uniformly in y, and $a_u(t,x,y,u)$ and $\Phi_u(t,x,y,u)$ are continuously differentiable in (t, x, u) for each $y \in Y$.

In addition, suppose that for all $(s,x,y,u) \in \overline{Q} \times U$

a) for suitable B_1

$$|a(s,x,y,u)| \le B_1(1+|x|),$$

b) for suitable $B_2(R)$

$$\|a_x(s,x,y,u)\| \le B_2(R),$$

whenever $|x| \le R$,

c) there exist $u_0 \in U$ such that for suitable C_1

$$\Phi(s,x,y,u) \ge -C_1 \text{ and } \Phi(s,x,y,u_0) \le C_1,$$

d) for suitable $C_2(R)$

$$|\Phi_x(s,x,y,u)| \le C_2(R),$$

whenever $|x| \le R$.

Let the terminal cost $\Psi(x,y): R^n \times Y \to R$ be continuous in x uniformly in y and differentiable in x for each $y \in Y$, with $\Psi_x(x,y)$ continuous in x uniformly in y.

Suppose also that for all $(x,y) \in R^n \times Y$

e) for suitable D_1

$$|\Psi(x,y)| \le D_1,$$

f) for suitable $D_2(R)$

$$|\Psi_x(x,y)| \le D_2(R),$$

whenever $|x| \le R$.

Let $\tau_t^0(\omega) = t$, and for $k = 1, 2, 3, \ldots$, let

$$\tau_t^k(\omega) = \inf\left\{s > \tau_t^{k-1}(\omega) : y(s,\omega) \neq y\left(\tau_t^k(\omega),\omega\right)\right\} \wedge T,$$

i.e. $\tau_t^k(\omega)$ is the time of k^{th} jump of the process $y(s,\omega)$ after time t.

Let $y_k(\omega) = y\left(\tau_t^k(\omega),\omega\right)$, $k = 0, 1, 2, \ldots$.

With these notation, (1) becomes

$$\frac{dx(s)}{ds} = a\big(s, x(s), y_k, u(s)\big),$$

$$\tau_t^k \le s < \tau_t^{k+1}, \qquad k = 0, 1, 2, \ldots.$$

Since $y(s)$ is independent of $x(s)$ and $u(s)$, the randomness of the system is due to the "outside" disturbance $y(s)$, neither the control $u(s)$ nor the trajectory $x(s)$ of the system have any impact on it.

Therefore, instead of considering controls of the form $u(s) = \underline{u}\big(s, x(s), y(s)\big)$, let us consider deterministic (non-random) intrajump control functions of the open loop nature $u(\cdot) = u(\cdot; t, x, y)$ which need to be specified from t to the terminal time T in case no jump of $y(s)$ occurs before T. If a jump occurs at time τ before T to $y' \in Y$ say, the control function $u\left(\cdot;\tau, x_{t,y}^{u(\cdot)}(\tau,x), y'\right)$ is used next, where

(5) $$x_{t,y}^{u(\cdot)}(s,x) = x + \int_t^s a\left(r, x_{t,y}^{u(\cdot)}(r,x), y, u(r;t,x,y)\right)dr\,, \qquad t \le s \le T\,.$$

Ideally, one could expect to find continuous $u(\cdot;t,x,y)$ for all $(t,x,y) \in Q$ (then $x(s)$ would be continuously differentiable between the jumps of $y(s)$). However, it would be more realistic to search for piecewise continuous $u(\cdot;t,x,y)$, and then $x(s)$ would be continuous with piecewise continuous derivative.

Without loss of generality, we can assume that $u(\cdot;t,x,y)$ is continuous from the right for all $(t,x,y) \in Q$.

Let $\mathcal{U}$ be the set of all collections of piecewise continuous intrajump open loop (deterministic) control functions

(6) $$\mathcal{U} = \left\{u(\cdot) = u(\cdot;t,x,y) : Q \to PC([t,T);U)\right\},$$

where $PC([t,T);U)$ denotes the set of all piecewise continuous functions $u(s):[t,T) \to U$ continuous from the right.

For $y \in Y$, $t < s$, let

$$\Lambda(t,s,y) = P\left(\tau_t^1 > s \mid y(t) = y\right) = \exp\left\{-\int_t^s \lambda(r,y)dr\right\}.$$

Then

$$P\left(\tau_t^1 \le s, y\left(\tau_t^1\right) \in A \mid y(t) = y\right) = \int_t^s \Lambda(t,r,y)\Pi(r,y,A \setminus \{y\})dr\,.$$

Let $\mathcal{D}_A$ be the set of all functions $f(t,x,y):\overline{Q} \to R$ that are bounded and continuous in (t, x) uniformly in y, and such that $f_x(t,x,y)$ exists for all $(t,x,y) \in \overline{Q}$, and $f_x(t,x,y)$ is also bounded and continuous in (t, x) uniformly in y.

For $v \in U$ define

(7) $$A^v f(t,x,y) = f_x(t,x,y) \cdot a(t,x,y,v) - G_t f(t,x,y)$$

for functions $f(t,x,y) \in \mathcal{D}_A$, where

$$G_t \phi(y) = -\int_Y [\phi(y') - \phi(y)] \Pi(t,y,dy').$$

Lemma 1.

a) For every $(t,x,y) \in \overline{Q}$,

$$|V(t,x,y)| \le C_1(T-t) + D_1.$$

b) For every $t \in [0,T]$, $y \in Y$ and every $x, x' \in R^n$,

$$|V(t,x,y) - V(t,x',y)| \le M_R |x - x'|,$$

whenever $|x| \le R$, $|x'| \le R$, where

$$M_R = \left(\frac{C_2(R_1)}{B_2(R_1)} + D_2(R_1) \right) e^{B_2(R_1)(T-t)} - \frac{C_2(R_1)}{B_2(R_1)},$$

where $R_1 = (1+R)e^{B_1(T-t)} - 1$.

Proof:

a) Since $\Phi(s,x,y,u) \ge -C_1$ and $\Psi(x,y) \ge -D_1$, $V(t,x,y) \ge -C_1(T-t) - D_1$.

On the other hand, by choosing $u(s;t',x',y') = u_0$ for all $(t',x',y') \in Q$, $t' \le s < T$,

$$V(t,x,y) = E_{t,x,y}\left\{\int_t^T \Phi(s,x(s),y(s),u_0)ds + \Psi(x(T),y(T))\right\} \le C_1(T-t) + D_1 .$$

b) For any control $u(\cdot) \in \mathcal{U}$, $t \le s \le T$,

$$1 + |x(s)| \le (1 + |x(t)|) + \int_t^s |a(r,x(r),y(r),u(r))| dr \le (1 + |x(t)|) + \int_t^s B_1(1 + |x(r)|)dr .$$

Thus,

$$|x(s)| \le (1 + |x(t)|)e^{B_1(s-t)} - 1 \text{ for } t \le s \le T,$$

and therefore, $|x(s)| \le R_1$, $t \le s \le T$, whenever $|x(t)| \le R$.

For a fixed $\varepsilon > 0$, there exist a control $u_\varepsilon(\cdot) \in \mathcal{U}$ such that

$$J(t,x,y,u_\varepsilon(\cdot)) \le V(t,x,y) + \varepsilon \qquad \text{for all } (t,x,y) \in Q .$$

Let $|x| \le R$, $|x'| \le R$, and let $x(s)$ be the solution of the system of differential equations

$$\frac{dx(s)}{ds} = a(s,x(s),y(s),u_\varepsilon(s)), \qquad t \le s \le T,$$

$$x(t) = x \in R^n .$$

Let $u'(\cdot)$ be such that

$$u'(r;\, s,\, x,\, y) = u_\varepsilon(r;\, s,\, x(s),\, y) \qquad \text{for all } (t,x,y) \in Q .$$

Let $x'(s)$ be the solution of the system of differential equations

$$\frac{dx'(s)}{ds} = a(s,x'(s),y(s),u'(s)), \qquad t \le s \le T,$$

$$x'(t) = x' \in R^n,$$

Then

$$|x(s) - x'(s)| \le |x - x'| + \int_t^s B_2(R_1)|x(r) - x'(r)|dr, \; t \le s \le T.$$

Therefore,

$$|x(s) - x'(s)| \le |x - x'| e^{B_2(R_1)(s-t)}, \qquad t \le s \le T.$$

Hence,

$$\left| \int_t^T \Phi(s, x(s), y(s), u_\varepsilon(s)) ds - \int_t^T \Phi(s, x'(s), y(s), u'(s)) ds \right| \le$$

$$\le C_2(R_1)|x - x'| \int_t^T e^{B_2(R_1)(s-t)} ds = \frac{C_2(R_1)}{B_2(R_1)} \left(e^{B_2(R_1)(T-t)} - 1 \right) |x - x'|.$$

Also

$$|\Psi(x(T), y(T)) - \Psi(x'(T), y(T))| \le D_2(R_1) e^{B_2(R_1)(T-t)} |x - x'|.$$

Therefore,

$$V(t, x, y) - V(t, x', y) \le J(t, x, y, u_\varepsilon(\cdot)) - J(t, x', y, u'(\cdot)) \le M_R |x - x'|,$$

and the assertion immediately follows after a similar symmetric argument.

Lemma 1 (b) states that $V(t, x, y)$ is locally Lipschitz in x. Then by Rademacher's Theorem, $V_x(t, x, y)$ exists for almost all $x \in R^n$, for all $t \in [0, T]$, $y \in Y$.

CHAPTER 2

DYNAMIC PROGRAMMING

Lemma 2.

For all $(t,x,y) \in Q$ and $t \le s \le T$,

$$V(t,x,y) = \inf_{u(\cdot)} E_{t,x,y}\left\{ \int_t^{s\wedge\tau_t^1} \Phi(r,x(r),y(r),u(r))dr + V\left(s\wedge\tau_t^1, x(s\wedge\tau_t^1), y(s\wedge\tau_t^1)\right)\right\}.$$

Proof:

Fix a control $u(\cdot) \in \mathcal{U}$.

For any $\varepsilon > 0$, there exists control $u_{\varepsilon,t}(\cdot) \in \mathcal{U}$ such that for every

$(t',x',y') \in (t,T) \times R^n \times Y$

$$J\left(t',x',y',u_{\varepsilon,t}(\cdot)\right) \le V(t',x',y') + \varepsilon.$$

Define

$$\tilde{u}_{\varepsilon,t,s}(r;t',x',y') = u(r;t',x',y') \cdot 1_{\{r<s,t'\le t\}} + u_{\varepsilon,t}(r;t',x',y') \cdot \left(1 - 1_{\{r<s,t'\le t\}}\right)$$

for all $(t',x',y') \in Q$ and $t \le r < T$.

Then

$$V(t,x,y) \le J\left(t,x,y,\tilde{u}_{\varepsilon,t,s}(\cdot)\right) =$$

$$= E_{t,x,y}\left\{ \int_t^{s\wedge\tau_t^1} \Phi(r,x(r),y(r),u(r))dr + J\left(s\wedge\tau_t^1, x(s\wedge\tau_t^1), y(s\wedge\tau_t^1), u_{\varepsilon,t}(\cdot)\right)\right\} \le$$

$$\le E_{t,x,y}\left\{ \int_t^{s\wedge\tau_t^1} \Phi(r,x(r),y(r),u(r))dr + V\left(s\wedge\tau_t^1, x(s\wedge\tau_t^1), y(s\wedge\tau_t^1)\right)\right\} + \varepsilon$$

Since $\varepsilon > 0$ is arbitrary, we get

$$V(t,x,y) \le E_{t,x,y}\left\{ \int_t^{s\wedge\tau_t^1} \Phi(r,x(r),y(r),u(r))dr + V\left(s\wedge\tau_t^1, x(s\wedge\tau_t^1), y(s\wedge\tau_t^1)\right)\right\}$$

for any $u(\cdot) \in \mathcal{U}$.

On the other hand, for any $\varepsilon > 0$ there exists $\hat{u}_\varepsilon(\cdot) \in \mathcal{U}$ for which

$$\varepsilon + V(t,x,y) \ge J\left(t,x,y,\hat{u}_\varepsilon(\cdot)\right) =$$

$$= E_{t,x,y}\left\{ \int_t^{s\wedge\tau_t^1} \Phi(r,x(r),y(r),u(r))dr + J\left(s\wedge\tau_t^1, x(s\wedge\tau_t^1), y(s\wedge\tau_t^1), \hat{u}_\varepsilon(\cdot)\right)\right\} \ge$$

$$\ge E_{t,x,y}\left\{ \int_t^{s\wedge\tau_t^1} \Phi(r,x(r),y(r),u(r))dr + V\left(s\wedge\tau_t^1, x(s\wedge\tau_t^1), y(s\wedge\tau_t^1)\right)\right\}.$$

Remark.

Only $u(\cdot;t,x,y)$ is used on $\left[t, s\wedge\tau_t^1\right)$, no transition to the next intrajump control occurs. Therefore, the assertion of Lemma 2 could be restated in the following way:

$$V(t,x,y) = \inf_{u(\cdot;t,x,y)} E_{t,x,y}\left\{ \int_t^{s\wedge\tau_t^1} \Phi\left(r, x_{t,y}^{u(\cdot)}(r,x), y, u(r;t,x,y)\right) dr + \right.$$

$$\left. + V\left(s\wedge\tau_t^1, x_{t,y}^{u(\cdot)}\left(s\wedge\tau_t^1, x\right), y(s\wedge\tau_t^1)\right)\right\},$$

where the infimum is taken over all $u(\cdot;t,x,y) \in PC([t,T);U)$.

Theorem 1.

If $V(t,x,y) \in \mathcal{D}_A$,

then $V_t(t,x,y)$ exists for all $(t,x,y) \in Q$, and $V(t, x, y)$ satisfies the dynamic programming equation

(8) $$V_t(t,x,y) + \inf_{v\in U}\left\{\Phi(t,x,y,v) + A^v V(t,x,y)\right\} = 0, \qquad (t,x,y) \in Q,$$

and boundary (terminal) condition

(9) $$V(T,x,y) = \Psi(x,y).$$

Proof:

Let $(t,x,y) \in Q$ and $t \le s \le T$.

Then for $u(\cdot) \in \mathcal{U}$,

$$E_{t,x,y}\left\{ \int_t^{s\wedge\tau_t^1} \Phi(r, x(r), y(r), u(r)) dr + V\left(s\wedge\tau_t^1, x(s\wedge\tau_t^1), y(s\wedge\tau_t^1)\right)\right\} =$$

$$= \Lambda(t,s,y)\left[\int_t^s \Phi\left(r, x_{t,y}^{u(\cdot)}(r,x), y, u(r;t,x,y)\right) dr + V\left(s, x_{t,y}^{u(\cdot)}(s,x), y\right)\right] +$$

$$+\int_t^s \Lambda(t,w,y)\lambda(w,y)\left[\int_t^w \Phi\left(r,x_{t,y}^{u(\cdot)}(r,x),y,u(r;t,x,y)\right)dr\right]dw+$$

$$+\int_t^s \Lambda(t,w,y)\left[\int_{Y\setminus\{y\}} V\left(w,x_{t,y}^{u(\cdot)}(w,x),y'\right)\Pi(w,y,dy')\right]dw=$$

$$=\int_t^s \Lambda(t,r,y)\Phi\left(r,x_{t,y}^{u(\cdot)}(r,x),y,u(r;t,x,y)\right)dr+\Lambda(t,s,y)V\left(s,x_{t,y}^{u(\cdot)}(s,x),y\right)+$$

$$+\int_t^s \Lambda(t,r,y)\left[\int_{Y\setminus\{y\}} V\left(r,x_{t,y}^{u(\cdot)}(r,x),y'\right)\Pi(r,y,dy')\right]dr.$$

Then Lemma 2 and the Fundamental Theorem of Calculus imply

$$0=\Lambda(t,s,y)[V(s,x,y)-V(t,x,y)]+$$

$$+\inf_{u(\cdot;t,x,y)}\left\{\int_t^s \Lambda(t,r,y)\Phi\left(r,x_{t,y}^{u(\cdot)}(r,x),y,u(r;t,x,y)\right)dr+\right.$$

$$+\int_t^s \Lambda(t,s,y)V_x\left(s,x_{t,y}^{u(\cdot)}(r,x),y\right)\cdot a\left(r,x_{t,y}^{u(\cdot)}(r,x),y,u(r;t,x,y)\right)dr+$$

$$\left.+\int_t^s \Lambda(t,r,y)\int_{Y\setminus\{y\}}\left[V\left(r,x_{t,y}^{u(\cdot)}(r,x),y'\right)-V(t,x,y)\right]\Pi(r,y,dy')dr\right\}.$$

For any $u(\cdot;t,x,y)\in PC([t,T);U)$,

$$\lim_{s\to t}\frac{1}{s-t}\int_t^s \Lambda(t,r,y)\Phi\left(r,x_{t,y}^{u(\cdot)}(r,x),y,u(r;t,x,y)\right)dr=\Phi(r,x,y,u(t;t,x,y)),$$

$$\lim_{s\to t}\frac{1}{s-t}\int_t^s \Lambda(t,s,y)V_x\left(s,x_{t,y}^{u(\cdot)}(r,x),y\right)\cdot a\left(r,x_{t,y}^{u(\cdot)}(r,x),y,u(r;t,x,y)\right)dr=$$

$$=V_x(t,x,y)\cdot a(t,x,y,u(t;t,x,y)),$$

$$\lim_{s\to t}\frac{1}{s-t}\int_t^s \Lambda(t,r,y)\int_{Y\setminus\{y\}}\left[V\left(r,x_{t,y}^{u(\cdot)}(r,x),y'\right)-V(t,x,y)\right]\Pi(r,y,dy')dr =$$

$$= -G_t V(t,x,y).$$

Therefore, $V_t(t,x,y)$ exists and

$$V_t(t,x,y)+\inf_{v\in U}\left\{\Phi(t,x,y,v)+A^v V(t,x,y)\right\}=0.$$

The boundary (terminal) condition immediately follows from the definition of $V(t,x,y)$.

In an easier case when $y(s)$ is a finite-state Markov chain, the dynamic programming equation has the form

$$V_t(t,x,y)+\inf_{v\in U}\left\{\Phi(t,x,y,v)+V_x(t,x,y)\cdot a(t,x,y,v)\right\}+$$

$$+\sum_{y'\in Y}\left[V(t,x,y')-V(t,x,y)\right]\Pi(t,y,\{y'\})=0,$$

a system of partial differential equations indexed by $y\in Y$ (note that Y is a finite set now).

Fleming and Rishel [7] derive the continuous-time dynamic programming equation of optimal stochastic control theory

$$V_t(t,y)+\min_{v\in U}\left\{\Phi(t,y,v)+A^v(s)V(t,y)\right\}=0,$$

for a somewhat more general case, where $A^{v}(s)$ is the generator of the controlled Markov process $y(s)$. In our problem, the control process $x(s)$ is not a Markov process, but the two-component process $(x(s), y(s))$ is a Markov process, and A^{v} is its generator.

Lemma 3. (Dynkin's Formula)

Let $f(t,x,y) \in \mathcal{D}_A$, and suppose $f_t(t,x,y)$ exists and is bounded and continuous in (t, x) uniformly in y.

Let $u(\cdot) \in \mathcal{U}$, and let $x(s)$ be the corresponding trajectory of the system.

Then for $0 \le t \le s \le T$

$$E_{t,x,y}\{f(s,x(s),y(s))\} - f(t,x,y) =$$

$$= E_{t,x,y}\left\{\int_t^s \left[f_t(r,x(r),y(r)) + A^{u(r)} f(r,x(r),y(r))\right] dr\right\}.$$

Proof:

$$f(s,x(s),y(s)) - f(t,x,y) =$$

$$= \sum_{k\ge 0}\left[f\left(s\wedge\tau_t^{k+1}, x(s\wedge\tau_t^{k+1}), y(s\wedge\tau_t^{k})\right) - f\left(s\wedge\tau_t^{k}, x(s\wedge\tau_t^{k}), y(s\wedge\tau_t^{k})\right)\right] +$$

$$+ \sum_{k\ge 0}\left[f\left(s\wedge\tau_t^{k+1}, x(s\wedge\tau_t^{k+1}), y(s\wedge\tau_t^{k+1})\right) - f\left(s\wedge\tau_t^{k+1}, x(s\wedge\tau_t^{k+1}), y(s\wedge\tau_t^{k})\right)\right] =$$

$$= \Sigma_1 + \Sigma_2 .$$

From the Fundamental Theorem of Calculus

$$E_{t,x,y}\{\Sigma_1\} = E_{t,x,y}\left\{\int_t^s\left[f_t(r,x(r),y(r)) + f_x(r,x(r),y(r))\cdot a(r,x(r),y(r),u(r))\right]dr\right\}.$$

Σ_2 requires more careful consideration.

$$E_{t,x,y}\left\{f\left(s\wedge\tau_t^{k+1},x(s\wedge\tau_t^{k+1}),y(s\wedge\tau_t^{k+1})\right) - f\left(s\wedge\tau_t^{k+1},x(s\wedge\tau_t^{k+1}),y(s\wedge\tau_t^{k})\right)\right\} =$$

$$= E_{t,x,y}\left\{\left[f\left(s\wedge\tau_t^{k+1},x(s\wedge\tau_t^{k+1}),y_{k+1}\right) - f\left(s\wedge\tau_t^{k+1},x(s\wedge\tau_t^{k+1}),y_k\right)\right]1_{\left\{\tau_t^{k+1}\le s\right\}}\right\} =$$

$$= -E_{t,x,y}\left\{1_{\left\{\tau_t^k\le s\right\}}\int_{\tau_t^k}^s \Lambda\left(\tau_t^k,r,y_k\right)G_r f\left(r,x(r),y_k\right)dr\right\} =$$

$$= -E_{t,x,y}\left\{\int_{s\wedge\tau_t^k}^s \Lambda\left(s\wedge\tau_t^k,r,y_k\right)G_r f\left(r,x(r),y_k\right)dr\right\}.$$

Also

$$E_{t,x,y}\left\{\int_{s\wedge\tau_t^k}^{s\wedge\tau_t^{k+1}} G_r f\left(r,x(r),y_k\right)dr\right\} = E_{t,x,y}\left\{1_{\left\{\tau_t^{k+1}\le s\right\}}\int_{\tau_t^k}^{\tau_t^{k+1}} G_r f\left(r,x(r),y_k\right)dr\right\} +$$

$$+ E_{t,x,y}\left\{1_{\left\{\tau_t^k\le s<\tau_t^{k+1}\right\}}\int_{\tau_t^k}^{s} G_r f\left(r,x(r),y_k\right)dr\right\} = E_1 + E_2.$$

$$E_1 = E_{t,x,y}\left\{1_{\left\{\tau_t^k\le s\right\}}\int_{\tau_t^k}^s \Lambda\left(\tau_t^k,w,y_k\right)\lambda\left(w,y_k\right)\left[\int_{\tau_t^k}^{w} G_r f\left(r,x(r),y_k\right)dr\right]dw\right\} =$$

$$= E_{t,x,y}\left\{1_{\left\{\tau_t^k\le s\right\}}\int_{\tau_t^k}^s \left[\int_r^s \Lambda\left(\tau_t^k,w,y_k\right)\lambda\left(w,y_k\right)dw\right]G_r f\left(r,x(r),y_k\right)dr\right\} =$$

$$= E_{t,x,y}\left\{1_{\left\{\tau_t^k\le s\right\}}\int_{\tau_t^k}^s \left[\Lambda\left(\tau_t^k,r,y_k\right) - \Lambda\left(\tau_t^k,s,y_k\right)\right]G_r f\left(r,x(r),y_k\right)dr\right\} =$$

$$= E_{t,x,y}\left\{ \int\limits_{s\wedge\tau_t^k}^{s} \left[\Lambda\left(s\wedge\tau_t^k, r, y_k\right) - \Lambda\left(s\wedge\tau_t^k, s, y_k\right)\right] G_r f\left(r, x(r), y_k\right) dr \right\}.$$

$$E_2 = E_{t,x,y}\left\{ 1_{\{\tau_t^k \le s\}} \Lambda\left(\tau_t^k, s, y_k\right) \int\limits_{\tau_t^k}^{s} G_r f\left(r, x(r), y_k\right) dr \right\} =$$

$$= E_{t,x,y}\left\{ \Lambda\left(s\wedge\tau_t^k, s, y_k\right) \int\limits_{s\wedge\tau_t^k}^{s} G_r f\left(r, x(r), y_k\right) dr \right\}.$$

Therefore,

$$E_{t,x,y}\{\Sigma_2\} = -E_{t,x,y}\left\{ \int\limits_{t}^{s} G_r f\left(r, x(r), y(r)\right) dr \right\},$$

and the assertion of the lemma immediately follows.

Theorem 2. (Verification Theorem)

Let $W(t,x,y) \in \mathcal{D}_A$ be such that $W_t(t,x,y)$ exists for all $(t,x,y) \in Q$, and suppose that $W_t(t,x,y)$ is bounded and continuous in (t, x) uniformly in y.

a) Suppose $W(t,x,y)$ satisfies for all $(t,x,y) \in Q$ the inequality

$$W_t(t,x,y) + \inf_{v \in U}\left\{\Phi(t,x,y,v) + A^v W(t,x,y)\right\} \ge 0$$

and boundary (terminal) condition

$$W(T,x,y) = \Psi(x,y).$$

Then for all $(t,x,y) \in \overline{Q}$ and $u(\cdot) \in \mathcal{U}$

$$W(t,x,y) \le J(t,x,y,u(\cdot)),$$

and therefore,

$$W(t,x,y) \le V(t,x,y).$$

b) If we can find for every $(t,x,y) \in Q$, $\underline{u}^*(t,x,y) \in U$ such that

$$W_t(t,x,y) + \Phi\left(t,x,y,\underline{u}^*(t,x,y)\right) + A^{\underline{u}^*(t,x,y)} W(t,x,y) = 0,$$

and if

$$u^*(\cdot) = \left\{u^*(s;t,x,y) = \underline{u}^*\left(s, x_{t,y}^{u(\cdot)}(s,x), y\right), (t,x,y) \in Q, t \le s < T\right\} \in \mathcal{U},$$

then $u^*(\cdot)$ is optimal.

c) If we can find $\underline{u}_\varepsilon(t,x,y) \in U$ such that

$$W_t(t,x,y) + \Phi(t,x,y,\underline{u}_\varepsilon(t,x,y)) + A^{\underline{u}_\varepsilon(t,x,y)} W(t,x,y) \le \frac{\varepsilon}{T},$$

and if

$$u_\varepsilon(\cdot) = \left\{u_\varepsilon(s;t,x,y) = \underline{u}_\varepsilon\left(s, x_{t,y}^{u(\cdot)}(s,x), y\right), (t,x,y) \in Q, t \le s < T\right\} \in \mathcal{U},$$

then $u_\varepsilon(\cdot)$ is ε - optimal, i.e.

$$J(t,x,y,u_\varepsilon(\cdot)) \le V(t,x,y) + \varepsilon \qquad \text{for all } (t,x,y) \in \overline{Q}.$$

Proof:

a) Let $u(\cdot) \in \mathcal{U}$.

Since for all $(t,x,y) \in Q$,

$$W_t(t,x,y) + A^{u(t)} W(t,x,y) \ge -\Phi(t,x,y,u(t)),$$

we obtain by Lemma 3

$$E_{t,x,y}\left\{\Psi(x(T),y(T))\right\} = E_{t,x,y}\left\{W(T,x(T),y(T))\right\} =$$

$$= W(t,x,y) + E_{t,x,y}\left\{\int_t^T\left[W_t(s,x(s),y(s)) + A^{u(s)}W(s,x(s),y(s))\right]ds\right\} \geq$$

$$\geq W(t,x,y) - E_{t,x,y}\left\{\int_t^T\Phi(s,x(s),y(s),u(s))ds\right\},$$

and the assertion immediately follows.

b) The same line of arguments works, only the inequality becomes equality.

c) Again, the same line of arguments as in (a) works. We obtain with the help of (a)

$$J(t,x,y,u_\varepsilon(\cdot)) \leq W(t,x,y) + \varepsilon \leq V(t,x,y) + \varepsilon .$$

CHAPTER 3

REDUCTION TO A FAMILY OF DETERMINISTIC CONTROL PROBLEMS

Let $\mathcal{E}$ denote the set of all functions $f(t,x,y):\overline{Q}\to R$ such that

$|f(t,x,y)|\le C_1(T-t)+D_1$ for all $(t, x, y) \in Q$, and let

$$\mathcal{U}_0 = \{u(\cdot) \in \mathcal{U} :\ J(t,x,y,u(\cdot)) \in \mathcal{E}\}.$$

Notice that $\mathcal{U}_0$ is non-empty since $u_0(\cdot)$ for which $u_0(s; t, x, y) = u_0$ for all $(t, x, y) \in Q$, $t \le s < T$, is in $\mathcal{U}_0$, and if there is an optimal control in the class $\mathcal{U}$, then that control is in $\mathcal{U}_0$. Lemma 1 (a) states that $V(t, x, y) \in \mathcal{E}$.

For every $u(\cdot) \in \mathcal{U}$, define the function $\hat{J}(t,x,y,u(\cdot))$: $[0, T] \times R^n \times Y \to R$ by

$$\hat{J}(t,x,y,u(\cdot))=\int_t^T \Lambda(t,s,y)\Phi\left(s,x_{t,y}^{u(\cdot)}(s,x),y,u(s;t,x,y)\right)ds+\Lambda(t,T,y)\Psi\left(x_{t,y}^{u(\cdot)}(T,x),y\right).$$

For every $u(\cdot) \in \mathcal{U}$, define operators

$$T^{u(\cdot)}f(t,x,y)=\hat{J}(t,x,y,u(\cdot))+\int_t^T \Lambda(t,s,y)\left\{\int_{Y\setminus\{y\}} f\left(s,x_{t,y}^{u(\cdot)}(s,x),y'\right)\Pi(s,y,dy')\right\}ds,$$

and define

$$T\, f(t,x,y) = \inf_{u(\cdot;t,x,y)} \left\{ \hat{J}(t,x,y,u(\cdot)) + \int_t^T \Lambda(t,s,y) \left\{ \int_{Y \setminus \{y\}} f\left(s, x_{t,y}^{u(\cdot)}(s,x), y'\right) \Pi(s,y,dy') \right\} ds \right\},$$

where the infimum is taken over all $u(\cdot;t,x,y) \in PC([t,T);U)$.

Lemma 4.

a) For any control $u(\cdot) \in \mathcal{U}$, $J(t,x,y,u(\cdot))$ satisfies the integral equation

$$J(t,x,y,u(\cdot)) = T^{u(\cdot)} J(t,x,y,u(\cdot))$$

i.e.

$$J(t,x,y,u(\cdot)) = \int_t^T \Lambda(t,s,y) \Phi\left(s, x_{t,y}^{u(\cdot)}(s,x), y, u(s;t,x,y)\right) ds + \Lambda(t,T,y) \Psi\left(x_{t,y}^{u(\cdot)}(T,x), y\right) +$$

$$+ \int_t^T \Lambda(t,s,y) \left\{ \int_{Y \setminus \{y\}} J\left(s, x_{t,y}^{u(\cdot)}(s,x), y', u(\cdot)\right) \Pi(s,y,dy') \right\} ds .$$

b) For any control $u(\cdot) \in \mathcal{U}_0$, if $W(t,x,y) \in \mathcal{E}$ satisfies $W(t,x,y) = T^{u(\cdot)} W(t,x,y)$,

then

$$W(t,x,y) = J(t,x,y,u(\cdot)).$$

c) $u(\cdot) \in \mathcal{U}$ is optimal if and only if $V(t,x,y) = T^{u(\cdot)} V(t,x,y)$.

Proof:

a) Fix a control $u(\cdot) \in \mathcal{U}$. Then

$$J(t,x,y,u(\cdot)) = E_{t,x,y}\left[1_{\{\tau_t^1 = T\}} \left\{ \int_t^T \Phi(s, x(s), y(s), u(s)) ds + \Psi(x(T), y(T)) \right\} \right] +$$

$$+ E_{t,x,y}\left[1_{\{\tau_t^1 < T\}} \left\{ \int_t^T \Phi(s, x(s), y(s), u(s)) ds + \Psi(x(T), y(T)) \right\} \right] = E_1 + E_2.$$

Straightforward calculations yield

$$E_1 = \Lambda(t,T,y) \left[\int_t^T \Phi\left(s, x_{t,y}^{u(\cdot)}(s,x), y, u(s;t,x,y)\right) ds + \Psi\left(x_{t,y}^{u(\cdot)}(T,x), y\right) \right].$$

$$E_2 = \int_t^T \Lambda(t,s,y) \lambda(s,y) \left[\int_t^s \Phi\left(r, x_{t,y}^{u(\cdot)}(r,x), y, u(r;t,x,y)\right) dr \right] ds +$$

$$+ E_{t,x,y}\left[1_{\{\tau_t^1 < T\}} J\left(\tau_t^1, x_{t,y}^{u(\cdot)}\left(\tau_t^1, x\right), y\left(\tau_t^1\right), u(\cdot)\right) \right] =$$

$$= \int_t^T \{\Lambda(t,r,y) - \Lambda(t,T,y)\} \Phi\left(r, x_{t,y}^{u(\cdot)}(r,x), y, u(r;t,x,y)\right) dr +$$

$$+ \int_t^T \Lambda(t,s,y) \left\{ \int_{Y \setminus \{y\}} J\left(s, x_{t,y}^{u(\cdot)}(s,x), y', u(\cdot)\right) \Pi(s,y,dy') \right\} ds,$$

and the assertion immediately follows.

b) Suppose $\sup\limits_{(t,x,y)} |W(t,x,y) - J(t,x,y,u(\cdot))| = \eta > 0$, and $(t, x, y) \in Q$ is such that

$$|W(t,x,y) - J(t,x,y,u(\cdot))| \geq \eta - e^{-K \cdot T} \cdot \eta/2 .$$

Then

$$\eta - e^{-K \cdot T} \cdot \eta/2 \leq |W(t,x,y) - J(t,x,y,u(\cdot))| = \left| T^{u(\cdot)} W(t,x,y) - T^{u(\cdot)} J(t,x,y,u(\cdot)) \right| =$$

$$= \left| \int_t^T \Lambda(t,s,y) \left\{ \int_{Y \setminus \{y\}} \left[W\left(s, x_{t,y}^{u(\cdot)}(s,x), y'\right) - V\left(s, x_{t,y}^{u(\cdot)}(s,x), y'\right) \right] \Pi(s,y,dy') \right\} ds \right| \leq$$

$$\leq \left(1 - \Lambda(t,T,y)\right)\eta \leq \left(1 - e^{-K \cdot T}\right) \cdot \eta . \qquad \text{Contradiction.}$$

Therefore, $W(t,x,y) = J\left(t,x,y,u(\cdot)\right)$.

c) Immediately follows from (a) and (b).

Corollary.

Finding the optimal $u(\cdot) \in \mathcal{U}$ is equivalent to finding, for each $(t, x, y) \in Q$, an optimal deterministic open loop control function $u(\cdot; t, x, y)$ minimizing

$$\int_t^T \Lambda(t,s,y) \tilde{\Phi}\left(s, x_{t,y}^{u(\cdot)}(s,x), y, u(s;t,x,y)\right) ds + \Lambda(t,T,y) \Psi\left(x_{t,y}^{u(\cdot)}(T,x), y\right),$$

where

$$\tilde{\Phi}(s,x,y,u) = \Phi(s,x,y,u) + \int_{Y \setminus \{y\}} V(s,x,y') \Pi(s,y,dy') .$$

A wider class of controls could be considered. Let us consider controls that have one functional form until a jump of $y(s)$ and then another functional form after each jump.

Let

$$\mathcal{M} = \left\{ u(\cdot) = u_k(\cdot; t, x, y) : Q \to PC\left([t,T); U\right), k = 0, 1, 2, \ldots \right\},$$

Define

$\nu(t) = \{$number of jumps of $y(s)$ on $[0, t]\}$.

The intrajump control function $u_{\nu(t)}(\cdot;t,x,y)$ is used from time t up to the terminal time T or the time of the first jump of $y(s)$ after t, whatever comes first. Now the control functions and thus the state of the system x(s) depend on the past of y(s).

Notice that controls $u(\cdot) \in \mathcal{U}$ also are in class $\mathcal{M}$, they are constant in the index k.

It can be seen from Corollary 4.1 that if there is an optimal control in the class $\mathcal{U}$, then that control is also optimal in class $\mathcal{M}$, and therefore, no advantage can be gained by going to the wider class $\mathcal{M}$.

<u>Theorem 3</u>. (Integral equation for the value function $V(t, x, y)$)

a) $V(t, x, y)$ satisfies the integral equation

(10) $$V(t,x,y) = T\,V(t,x,y)$$

i.e.

$$V(t,x,y) = \inf_{u(\cdot;t,x,y)} \left\{ \int_t^T \Lambda(t,s,y)\Phi\left(s, x_{t,y}^{u(\cdot)}(s,x), y, u(s;t,x,y)\right)ds + \Lambda(t,T,y)\Psi\left(x_{t,y}^{u(\cdot)}(T,x), y\right) + \right.$$

$$\left. + \int_t^T \Lambda(t,s,y)\left\{ \int_{\mathcal{Y}\setminus\{y\}} V\left(s, x_{t,y}^{u(\cdot)}(s,x), y'\right)\Pi(s,y,dy')\right\}ds \right\},$$

where the infimum is taken over all $u(\cdot;t,x,y) \in PC([t,T);U)$.

b) If $W(t, x, y) \in \mathcal{E}$ satisfies $W(t,x,y) = T\,W(t,x,y)$, then $W(t, x, y) = V(t, x, y)$.

c) $u(\cdot) \in \mathcal{U}$ is optimal if and only if $J(t,x,y,u(\cdot)) = T\,J(t,x,y,u(\cdot))$.

Proof:

Let $(t, x, y) \in Q$. Fix a control $u(\cdot) \in \mathcal{U}$.

For any $\varepsilon > 0$, there exists control $u_{\varepsilon,t}(\cdot) \in \mathcal{U}$ such that for every

$(t', x', y') \in (t, T) \times R^n \times Y$

$$J(t', x', y', u_{\varepsilon,t}(\cdot)) \le V(t', x', y') + \varepsilon .$$

Define

$$\tilde{u}_{\varepsilon,t}(r; t', x', y') = u(r; t', x', y') \cdot 1_{\{t' \le t\}} + u_{\varepsilon,t}(r; t', x', y') \cdot 1_{\{t' > t\}}$$

for all $(t', x', y') \in Q$ and $t \le r < T$.

Then by Lemma 4

$$V(t, x, y) \le J\left(t, x, y, \tilde{u}_{\varepsilon,t}(\cdot)\right) =$$

$$= \int_t^T \Lambda(t, s, y)\Phi\left(s, x_{t,y}^{u(\cdot)}(s, x), y, u(s; t, x, y)\right)ds + \Lambda(t, T, y)\Psi\left(x_{t,y}^{u(\cdot)}(T, x), y\right) +$$

$$+ \int_t^T \Lambda(t, s, y)\left\{ \int_{Y \setminus \{y\}} J\left(s, x_{t,y}^{u(\cdot)}(s, x), y', u_{\varepsilon,t}(\cdot)\right)\Pi(s, y, dy')\right\}ds \le$$

$$\le \int_t^T \Lambda(t, s, y)\Phi\left(s, x_{t,y}^{u(\cdot)}(s, x), y, u(s; t, x, y)\right)ds + \Lambda(t, T, y)\Psi\left(x_{t,y}^{u(\cdot)}(T, x), y\right) +$$

$$+ \int_t^T \Lambda(t, s, y)\left\{ \int_{Y \setminus \{y\}} V\left(s, x_{t,y}^{u(\cdot)}(s, x), y'\right)\Pi(s, y, dy')\right\}ds + \varepsilon \cdot (1 - \Lambda(t, T, y))$$

Since $\varepsilon > 0$ is arbitrary, we get

$$V(t,x,y) \le \int_t^T \Lambda(t,s,y)\Phi\left(s, x_{t,y}^{u(\cdot)}(s,x), y, u(s;t,x,y)\right)ds + \Lambda(t,T,y)\Psi\left(x_{t,y}^{u(\cdot)}(T,x), y\right) +$$

$$+ \int_t^T \Lambda(t,s,y)\left\{ \int_{Y\setminus\{y\}} V\left(s, x_{t,y}^{u(\cdot)}(s,x), y'\right)\Pi(s,y,dy')\right\}ds \Bigg\},$$

for any $u(\cdot;t,x,y) \in PC([t,T);U)$.

Also by Lemma 4

$$J(t,x,y,u(\cdot)) = \int_t^T \Lambda(t,s,y)\Phi\left(s, x_{t,y}^{u(\cdot)}(s,x), y, u(s;t,x,y)\right)ds + \Lambda(t,T,y)\Psi\left(x_{t,y}^{u(\cdot)}(T,x), y\right) +$$

$$+ \int_t^T \Lambda(t,s,y)\left\{ \int_{Y\setminus\{y\}} J\left(s, x_{t,y}^{u(\cdot)}(s,x), y', u(\cdot)\right)\Pi(s,y,dy')\right\}ds \ge$$

$$\ge \int_t^T \Lambda(t,s,y)\Phi\left(s, x_{t,y}^{u(\cdot)}(s,x), y, u(s;t,x,y)\right)ds + \Lambda(t,T,y)\Psi\left(x_{t,y}^{u(\cdot)}(T,x), y\right) +$$

$$+ \int_t^T \Lambda(t,s,y)\left\{ \int_{Y\setminus\{y\}} V\left(s, x_{t,y}^{u(\cdot)}(s,x), y'\right)\Pi(s,y,dy')\right\}ds,$$

and the assertion follows.

b) Suppose $\sup\limits_{(t,x,y)} |W(t,x,y) - V(t,x,y)| = \eta > 0$, and $(t, x, y) \in Q$ is such that

$$|W(t,x,y) - V(t,x,y)| \ge \eta - e^{-K\cdot T} \cdot \eta/3 .$$

Without loss of generality, $W(t, x, y) > V(t, x, y)$.

There exists a control $u(\cdot) \in \mathcal{U}$ such that

$$T^{u(\cdot)}V(t,x,y) \le T\,V(t,x,y) + e^{-K\cdot T} \cdot \eta/3 .$$

Then

$$\eta - e^{-K \cdot T} \cdot \eta/3 \leq W(t,x,y) - V(t,x,y) = T\,W(t,x,y) - T\,V(t,x,y) \leq$$

$$\leq T^{u(\cdot)} W(t,x,y) - T^{u(\cdot)} V(t,x,y) + e^{-K \cdot T} \cdot \eta/3 \leq$$

$$\leq \int_t^T \Lambda(t,s,y) \left\{ \int_{Y \setminus \{y\}} \left[W\left(s, x_{t,y}^{u(\cdot)}(s,x), y'\right) - V\left(s, x_{t,y}^{u(\cdot)}(s,x), y'\right) \right] \Pi(s,y,dy') \right\} ds + e^{-K \cdot T} \cdot \eta/3 \leq$$

$$\leq \left(1 - \Lambda(t,T,y)\right)\eta + e^{-K \cdot T} \cdot \eta/3 \leq \eta - e^{-K \cdot T} \cdot 2\eta/3 .$$

Contradiction. Therefore, $W(t, x, y) = V(t, x, y)$.

c) Immediately follows from (a) and (b).

Consider $u_0(\cdot) \in \mathcal{U}_0$ for which $u_0(s;\, t, x, y) = u_0$ for all $(t, x, y) \in Q$, $t \leq s < T$.

Let $x_0(s)$ be the solution of the system of differential equations

$$\frac{dx_0(s)}{ds} = a\left(s, x_0(s), y(s), u_0\right), \qquad 0 \leq t \leq s \leq T,$$

with initial condition

$$x_0(t) = x \in R^n.$$

Define

$$(11) \quad W_0(t,x,y) = J\left(t,x,y,u_0(\cdot)\right) = E_{t,x,y}\left\{ \int_t^T \Phi\left(s, x_0(s), y(s), u_0\right) ds + \Psi\left(x_0(T), y(T)\right) \right\},$$

and for $k = 1, 2, 3, \ldots$, let

$$(12) \qquad W_k(t,x,y) = T\,W_{k-1}(t,x,y).$$

Theorem 4.

$W_k(t, x, y)$ converges to $V(t, x, y)$ from above uniformly on Q.

Proof:

It is easy to see that $W_k(t, x, y) \geq V(t, x, y)$ for all $k = 0, 1, 2, \ldots$.

Since $W_0(t, x, y) \in \mathcal{E}$ and $V(t, x, y) \in \mathcal{E}$,

$$\|W_0(t,x,y) - V(t,x,y)\| = \sup_{(t,x,y)} \{W_0(t,x,y) - V(t,x,y)\} \leq 2 \cdot [C_1 T + D_1].$$

Suppose that for some k,

$$\|W_k(t,x,y) - V(t,x,y)\| \leq 2 \cdot [C_1 T + D_1] \cdot \left(1 - e^{-K \cdot T}\right)^k .$$

Then

$W_{k+1}(t, x, y) = T\, W_k(t, x, y) \leq$

$$\leq T\, V(t, x, y) + \int_t^T \Lambda(t,s,y) \left\{ \int_{Y \setminus \{y\}} 2 \cdot [C_1 T + D_1] \cdot \left(1 - e^{-K \cdot T}\right)^k \Pi(s, y, dy') \right\} ds =$$

$$= V(t, x, y) + 2 \cdot [C_1 T + D_1] \cdot \left(1 - e^{-K \cdot T}\right)^{k+1} .$$

The assertion of the lemma follows by the mathematical induction.

CHAPTER 4

FAMILY OF CONTROL PROBLEMS WITH "STOPPED" $y(s)$

Let

$$z_t^0(s) = y(t) = y(\tau_t^0), \qquad t \le s \le T,$$

and let for $k = 1, 2, \ldots$

$$z_t^k(s) = y(s) \qquad \text{for } t \le s < \tau_t^k$$

$$z_t^k(s) = y(\tau_t^k) \qquad \text{for } \tau_t^k \le s \le T.$$

($z_t^k(s)$ is y(s) stopped at the time of the kth jump after t.)

Note that for $k = 1, 2, \ldots$, for $l = 0, 1, \ldots, k$,

$$z_t^k(s) = z_{\tau_t^l}^{k-l}(s), \qquad \tau_t^l \le s \le T.$$

For $k = 0, 1, 2, \ldots$, consider the control problem where the state of a system is described by the system of differential equations

$$\frac{dx_k(s)}{ds} = a\left(s, x_k(s), z_t^k(s), u_k(s)\right), \qquad t \le s \le T,$$

with initial condition

$$x(t) = x \in R^n,$$

and the performance criterion is given by

$$J_k\left(t, x, y, u_k(\cdot)\right) = E_{t,x,y}\left\{\int_t^T \Phi\left(s, x_k(s), z_t^k(s), u_k(s)\right)ds + \Psi\left(x_k(T), z_t^k(T)\right)\right\}.$$

Since, for $k = 0, 1, 2, \dots$, $z_t^k(s)$ has at most k jumps on $[t, T]$, let us consider controls $u_k(\cdot)$ of the form

$$u_k(s) = u_{k,l}(s) = \underline{u}_{k,l}(s; \tau_t^l, x_k(\tau_t^l), z_t^k(\tau_t^l)),$$

$$\tau_t^l \leq s < \tau_t^{l+1},\ l = 0, 1, \dots, k-1,$$

$$u_k(s) = u_{k,k}(s) = \underline{u}_{k,k}(s; \tau_t^k, x_k(\tau_t^k), z_t^k(\tau_t^k)),$$

$$\tau_t^k \leq s \leq T.$$

Let $\mathcal{U}_k$ be the set of all controls $u_k(\cdot)$ of this form. Note that $\bigcup_k \mathcal{U}_k = \mathcal{M}$.

For $k = 0, 1, 2, \dots$, consider the value function

$$V_k(t,x,y) = \inf_{u_k(\cdot)} \{J_k(t,x,y,u_k(\cdot))\}.$$

Using arguments similar to the ones used earlier, we can prove the following lemma.

<u>Lemma 5</u>.

If $J_k(t,x,y,u_k^*(\cdot)) = V_k(t,x,y)$ and $J_{k-1}(t,x,y,u_{k-1}^*(\cdot)) = V_{k-1}(t,x,y)$, then

a) $u_{k,l}^*(\cdot) = u_{k-1,l-1}^*(\cdot)$, $l = 1, 2, \dots, k$,

b) $u_{k,0}^*(\cdot)$ minimizes

$$\int_t^T \Phi_k\left(s, x_{t,y}^{u_{k,0}(\cdot)}(s,x), y, u_{k,0}(s)\right) ds + \Psi_k\left(x_{t,y}^{u_{k,0}(\cdot)}(T,x), y\right),$$

where $\Phi_0(s,x,y,u) = \Phi(s,x,y,u)$, $\Psi_0(x,y) = \Psi(x,y)$,

$$\Phi_k(s,x,y,u) = \Lambda(t,s,y)\left\{\Phi(s,x,y,u) + \int_{Y\setminus\{y\}} V_{k-1}(s,x,y')\Pi(s,y,dy')\right\},$$

$$\Psi_k(x,y) = \Lambda(t,T,y)\Psi(x,y), \qquad k = 1, 2, \ldots .$$

For every $y \in$ Y, consider the deterministic control problem where the state of a system is described by the system of differential equations

$$(13) \qquad \frac{dx(s)}{ds} = a(s,x(s),y,u(s)), \qquad t \le s \le T,$$

with initial condition

$$(14) \qquad x(t) = x \in R^n,$$

and the performance criterion is given by

$$F_0(t,x,y,u(\cdot)) = \int_t^T \Phi_k(s,x(s),y,u(s))ds + \Psi_k(x(T),y).$$

Suppose that $(\Phi_k)_x(s,x,y,u)$ exists for all $(s,x,y,u) \in (t,T) \times R^n \times Y \times U$.

Suppose that $(\hat{x}(s),\hat{u}(s))$, $t \le s \le T$, is the optimal solution.

Let $v(s)$:$[t, T] \to R^m$ be a piecewise continuous function and set

$$u(s,\delta) = \hat{u}(s) + \delta \cdot v(s), \qquad t \le s \le T.$$

Then (13) and (14) with $u(s) = u(s,\delta)$ have a one-parameter family of solutions $x(s,\delta)$, $t \le s \le T$, for $|\delta| < \delta_0$, which contains $\hat{x}(s)$ for $\delta = 0$. The functions $x(s,\delta)$ are continuous and have continuous derivatives with respect to δ.

Let

$$\Delta(s) = \frac{\partial x}{\partial \delta}(s,0), \qquad t \le s \le T.$$

Then $\Delta(s)$ satisfies the system of differential equations

$$\frac{d\Delta(s)}{ds} = a_x(s, \hat{x}(s), y, \hat{u}(s)) \cdot \Delta(s) + a_u(s, \hat{x}(s), y, \hat{u}(s)) \cdot v(s), \quad t \le s \le T,$$

with initial condition

$$\Delta(t) = \mathbf{0} \in R^n.$$

Denote by $A(s)$, $t \le s \le T$, the $n \times n$ matrix-valued function satisfying the differential equation

$$\frac{dA(s)}{ds} = a_x(s, \hat{x}(s), y, \hat{u}(s)) \cdot A(s), \qquad t \le s \le T,$$

with initial condition

$$A(t) = \mathbf{I}_n.$$

Then

$$\Delta(s) = A(s) \cdot \int A^{-1}(r) \cdot a_u(r, \hat{x}(r), y, \hat{u}(r)) \cdot v(r) dr, \qquad t \le s \le T.$$

Set $f(\delta; v(\cdot)) = F_0(t, x, y, u(\cdot, \delta))$, for $|\delta| < \delta_0$.

Then $f'(\delta; v(\cdot))$ exists and

$$f'(\delta; v(\cdot)) = (\Psi_k)_x(\hat{x}(T), y) + \int_t^T [(\Phi_k)_x(s, \hat{x}(s), y, \hat{u}(s))\Delta(s) + (\Phi_k)_u(s, \hat{x}(s), y, \hat{u}(s))v(s)] ds$$

$$= \int_t^T \varphi(r) v(r) dr,$$

where

$$\varphi(s) = (\Phi_k)_u(s, \hat{x}(s), y, \hat{u}(s))$$

(15) $$+\left\{(\Psi_k)_x(\hat{x}(T), y)A(T) + \int_s^T (\Phi_k)_x(r, \hat{x}(r), y, \hat{u}(r))A(r)dr\right\}A^{-1}(s)a_u(s, \hat{x}(s), y, \hat{u}(s)),$$

$$t \le s \le T.$$

If $(\hat{x}(s), \hat{u}(s))$, $t \le s \le T$, is the optimal solution, then for any piecewise continuous $v(s)$, $f(\delta; v(\cdot))$ has a local minimum at $\delta = 0$. Then

$$f'(\delta; v(\cdot)) = 0, \qquad f''(\delta; v(\cdot)) \ge 0.$$

This proves the following lemma:

Lemma 6.

If $(\hat{x}(s), \hat{u}(s))$, $t \le s \le T$, is the optimal solution, then

$$f'(\delta; v(\cdot)) = 0, \qquad f''(\delta; v(\cdot)) \ge 0$$

for any piecewise continuous $v(\cdot)$.

The relation

$$f'(\delta; v(\cdot)) = 0$$

holds for any piecewise continuous $v(\cdot)$ if and only if

(16) $$\varphi(\mathrm{s}) = \mathbf{0} \in R^m$$

almost everywhere on $[t, T]$.

Set

$$P(s) = \left\{(\Psi_k)_x(\hat{x}(T), y)A(T) + \int_s^T (\Phi_k)_x(r, \hat{x}(r), y, \hat{u}(r))A(r)dr\right\} \cdot A^{-1}(s), \quad t \le s \le T.$$

Then since $\frac{d}{ds}A^{-1}(s) = -A^{-1}(s)\cdot a_x(s,\hat{x}(s),y,\hat{u}(s))$,

(17) $$\frac{d}{ds}P(s) = -(\Phi_k)_x(s,\hat{x}(s),y,\hat{u}(s)) - P(s)\cdot a_x(s,\hat{x}(s),y,\hat{u}(s)),\quad t \le s \le T,$$

and

$$P(T) = (\Psi_k)_x(\hat{x}(T),y)$$

Define the function $H^y(s,x,u,p):[0,T]\times R^n \times U \times R^n \to R$ by

$$H^y(s,x,u,p) = \Phi_k(s,x,y,u) + p\cdot a(s,x,y,u).$$

Then (13), (17) and (16) can be rewritten as

(13') $$\frac{d\hat{x}(s)}{ds} = H_p^y(s,\hat{x}(s),\hat{u}(s),P(s)),\qquad t \le s \le T,$$

(17') $$\frac{dP(s)}{ds} = -H_x^y(s,\hat{x}(s),\hat{u}(s),P(s)),\qquad t \le s \le T,$$

(16') $$H_u^y(s,\hat{x}(s),\hat{u}(s),P(s)) = 0,\qquad t \le s \le T.$$

Suppose that

(18) $$\det\left|H_{uu}^y(s,\hat{x}(s),\hat{u}(s),P(s))\right| \ne 0,\qquad t \le s \le T.$$

Then since $\hat{x}(s)$ and $P(s)$ are AC, (16') implies that $\hat{u}(s)$ is continuous. Then (13') and (17') imply that $\hat{x}(s)$ and $P(s)$ are of class C^1. Then (16') implies that $\hat{u}(s)$ is also of class C^1.

Also from (16') we get

(19) $$\frac{d\hat{u}(s)}{ds} = -\left[H_{uu}^y\right]^{-1}\cdot\left[H_{su}^y + H_{xu}^y\cdot a - H_{up}^y\cdot\left((\Phi_k)_x + P(s)\cdot a_x\right)\right],\quad t \le s \le T,$$

where the arguments in the derivatives of H^y are $(s, \hat{x}(s), \hat{u}(s), P(s))$, and the arguments in a, Φ_k and their derivatives are $(s, \hat{x}(s), y, \hat{u}(s))$.

From (16') we also get

$$P(s) = (\Phi_k)_u \cdot a_u^T \cdot \left(a_u \cdot a_u^T\right)^{-1} = (\Phi_k)_u \cdot a_u^- , \qquad t \le s \le T, \tag{20}$$

where a_u^- is the generalized inverse of $n \times m$ matrix a_u, and the arguments in a, Φ_k and their derivatives are $(s, \hat{x}(s), y, \hat{u}(s))$.

Then (19) and (20) imply that $\hat{u}(s)$ satisfies the differential equation

$$\frac{d\hat{u}(s)}{ds} = f^y(s, \hat{x}(s), \hat{u}(s)), \qquad t \le s \le T,$$

where the function $f^y(s, x, u) : [t, T] \times R^n \times U \to R^m$ is continuous and is entirely determined by functions a, Φ_k and their derivatives (excluding a_{xx} and $(\Phi_k)_{xx}$).

Denote by $x(s; t, x, y, v)$ and $u(s; t, x, y, v)$ the solution of the following system of differential equations

$$\frac{dx(s)}{ds} = a(s, x(s), y, u(s)), \qquad t \le s \le T,$$

$$\frac{du(s)}{ds} = f^y(s, x(s), u(s)), \qquad t \le s \le T,$$

$$x(t) = x,$$

$$u(t) = v \in U.$$

Then we get the following lemma:

Lemma 7.

If $(\hat{x}(s), \hat{u}(s))$, $t \le s \le T$, is the optimal solution and (18) holds, then

a) $\hat{u}(s)$ is of class C^1,

b) $\hat{x}(s) = x(s;t,x,y,\hat{u}(t))$,

$$\hat{u}(s) = u(s;t,x,y,\hat{u}(t)), \qquad t \le s \le T.$$

CONCLUSION

While dynamic programming is a simple mathematical technique that has been used for many years by mathematicians and engineers in a variety of contexts, it is a powerful systematic tool for optimization problems. The partial differential equation of dynamic programming provides us with the means for finding both the value function and the optimal control, along with the necessary and sufficient conditions for optimality.

The method of dynamic programming encounters the difficulty that for many problems the value function is not differentiable everywhere. In this case, the (Hamilton-Jacobi-)Bellman equation cannot be solved in the classical sense, and the value function is its generalized viscosity solution.

The partial differential equation of dynamic programming is generally difficult to solve explicitly, the value function can be found this way only in a few special cases. In many other cases, numerical methods are needed to solve the (Hamilton-Jacobi-)Bellman equation approximately, that creates another set of difficulties especially for multidimensional control problems.

The sequence of functions W_k constructed in this work can be used to approximate the value function V for an optimal control problem without having to solve the (Hamilton-Jacobi-)Bellman equation. The sequence W_k converges to V uniformly from above with an exponential convergence rate. Once the value function is known or approximated with the desired accuracy, we can use the ideas of the dynamic programming or the Corollary to Lemma 4 to search for an optimal or an ε–optimal control depending on the techniques used to assist the dynamic programming approach and whether an optimal control exists or not.

In the near future, we plan to investigate under which conditions

a) the sequence of control functions used to obtain $W_k(t, x, y)$ converges, and if it can be used to obtain (or approximate) the optimal control $u^*(\cdot)$ for the main problem.

b) the sequence of functions $V_k(t, x, y)$ converges to $V(t, x, y)$;

c) functions $V_k(t, x, y)$ are differentiable in x;

d) the sequence of control functions $u^*_{k,0}(s)$ converges, and if it can be used to obtain (or approximate) the optimal control $u^*(\cdot)$ for the main problem.

Allow us to suggest a couple of potential further problems. Suppose the Markov disturbance $y(s)$ in (1) is not a jump Markov process. If $y(s)$ can be approximated by a jump Markov process $\tilde{y}(s)$, would an optimal control $\tilde{u}^*(s)$ for the control problem with the Markov disturbance $\tilde{y}(s)$, if it exist, perform well for the original control problem with $y(s)$?

For another possible future problem, let $\{y_n, n \geq 1\}$ be an ergodic stochastic process in Y with ergodic distribution ρ for y_n. For $\varepsilon < \varepsilon_0$, consider the process $y_\varepsilon(s)$, $0 \leq s \leq T$, defined by

$$y_\varepsilon(s) = y_n \quad \text{for} \quad (n-1)\,\varepsilon \leq s < n\,\varepsilon.$$

Consider a stochastic optimal control problem where the evolution of the system is described by the differential equation

$$\frac{dx_\varepsilon(s)}{ds} = a(s, x_\varepsilon(s), y_\varepsilon(s), u_\varepsilon(s)), \qquad 0 \le t \le s \le T,$$

with initial condition $x_\varepsilon(t) = x$, and the performance criterion is given by

$$J_\varepsilon(t, x, y; u_\varepsilon(\cdot)) = E_{t,x,y}\left\{\int_t^T \Phi(s, x_\varepsilon(s), y_\varepsilon(s), u_\varepsilon(s))ds + \Psi(x_\varepsilon(T))\right\}.$$

Let

$$\bar{a}(s, x, u) = \int_Y a(s, x, y, u)\rho(dy), \qquad \bar{\Phi}(s, x, u) = \int_Y \Phi(s, x, y, u)\rho(dy),$$

and consider a deterministic optimal control problem where the evolution of the system is described by the differential equation

$$\frac{d\bar{x}(s)}{ds} = \bar{a}(s, \bar{x}(s), \bar{u}(s)), \qquad 0 \le t \le s \le T,$$

with initial condition $\bar{x}(t) = x$, and the performance criterion is given by

$$J(t, x; \bar{u}(\cdot)) = \int_t^T \Phi(s, \bar{x}(s), \bar{u}(s))ds + \Psi(\bar{x}(T)).$$

Suppose an optimal control $\bar{u}^*(s)$ exists. Would $\bar{u}^*(s)$ perform well for the original stochastic control problem for small ε? If $u_\varepsilon^*(s)$ is an optimal control for the original control problem for $\varepsilon < \varepsilon_0$, can anything be said about the behavior of $u_\varepsilon^*(s)$ and their relationship to $\bar{u}^*(s)$ as $\varepsilon \to 0$?

BIBLIOGRAPHY

[1] R. Bellman, *Dynamic Programming*, Princeton Univ. Press, New Jersey (1957)

[2] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., **277** (1984) pp. 1--42.

[3] M. H. A. Davis, *Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models*, J. Roy. Statist. Soc., **B46** (1984), pp. 353--388.

[4] M. H. A. Davis, *Control of piecewise-deterministic processes via discrete-time dynamic programming*, Stochastic differential systems (Bad Honnef, 1985), pp. 140--150, Lecture Notes in Control and Inform. Sci., **78**, Springer, Berlin-New York, 1986.

[5] M. A. H. Dempster and J. J. Ye, *Necessary and sufficient optimality conditions for control of piecewise deterministic Markov processes*, Stochastics Stochastics Rep., **40** (1992), no. 3-4, pp. 125--145.

[6] M. A. H. Dempster and J. J. Ye, *Generalized Bellman-Hamilton-Jacobi optimality conditions for a control problem with a boundary condition*, Appl. Math. Optim., **33** (1996), no. 3, pp. 211--225.

[7] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Applications of Mathematics, **1**, Springer-Verlag, New York, 1975

[8] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Applications of Mathematics, **25**, Springer-Verlag, New York, 1993.

[9] J. J. Florentin, *Optimal control of systems with generalized Poisson inputs*, J. Basic Engineering, **85** (1963), pp. 217--221.

[10] R. M. Goor, *Existence of an optimal control for systems with jump Markov disturbances*, SIAM J. Control Optimization, **14** (1976), no. 5, pp. 899--918.

[11] N. N. Krassovskii and E. A. Lidskii, *Analytical design of controllers in stochastic systems with velocity limited controlling action*, Appl. Math. Mech., **25** (1961), pp. 627--643.

[12] N. V. Krylov, *Introduction to the theory of diffusion processes*, Amer. Math. Soc., Providence, R.I. (1995)

[13] H. J. Kushner, *On the stochastic maximum principle: Fixed time of control*, J. Math. Anal. Appl., **11** (1965), pp. 78--92.

[14] E. A. Lidskii, *Optimal control of systems with random properties*, Appl. Math. Mech., **27** (1963), pp. 42--59.

[15] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, John Wiley, New York (1962)

[16] R. W. Rishel, *Dynamic programming and minimum principles for systems with jump Markov disturbances*, SIAM J. Control, **13** (1975), pp. 338--371.

[17] H. M. Soner, *Optimal control with state-space constraint. II*, SIAM J. Control Optim., **24** (1986), no. 6, pp. 1110--1122.

[18] D. D. Sworder, *Feedback control of a class of linear systems with jump parameters*, IEEE Trans. Automatic Control, **AC-14** (1969), pp. 9--14.

[19] D. Vermes, *Optimal control of piecewise deterministic Markov process*, Stochastics, **14** (1985), no. 3, pp. 165--207.

[20] W. M. Wonham, *Random differential equations in control theory*, Probabilistic Methods in Applied Mathematics, vol. 2, A. T. Bharucha-Reid, ed., Academic Press, New York, 1970, pp. 191--199.