Controlled Markov Processes and Viscosity Solutions

Wendell H. Fleming, H. Mete Soner

Controlled Markov Processes and Viscosity Solutions

Second Edition

Springer

Wendell H. Fleming
Div. Applied Mathematics
Brown University
182 George Street
Providence, RI 02912
USA
whf@cfm.brown.edu

H.M. Soner
Department of Mathematics
Carnegie-Mellon University
Schenley Park
Pittsburgh PA 15213
USA

Mathematics Subject Classification 60JXX, 93EXX, 35K55

Library of Congress Control Number: 2005929857

ISBN-10: 0-387-26045-5
ISBN-13: 978-0387-260457

Printed on acid-free paper.

Printed in the United States of America. EB

9 8 7 6 5 4 3 2 1

springeronline.com

We dedicate this edition to
Florence Fleming
Serpil Soner

Contents

Preface to Second Edition

This edition differs from the previous one in several respects. The use of stochastic calculus and control methods to analyze financial market models has expanded at a remarkable rate. A new Chapter X gives an introduction to the role of stochastic optimal control in portfolio optimization and in pricing derivatives in incomplete markets. Risk-sensitive stochastic control has been another active research area since the First Edition of this book appeared. Chapter VI of the First Edition has been completely rewritten, to emphasize the relationships between logarithmic transformations and risk sensitivity. Risk-sensitive control theory provides a link between stochastic control and H-infinity control theory. In the H-infinity approach, disturbances in a control system are modelled deterministically, instead of in terms of stochastic processes. A new Chapter XI gives a concise introduction to two-controller, zero-sum differential games. Included are differential games which arise in nonlinear H-infinity control and as totally risk-averse limits in risk-sensitive stochastic control. Other changes from the First Edition include an updated treatment in Chapter V of viscosity solutions for second-order PDEs. Material has also been added in Section I.11 on existence of optimal controls in deterministic problems. This simplifies the presentation in later sections, and also is of independent interest.

We wish to thank D. Hernandez-Hernandez, W.M. McEneaney and S.-J. Sheu who read various new chapters of this edition and made helpful comments. We are also indebted to Madeline Brewster and Winnie Isom for their able, patient help in typing and revising the text for this edition.

W.H. Fleming
May 1, 2005 *H.M. Soner*

Preface

This book is intended as an introduction to optimal stochastic control for continuous time Markov processes and to the theory of viscosity solutions. We approach stochastic control problems by the method of dynamic programming. The fundamental equation of dynamic programming is a nonlinear evolution equation for the value function. For controlled Markov diffusion processes on n - dimensional euclidean space, the dynamic programming equation becomes a nonlinear partial differential equation of second order, called a Hamilton – Jacobi – Bellman (HJB) partial differential equation. The theory of viscosity solutions, first introduced by M. G. Crandall and P.-L. Lions, provides a convenient framework in which to study HJB equations. Typically, the value function is not smooth enough to satisfy the HJB equation in a classical sense. However, under quite general assumptions the value function is the unique viscosity solution of the HJB equation with appropriate boundary conditions. In addition, the viscosity solution framework is well suited to proving continuous dependence of solutions on problem data.

The book begins with an introduction to dynamic programming for deterministic optimal control problems in Chapter I, and to the corresponding theory of viscosity solutions in Chapter II. A rather elementary introduction to dynamic programming for controlled Markov processes is provided in Chapter III. This is followed by the more technical Chapters IV and V, which are concerned with controlled Markov diffusions and viscosity solutions of HJB equations. We have tried, through illustrative examples in early chapters and the selection of material in Chapters VI – VII, to connect stochastic control theory with other mathematical areas (e.g. large deviations theory) and with applications to engineering, physics, management, and finance. Chapter VIII is an introduction to singular stochastic control. Dynamic programming leads in that case not to a single partial differential equation, but rather to a system of partial differential inequalities. This is also a feature of other important classes of stochastic control problems not treated in this book, such as impulsive control and problems with costs for switching controls.

Value functions can be found explicitly by solving the HJB equation only in a few cases, including the linear–quadratic regulator problem, and some special problems in finance theory. Otherwise, numerical methods for solving the HJB equation approximately are needed. This is the topic of Chapter IX.

Chapters III, IV and VI rely on probabilistic methods. The only results about partial differential equations used in these chapters concern classical solutions (not viscosity solutions.) These chapters can be read independently of Chapters II and V. On the other hand, readers wishing an introduction to viscosity solutions with little interest in control may wish to focus on Chapter II, Secs. 4–6, 8 and on Chapter V, Secs. 4–8.

We wish to thank M. Day, G. Kossioris, M. Katsoulakis, W. McEneaney, S. Shreve, P. E. Souganidis, Q. Zhang and H. Zhu who read various chapters and made helpful comments. Thanks are also due to Janice D'Amico who typed drafts of several chapters. We are especially indebted to Christy Newton. She not only typed several chapters, but patiently helped us through many revisions to prepare the final version.

W.H. Fleming
June 1, 1992 *H.M. Soner*

Notation

In this book the following system of numbering definitions, theorems, formulas etc. is used. Roman numerals are used to refer to chapters. For example, Theorem II.5.1 refers to Theorem 5.1 in Chapter II. Similarly, IV(3.7) refers to formula (3.7) of Chapter IV; and within Chapter IV we write simply (3.7) for such a reference.

$\mathbb{R}^n$ denotes n-dimensional euclidean space, with elements $x = (x_1, \cdots, x_n)$. We write

$$x \cdot y = \sum_{i=1}^{n} x_i y_i$$

and $|x| = (x \cdot x)^{\frac{1}{2}}$ for the euclidean norm. If A is a $m \times n$ matrix, we denote by $|A|$ the operator norm of the corresponding linear transformation from $\mathbb{R}^n$ into $\mathbb{R}^d$:

$$|A| = \max_{|x| \le 1} |Ax|.$$

The transpose of A is denoted by A'. If a and A are $n \times n$ matrices,

$$\operatorname{tr} aA = \sum_{i,j=1}^{n} a_{ij} A_{ij}.$$

$\mathcal{S}^n$ denotes the set of symmetric $n \times n$ matrices and $\mathcal{S}_+^n$ the set of nonnegative definite $A \in \mathcal{S}^n$. The interior, closure, and boundary of a set B are denoted by $intB$, $\bar{B}$ and ∂B respectively. If Σ is a metric space,

$\mathcal{B}(\Sigma) = \sigma -$ algebra of Borel sets of Σ

$\mathcal{M}(\Sigma) =$ {all real – valued functions on Σ which are bounded below}

$C(\Sigma) =$ {all real – valued continuous functions on Σ}

$C_b(\Sigma) =$ bounded functions in $C(\Sigma)\}$.

If Σ is a Banach space

$C_p(\Sigma) =$ {polynomial growing functions in $C(\Sigma)$}.

A function ϕ is called polynomially growing if there exist constants $K, m \geq 0$ such that

$$|\phi(x)| \leq K(1 + |x|^m), \ \forall x \in \Sigma.$$

For an open set $O \subset I\!\!R^n$, and a positive integer k,

$C^k(O) =$ {all $k-$ times continuously differentiable functions on O}

$C_b^k(O) = \{\phi \in C^k(O):$ ϕ and all partial derivatives of ϕ or orders $\leq k$ are bounded}

$C_p^k(O) = \{\phi \in C^k(O):$ all partial derivatives of ϕ of orders $\leq k$ are polynomially growing}.

For a measurable set $E \subset I\!\!R^n$, we say that $\phi \in C^k(E)$ if there exist $\tilde{E}$ open with $E \subset \tilde{E}$ and $\tilde{\phi} \in C^k(\tilde{E})$ such that $\phi(x) = \tilde{\phi}(x)$ for all $x \in E$. Spaces $C_b^k(E), C_p^k(E)$ are defined similarly. $C^\infty(E), C_b^\infty(E), C_p^\infty(E)$ denote the intersections over $k = 1, 2, \cdots$ of $C^k(E), C_b^k(E), C_p^k(E)$.

We denote the gradient vector and matrix of second order partial derivatives of ϕ by

$$D\phi = (\phi_{x_1}, \cdots, \phi_{x_n})$$

$$D^2\phi = (\phi_{x_i x_j}), i, j = 1, \cdots, n.$$

Sometimes these are denoted instead by ϕ_x, ϕ_{xx} respectively.

If ϕ is a vector-valued function, with values in $I\!\!R^m$, then we write $\phi \in C^k(E), \phi \in C_b^k(E)$ etc if each component of ϕ belongs to $C^k(E), C_b^k(E)$ etc. For vector-valued functions, $D\phi$ and $D^2\phi$ are identified with the differentials of ϕ of first and second orders. For vector-valued ϕ, $|D\phi|, |D^2\phi|$ are the operator norms. We denote intervals of $I\!\!R^1$, respectively closed and half-open to the right, by

$$[a, b], \quad [a, b).$$

Given $t_0 < t_1$

$$Q_0 = [t_0, t_1) \times I\!\!R^n, \qquad \overline{Q}_0 = [t_0, t_1) \times I\!\!R^n.$$

Given $O \subset I\!R^n$ open

$$Q = [t_0, t_1) \times O, \qquad \overline{Q} = [t_0, t_1] \times \overline{O}$$

$$\partial^* Q = ([t_0, t_1] \times \partial O) \cup (\{t_1\} \times O).$$

We call $\partial^* Q$ the *parabolic boundary* of the cylindrical region Q. If $\Phi = \phi(t, x), G \subset I\!R^{n+1}$, we say that $\Phi \in C^{\ell,k}(G)$ if there exist $\tilde{G}$ open with $G \subset \tilde{G}$ and $\tilde{\Phi}$ such that $\tilde{\Phi}(t, x) = \Phi(t, x)$ for all $(t, x) \in G$ and all partial derivatives of $\tilde{\Phi}$ or orders $\leq \ell$ in t and of orders $\leq k$ in x are continuous on $\tilde{G}$. For example, we often consider $\Phi \in C^{1,2}(G)$, where either $G = Q$ or $G = \overline{Q}$. The spaces $C_b^{\ell,k}(G), C_p^{\ell,k}(G)$ are defined similarly as above.

The gradient vector and matrix of second-order partial derivatives of $\Phi(t, \cdot)$ are denoted by $D_x\Phi, D_x^2\Phi$, or sometimes by Φ_x, Φ_{xx}.

If F is a real-valued function on a set U which has a minimum on U, then

$$\arg\min_{v \in U} F(v) = \{v^* \in U : F(v^*) \leq F(v) \ \forall v \in U\}.$$

The supnorm of a bounded function is denoted by $\| \ \|$, and L_p-norms are denoted by $\| \ \|_p$.

I

Deterministic Optimal Control

I.1 Introduction

The concept of control can be described as the process of influencing the behavior of a dynamical system to achieve a desired goal. If the goal is to optimize some payoff function (or cost function) which depends on the control inputs to the system, then the problem is one of *optimal control.*

In this introductory chapter we are concerned with deterministic optimal control models in which the dynamics of the system being controlled are governed by a set of ordinary differential equations. In these models the system operates for times s in some interval I. The state at time $s \in I$ is a vector in n-dimensional euclidean $I\!R^n$. At each time s, a control $u(s)$ is chosen from some given set U (called the *control space.*) If I is a finite interval, namely,

$$I = [t, t_1] = \{s : t \leq s \leq t_1\},$$

then the differential equations describing the time evolution of $x(s)$ are (3.2) below. The cost functional to be optimized takes the form (3.4).

During the 1950's and 1960's aerospace engineering applications greatly stimulated the development of deterministic optimal control theory. Among such applications was the problem of optimal flight trajectories for aircraft and space vehicles. However, deterministic control theory provides methods of much wider applicability to problems from diverse areas of engineering, economics and management science. Some illustrative examples are given in Section 2.

It often happens that a system is being controlled only for $x(s) \in \overline{O}$, where $\overline{O}$ is the closure of some given open set $O \subset I\!R^n$. Two versions of that situation are formulated in Section 3. In one version, control occurs only until the time of exit from a closed cylindrical region $\overline{Q} = [t_0, t_1] \times \overline{O}$. In the other version, only controls which keep $x(s) \in \overline{O}$ for $t \leq s \leq t_1$ are allowed (this is called a *state constrained control problem.*)

The method of dynamic programming is the one which will be followed in this book, to study both deterministic and stochastic optimal control problems. In dynamic programming, a *value function* V is introduced which is the optimum value of the payoff considered as a function of initial data. See Section 4, and also Section 7 for infinite time horizon problems. The value function V for a deterministic optimal control problem satisfies, at least formally, a first order nonlinear partial differential equation. See (5.3) or (7.10) below. In fact, the value function V often does not have the smoothness properties needed to interpret it as a solution to the dynamic programming partial differential equation in the usual ("classical") sense. However, in such cases V can be interpreted as a viscosity solution, as will be explained in Chapter II.

Closely related to dynamic programming is the idea of feedback controls, which will also be called in this book *Markov control policies.* According to a Markov control policy, the control $u(s)$ is chosen based on knowing not only time s but also the state $x(s)$. The Verification Theorems 5.1, 5.2 and 7.1 provide a way to find optimal Markov control policies, in cases when the value function V is indeed a classical solution of the dynamic programming partial differential equation with the appropriate boundary data.

Another approach to optimal deterministic control is via Pontryagin's principle, which provides a general set of necessary conditions for an extremum. In Section 6 we develop, rather briefly, the connection between dynamic programming and Pontryagin's principle. We also give a proof of Pontryagin's principle, for the special case of control on a fixed time interval ($O = \mathbb{R}^n$).

In Section 8 and 9 we consider a special class of control problems, in which the control is the time derivative of the state ($u(s) = \dot{x}(s)$) and there are no control constraints. Such problems belong to the classical calculus of variations. For a calculus of variations problem, the dynamic programming equation is called a Hamilton-Jacobi partial differential equation. Many first-order nonlinear partial differential equations can be interpreted as Hamilton-Jacobi equations, by using duality for convex functions. This duality corresponds to the dual Lagrangian and Hamiltonian formulations in classical mechanics. These matters are treated in Section 10.

Another part of optimal control theory concerns the existence of optimal controls. In Section 11 we prove two special existence theorems which are used elsewhere in this book. The proofs rely on lower semicontinuity of the cost function in the control problem.

The reader should refer to Section 3 for notations and assumptions used in this chapter, for finite-time horizon deterministic optimal control problems. For infinite-time horizon problems, these are summarized in Section 7.

I.2 Examples

We start our discussion by giving some examples. In choosing examples, in this section and later in the book, we have included several highly simplified models chosen from such diverse applications as inventory theory, control of physical devices, financial economics and classical mechanics.

Example 2.1. Consider the production planning of a factory producing n commodities. Let $x_i(s),\ u_i(s)$ denote respectively the inventory level and production rate for commodity $i = 1, \cdots, n$ at time s. In this simple model we assume that the demand rates d_i are fixed constants, known to the planner. Let

$$x(s) = (x_1(s), \cdots, x_n(s)),\ u(s) = (u_1(s), \cdots, u_n(s)),\ d = (d_1, \cdots, d_n).$$

They are, respectively, the inventory and control vectors at time s, and the demand vector. The rate of change of the inventory $x(s) \in I\!R^n$ is

$$\frac{d}{ds}x(s) = u(s) - d. \tag{2.1}$$

Let us consider the production planning problem on a given finite time interval $t \leq s \leq t_1$. Given an initial inventory $x(t) = x$, the problem is to choose the production rate $u(s)$ to minimize

$$\int_t^{t_1} h(x(s))ds + \psi(x(t_1)). \tag{2.2}$$

We call t_1 the *terminal time*, h the *running cost*, and ψ the *terminal cost*. It is often assumed that h and ψ are convex functions, and that $h(x),\ \psi(x)$ have a unique minimum at $x = 0$. A typical example of h is

$$h(x) = \sum_{i=1}^{n} \left[\alpha_i(x_i)^+ + \gamma_i(x_i)^-\right],$$

where α_i, γ_i are positive constants interpreted respectively as a unit holding cost and a unit shortage cost. Here, $a^+ = \max\{a, 0\},\ a^- = \max(-a, 0)$.

The production rate $u(s)$ must satisfy certain constraints related to the physical capabilities of the factory and the workforce. These capacity constraints translate into upper bounds for the production rates. We assume that these take the form $c_1u_1 + \cdots + c_nu_n \leq 1$ for suitable constants $c_i > 0$.

To summarize, this simple production planning problem is to minimize (2.2) subject to (2.1), the initial condition $x(t) = x$, and the control constraint $u(s) \in U$ where

$$U = \{v \in I\!R^n :\ v_i \geq 0, i = 1 \cdots, n, \sum_{i=1}^{n} c_iv_i \leq 1\}. \tag{2.3}$$

An infinite time horizon, discounted cost version of this problem will be mentioned in Example 7.4, and the solution to it will be outlined there.

Example 2.2. Consider a simple harmonic oscillator, in which a forcing term $u(s)$ is taken as the control. Let $x_1(s), x_2(s)$ denote respectively the position and velocity at time s. Then

$$\begin{aligned} &\frac{d}{ds}x_1(s) = x_2(s) \\ (2.4) \qquad & \\ &\frac{d}{ds}x_2(s) = -x_1(s) + u(s). \end{aligned}$$

We require that $u(s) \in U$, where U is a closed interval. For instance, if $U = [-a, a]$ with $a < \infty$, then the bound $|u(s)| \le a$ is imposed on the forcing term.

Let us consider the problem of controlling the simple harmonic oscillator on a finite time interval $t \le s \le t_1$. An initial position and velocity $(x_1(t), x_2(t)) = (x_1, x_2)$ are given. We seek to minimize a quadratic criterion of the form

$$(2.5) \qquad \int_t^{t_1} \left[m_1 x_1(s)^2 + m_2 x_2(s)^2 + u(s)^2\right] ds + d_1 x_1(t_1)^2 + d_2 x_2(t_2)^2,$$

where m_1, m_2, d_1, d_2 are nonnegative constants. If there is no constraint on the forcing term ($U = I\!R^1$), this is a particular case of the linear quadratic regulator problem considered in Example 2.3. If $U = [-a, a]$ with $a < \infty$, it is an example of a linear regulator problem with a saturation constraint.

One can also consider the problem of controlling the solution $x(s) = (x_1(s), x_2(s))$ to (2.4) on an infinite time horizon, say on the time interval $[0, \infty)$. A suitable modification of the quadratic criterion (2.5) could be used as the quantity to be minimized. Another possible criterion to be minimized is the time for $x(s)$ to reach a given target. If the target is the point $(0, 0)$, then the control function $u(\cdot)$ is to be chosen such that the first time θ when $x(\theta) = (0, 0)$ is minimized.

Example 2.3. We will now describe the *linear quadratic regulator problem* (LQRP). Due to the simplicity of its solution, it has been applied to a large number of engineering problems. Let $x(s) \in R^n$, $u(s) \in R^m$ satisfy

$$(2.6) \qquad \frac{d}{ds}x(s) = A(s)x(s) + B(s)u(s)$$

with given matrices $A(s)$ and $B(s)$ of dimensions $n \times n$, $n \times m$ respectively. Suppose we are also given $M(s), N(s)$, and D, such that $M(s)$ and D are nonnegative definite, symmetric $n \times n$ matrices and $N(s)$ is a symmetric, positive definite $m \times m$ matrix. The LQRP is to choose $u(s)$ so that

$$(2.7) \qquad \int_t^{t_1} [x(s) \cdot M(s)x(s) + u(s) \cdot N(s)u(s)]\, ds + x(t_1) \cdot Dx(t_1)$$

is minimized. Here $x \cdot y$ denotes the inner product between two vectors. The solution to this problem will be discussed in Example 5.1.

Example 2.4. The simplest kind of problem in classical calculus of variations is to determine a function $x(\cdot)$ which minimizes a functional

$$(2.8) \qquad \int_t^{t_1} L(s, x(s), \dot{x}(s))ds + \psi(x(t_1)),$$

subject to given conditions on $x(t)$ and $x(t_1)$). Here, $\dot{} = d/ds$. Let us fix the left endpoint, by requiring $x(t) = x$ where $x \in I\!R^n$ is given. For the right endpoint, let us fix t_1 and require that $x(t_1) \in \mathcal{M}$, where $\mathcal{M}$ is given closed subset of $I\!R^n$. If $\mathcal{M} = \{x_1\}$ consists of a single point, then the right endpoint (t_1, x_1) is fixed. At the opposite extreme, there is no restriction on $x(t_1)$ if $\mathcal{M} = I\!R^n$.

We will discuss calculus of variations problems in some detail in Sections 8 − 10. In the formulation in Section 8, we allow the possibility that the fixed upper limit t_1 in (2.8) is replaced by a time τ which is the smaller of t_1 and the exit time of $x(s)$ from a given closed region $\bar{O} \subset I\!R^n$. This is a particular case of the class of control problems to be formulated in Section 3.

I.3 Finite time horizon problems

In this section we formulate some classes of deterministic optimal control problems, which will be studied in the rest of this chapter and in Chapter II. At the end of the section, each of these classes of problems appears as a particular case of a general formulation.

A terminal time t_1 will be fixed throughout. Let $t_0 < t_1$ and consider initial times t in the finite interval $[t_0, t_1)$. (One could equally well take $-\infty < t < t_1$, but then certain assumptions in the problem formulation become slightly more complicated.) The objective is to minimize some payoff functional J, which depends on states $x(s)$ and controls $u(s)$ for $t \le s \le t_1$.

Let us first formulate the state dynamics for the control problem. Let $Q_0 = [t_0, t_1) \times I\!R^n$ and $\overline{Q}_0 = [t_0, t_1] \times I\!R^n$, the closure of Q_0. Let U be a closed subset of m-dimensional $I\!R^m$. We call U the *control space*. The state dynamics are given by a function

$$f : \overline{Q}_0 \times U \to I\!R^m.$$

It is assumed that $f \in C(\overline{Q}_0 \times U)$. Moreover, for suitable K_ρ:

$$(3.1) \qquad |f(t, x, v) - f(t, y, v)| \le K_\rho |x - y|$$

for all $t \in [t_0, t_1],\ x, y \in I\!R^n$ and $v \in U$ such that $|v| \le \rho$. If the control space U is compact, we can replace K_ρ by a constant K, since $U \subset \{v : \ |v| \le \rho\}$ for large enough ρ. If $f(t, \cdot, v)$ has a continuous gradient f_x, (3.1) is equivalent to the condition $|f_x(t, x, v)| \le K_\rho$ whenever $|v| \le \rho$.

A *control* is a bounded, Lebesgue measurable function $u(\cdot)$ on $[t, t_1]$ with values in U. Assumption (3.1) implies that, given any control $u(\cdot)$, the differential equation

$$(3.2) \qquad \frac{d}{ds} x(s) = f(s, x(s), u(s)),\ t \le s \le t_1$$

with initial condition

$$x(t) = x \tag{3.3}$$

has a unique solution. The solution $x(s)$ of (3.2) and (3.3) is called the *state* of the system at time s. Clearly the state depends on the control $u(\cdot)$ and the initial condition, but this dependence is suppressed in our notation.

Let $\mathcal{U}^0(t)$ denote the set of all controls $u(\cdot)$. In notation which we shall use later (Section 9)

$$\mathcal{U}^0(t) = L^\infty([t, t_1];\ U).$$

This is the space of all bounded, Lebesgue measurable, U - valued functions on $[t, t_0]$. In order to complete the formulation of an optimal control problem, we must specify for each initial data (t, x) a set $\mathcal{U}(t, x) \subset \mathcal{U}^0(t)$ of admissible controls and a payoff functional $J(t, x; u)$ to be minimized. Let us first formulate some particular classes of problems (A through D below). Then we subsume all of these classes in a more general formulation. For classes A and B, all controls $u(\cdot) \in \mathcal{U}^0(t)$ are admitted. However, for classes C and D only controls $u(\cdot)$ in a smaller $\mathcal{U}(t, x)$ are admitted.

A. Fixed finite time horizon. The problem is to find $u(\cdot) \in \mathcal{U}^0(t)$ which minimizes

$$J(t, x; u) = \int_t^{t_1} L(s, x(s), u(s))ds + \psi(x(t_1)), \tag{3.4}$$

where $L \in C(\overline{Q}_0 \times U)$. We call L the *running cost* function and ψ the *terminal cost* function.

B. Control until exit from a closed cylindrical region $\overline{Q}$. Consider the following payoff functional J, which depends on states $x(s)$ and controls $u(s)$ for times $s \in [t, \tau)$, where τ is the smaller of t_1 and the exit time of $x(s)$ from the closure $\overline{O}$ of an open set $O \subset I\!R^n$. We let $Q = [t_0, t_1) \times O$, $\overline{Q} = [t_0, t_1] \times \overline{O}$ the closure of the cylindrical region Q, and

$$\partial^* Q = ([t_0, t_1) \times \partial O) \cup (\{t_1\} \times \overline{O}).$$

We call $[t_0, t_1) \times \partial O$ and $\{t_1\} \times \overline{O}$ the *lateral boundary* and *terminal boundary,* respectively, of Q. Given initial data $(t, x) \in \overline{Q}$, let τ denote the exit time of $(s, x(s))$ from $\overline{Q}$. Thus,

$$\tau = \begin{cases} \inf\{s \in [t, t_1) :\ x(s) \notin \overline{O}\} \text{ or} \\ \\ t_1 \text{ if } x(s) \in \overline{O} \text{ for all } s \in [t, t_1) \end{cases}$$

Note that $(\tau, x(\tau)) \in \partial^* Q$. We let

$$J(t, x; u) = \int_t^{\tau} L(s, x(s), u(s))ds \tag{3.5}$$

$$+g(\tau, x(\tau))\chi_{\tau<t_1} + \psi(x(t_1))\chi_{\tau=t_1}$$

Here χ denotes an indicator function. Thus, for real numbers a, b,

$$\chi_{a<b} = \begin{cases} 1 \text{ if } & a < b \\ \\ 0 \text{ if } & a \geq b, \end{cases}$$

and $\chi_{a \leq b}$ is defined similarly. The function g is called a *boundary cost* function, and is assumed continuous.

B′. Control until exit from Q. Let $(t, x) \in Q$, and let τ' be the first time s such that $(s, x(s)) \in \partial^* Q$. Thus, τ' is the exit time of $(s, x(s))$ from Q, rather that from $\overline{Q}$ as for class B above. In (3.5) we now replace τ by τ'. We will give conditions under which B and B' are equivalent optimal control problems.

C. Final endpoint constraint. Suppose that in case A, the additional restriction $x(t_1) \in \mathcal{M}$ is imposed, where $\mathcal{M}$ is a given closed subset of $I\!R^n$. In particular, if $\mathcal{M} = \{x_1\}$ consists of a single point, then both endpoints (t, x) and (t_1, x_1) of the curve $\gamma = \{(s, x(s)) : \ t \leq s \leq t_1\}$ are given. We now admit controls $u(\cdot) \in \mathcal{U}(t, x)$, where

$$\mathcal{U}(t, x) = \{u(\cdot) \in \mathcal{U}^0(t) : x(t_1) \in \mathcal{M}\}.$$

The condition that $\mathcal{U}(t, x)$ is nonempty is called a reachability condition. See Sontag [Sg]. If $U = I\!R^m$, it is related to the concept of controllability.

In a similar way, one can consider the problem of minimizing J in (3.5) subject to an endpoint constraint $(\tau, x(\tau)) \in \mathcal{S}$, where $\mathcal{S}$ is a given closed subset of $\partial^* Q$.

D. State constraint. This is the problem of minimizing $J(t, x; u)$ in (3.4) subject to the constraint $x(s) \in \overline{O}$. In this case,

$$\mathcal{U}(t, x) = \{u(\cdot) \in \mathcal{U}^0(t) : \ x(s) \in \overline{O} \text{ for } t \leq s \leq t_1\}.$$

General problem formulation. Let us now formulate a general class of control problems, which includes each of the classes A through D above. Let $O \subset I\!R^n$ be open, with either: (i) $O = I\!R^n$, or (ii) ∂O a compact manifold of class C^2. Let $Q = [t_0, t_1) \times O$. In case $O = I\!R^n$, we have $Q = Q_0$. Let Ψ be a function, such that

$$\Psi(t, x) = \begin{cases} g(t, x) \text{ if } (t, x) \in [t_0, t_1) \times I\!R^n \\ \\ \psi(x) \ \text{ if } (t, x) \in \{t_1\} \times I\!R^n \end{cases} \tag{3.6}$$

We let

$$J(t, x; \ u) = \int_t^\tau L(s, x(s), u(s))ds + \Psi(\tau, x(\tau)), \tag{3.7}$$

where τ is the exit time of $(s, x(s))$ from $\overline{Q}$. This agrees with (3.5), and also with (3.4) in case $O = \mathbb{R}^n$. We admit controls $u(\cdot) \in \mathcal{U}(t,x)$, where $\mathcal{U}(t,x)$ is nonempty and satisfies the following "switching" condition (3.9). Roughly speaking, condition (3.9) states that if we replace an admissible control by another admissible one after a certain time, then the resulting control is still admissible. More precisely, let $u(\cdot) \in \mathcal{U}(t,x)$ and $u'(\cdot) \in \mathcal{U}(r, x(r))$ for some $r \in [t, \tau]$. Define a new control by

$$\tilde{u}(s) = \begin{cases} u(s), \ t \le s \le r \\ \\ u'(s), r < s \le t_1. \end{cases} \tag{3.8}$$

Let $\tilde{x}(s)$ be the solution to (3.2) corresponding to control $\tilde{u}(\cdot)$ and initial condition $\tilde{x}(t) = x$. Then we assume that

$$\tilde{u}_s(\cdot) \in \mathcal{U}(s, \tilde{x}(s)), \ t \le s \le \tilde{\tau}, \tag{3.9}$$

where $\tilde{u}_s(\cdot)$ denotes the restriction to $[s, t_1]$ of $\tilde{u}(\cdot)$ and $\tilde{\tau}$ is the exit time from $\overline{Q}$ of $(s, \tilde{x}(s))$. Note that (3.9) implies, in particular, that an admissible control always stays admissible. Indeed, simply take in (3.7) $r = \tau$ and $\tilde{u}_s(\cdot) = u_s(\cdot)$.

The control problem is as follows: given initial data $(t,x) \in \overline{Q}$, find $u^*(\cdot) \in \mathcal{U}(t,x)$ such that

$$J(t,x;\ u^*) \le J(t,x;u) \text{ for all } u(\cdot) \in \mathcal{U}(t,x).$$

Such a $u^*(\cdot)$ is called an *optimal* control.

Relation between classes B and B'. Let us conclude this section by giving some conditions (3.10), (3.11) under which the problem of controlling until the time τ of exit of $(s, x(s))$ from $\overline{Q}$ is equivalent of that of controlling until the time τ' of exit from Q. Let us assume:

$$L \ge 0, \psi \ge 0, \ \psi(x) = 0 \text{ for } x \epsilon \partial O \text{ and } g \equiv 0. \tag{3.10}$$

(3.11) For every $(s, \xi) \in [t_0, t_1] \times \partial O$ there exists $v(s,\xi) \in U$ such that

$$f(s, \xi, v(s,\xi)) \cdot \eta(\xi) > 0,$$

where $\eta(\xi)$ is the exterior unit normal at $\xi \in \partial O$.

We always have $\tau' \le \tau \le t_1$. In particular, $\tau' = t_1$ implies that $\tau' = \tau$. If $\tau' < t_1$, then by (3.5) and the assumption $g \equiv 0$,

$$J(t,x;u) = \int_t^{\tau'} L ds + \left[\int_{\tau'}^{\tau} L ds + \psi(x(t_1)) \chi_{\tau = t_1} \right].$$

Let us denote the first term on the right side by $J'(t,x;u)$. J' is the payoff for the problem of control up to time τ', in case $\tau' < t_1$. Since $L \ge 0$ and $\psi \ge 0$,

$$J(t,x;u) \geq J'(t,x;u) \tag{3.12}$$

for all $u(\cdot) \in \mathcal{U}^0(t)$. On the other hand, given $u(\cdot)$ with $\tau' < t_1$, let

$$\tilde{u}(s) = \begin{cases} u(s), \; t \leq s \leq \tau', \\ \\ v(\tau', x(\tau')), \; \tau' < s \leq t_1, \end{cases}$$

with $v(s,\xi)$ as in (3.11). The corresponding solution $\tilde{x}(s)$ of (3.2) with $\tilde{x}(t) = x$ coincides with $x(s)$ for $t \leq s \leq \tau'$, and exits from $\overline{Q}$ at time τ'. Thus,

$$J'(t,x;u) = J(t,x;\tilde{u}). \tag{3.13}$$

From (3.12) and (3.13), it suffices to minimize J among controls $\tilde{u}(\cdot)$ for which the exit times from Q and $\overline{Q}$ are the same.

I.4 Dynamic programming principle

It is convenient to consider a family of optimization problems with different initial conditions (t,x). Consider the minimum value of the payoff function as a function of this initial point. Thus define a *value function* by

$$V(t,x) = \inf_{u(\cdot) \in \mathcal{U}(t,x)} J(t,x;u), \tag{4.1}$$

for all $(t,x) \in \overline{Q}$. We shall assume that $V(t,x) > -\infty$. This is always true if the control space U is compact, or if U is not compact but the cost functions are bounded below ($L \geq -M, \Psi \geq -M$ for some constant $M \geq 0$.)

The method of dynamic programming uses the value function as a tool in the analysis of the optimal control problem. In this section and the following one we study some basic properties of the value function. Then we illustrate the use of these properties in an example for which the problem can be explicitly solved (the linear quadratic regulator problem) and introduce the idea of feedback control policy.

We start with a simple property of V. Let $r \wedge \tau = \min(r,\tau)$. Recall that g is the boundary cost (see (3.5)).

Lemma 4.1. *For every initial condition $(t,x) \in \overline{Q}$, admissible control $u(\cdot) \in \mathcal{U}(t,x)$ and $t \leq r \leq t_1$, we have*

$$\begin{aligned} V(t,x) \leq \int_t^{r \wedge \tau} L(s,x(s),u(s))ds + g(\tau,x(\tau))\chi_{\tau<r} \\ +V(r,x(r))\chi_{r \leq \tau}. \end{aligned} \tag{4.2}$$

Proof. Suppose that $\tau < r \leq t_1$. Then $\Psi(r \wedge \tau, x(r \wedge \tau)) = g(\tau, x(\tau))$, and (4.2) follows from the definition of the value function. Now suppose that

$r \le \tau$. For any $\delta > 0$, choose an admissible control $u^1(\cdot) \in \mathcal{U}(r, x(r))$ such that

$$\int_r^{\tau^1} L(s, x^1(s), u^1(s))ds + \Psi(\tau^1, x^1(\tau^1)) \le V(r, x(r)) + \delta.$$

Here $x^1(s)$ is the state at time s corresponding to the control $u^1(\cdot)$ and initial condition $(r, x(r))$, and τ^1 is the exit time of $(s, x^1(s))$ from $\overline{Q}$. (Such a control $u^1(\cdot)$ is called δ - *optimal*.) As in (3.8) define an admissible control $\tilde{u}(\cdot) \in \mathcal{U}(t, x)$ by

$$\tilde{u}(s) = \begin{cases} u(s), & s \le r \\ \\ u^1(s), & s > r. \end{cases}$$

Let $\tilde{x}(s)$ be the state corresponding to $\tilde{u}(\cdot)$ with initial condition (t, x), and $\tilde{\tau}$ the exit time of $(s, \tilde{x}(s))$ from $\overline{Q}$. Since $r < \tau, \tau^1 = \tilde{\tau}$ and we have

$$\begin{aligned} V(t,x) &\le J(t,x;\tilde{u}) \\ &= \int_t^{\tilde{\tau}} L(s, \tilde{x}(s), \tilde{u}(s))ds + \Psi(\tilde{\tau}, \tilde{x}(\tilde{\tau})) \\ &= \int_t^r L(s, x(s), u(s))ds + \int_r^{\tau^1} L(s, x^1(s), u^1(s))ds \\ &\qquad\qquad + \Psi(\tau^1, x^1(\tau^1)) \\ &\le \int_t^r L(s, x(s), u(s))ds + V(r, x(r)) + \delta. \qquad \square \end{aligned}$$

The proof of Lemma 4.1 shows that the right side of (4.2) is a nondecreasing function of r. However, if $u(\cdot)$ is optimal (or nearly optimal), then this function is constant (or nearly constant). Indeed, for a small positive δ, choose a δ-optimal admissible control $u(\cdot) \in \mathcal{U}(t, x)$. Then for any $r \in [t, t_1]$ we have

$$\begin{aligned} \delta + V(t,x) &\ge J(t,x;u) \\ &= \int_t^{\tau} L(s, x(s), u(s))ds + \Psi(\tau, x(\tau)) \\ &= \int_t^{\tau\wedge r} L(s, x(s), u(s))ds + \int_{\tau\wedge r}^{\tau} L(s, x(s), u(s))ds + \Psi(\tau, x(\tau)) \\ &= \int_t^{\tau\wedge r} L(s, x(s), u(s))ds + J(r\wedge\tau, x(r\wedge\tau); u) \\ &\ge \int_t^{\tau\wedge r} L(s, x(s), u(s))ds + g(\tau, x(\tau))\chi_{\tau<r} + V(r, x(r))\chi_{r\le\tau}. \end{aligned}$$

Since δ is arbitrary, we have proved the following.

Lemma 4.2. *For any initial condition* $(t, x) \in \overline{Q}$ *and* $r \in [t, t_1]$,

$$V(t,x) = \inf_{u(\cdot)\in\mathcal{U}(t,x)} \left[\int_t^{r\wedge\tau} L(s, x(s), u(s))ds + g(\tau, x(\tau))\chi_{\tau<r} + V(r, x(r))\chi_{r\leq\tau} \right]. \tag{4.3}$$

The above identity is called the *dynamic programming principle.* It is the basis of the solution technique developed by Bellman in the 1950's [Be]. An interesting observation is that an optimal control $u^*(\cdot) \in \mathcal{U}(t,x)$ minimizes (4.3) at every r. Hence to determine the optimal control $u^*(t)$, it suffices to analyze (4.3) with r arbitrarily close to t. Intuitively this yields a simple optimization problem that is minimized by $u^*(t)$. However, as we shall see in later chapters, this approach requires a knowledge of the value function.

Another corollary of the above computations is the following.

Corollary 4.1. *An admissible control* $u(\cdot) \in \mathcal{U}(t,x)$ *is* δ*-optimal at* (t, x) *if any only if it is* δ*-optimal at every* $(r, x(r))$ *with* $r \in [t, \tau]$.

I.5 Dynamic programming equation

In this section, we assume that the value function is continuously differentiable and proceed formally to obtain a nonlinear partial differential equation satisfied by the value function. In general however, the value function is not differentiable. In that case a notion of "weak" solutions to this equation is needed. This will be the subject of Chapter 2. After formally deriving the dynamic programming partial differential equation (5.3), we prove two Verification Theorems (Theorems 5.1 and 5.2) which give sufficient conditions for a solution to the optimal control problem.

Let $0 < h \leq t_1 - t$, and take $r = t + h$ in the dynamic programming principle (4.3). Subtract $V(t, x)$ from both sides of (4.3) and then divide by h. This yields

$$\inf_{u(\cdot)\in\mathcal{U}(t,x)} \left\{ \frac{1}{h} \int_t^{(t+h)\wedge\tau} L(s, x(s), u(s))ds + \frac{1}{h} g(\tau, x(\tau))\chi_{\tau<t+h} + \frac{1}{h}[V(t+h, x(t+h))\chi_{t+h\leq\tau} - V(t,x)] \right\} = 0. \tag{5.1}$$

Let us assume that:

(5.2) For every $(t, x) \in Q$ and $v \in U$ there exists $u(\cdot) \in \mathcal{U}(t,x)$ such that

$$v = \lim_{s\downarrow t} u(s).$$

If $\mathcal{U}(t,x) = \mathcal{U}^0(t)$, clearly (5.2) holds. For instance, we may take $u(s) \equiv v$. For the state constrained problem (Case D, Section 3), (5.2) holds provided $\mathcal{U}(r,\xi)$ is not empty for every $(r,\xi) \in Q$ (See Theorem II.12.1.). Note that we assume (5.2) for $(t,x) \in Q$, not $(t,x) \in \overline{Q}$. In the state constrained problem, only controls in some subset of U can be used at time t when (t,x) is on the lateral boundary of Q. If $(t,x) \in Q$, then $x \in O$ and $t + h \leq \tau$ if h is sufficiently small. If we formally let $h \downarrow 0$ in (5.2) we obtain, for $(t,x) \in Q$,

$$\frac{\partial}{\partial t} V(t,x) + \inf_{v \in U} \{L(t,x,v) + f(t,x,v) \cdot D_x V(t,x)\} = 0. \tag{5.3}$$

This is a nonlinear partial differential equation of first order, which we refer to as the *dynamic programming equation* or simply DPE. In (5.3), $D_x V$ denotes the gradient of $V(t,\cdot)$. It is notationally convenient to rewrite (5.3) as

$$-\frac{\partial}{\partial t} V(t,x) + H(t,x,D_x V(t,x)) = 0, \tag{5.3$'$}$$

where for $(t,x,p) \in \overline{Q}_0 \times I\!\!R^n$

$$H(t,x,p) = \sup_{v \in U} \{-p \cdot f(t,x,v) - L(t,x,v)\}. \tag{5.4}$$

In analogy with a quantity occurring in classical mechanics, we call this function the *Hamiltonian*. The dynamic programming equation (5.3$'$) is sometimes also called a *Hamilton–Jacobi–Bellman PDE*.

Equation (5.3) is to be considered in Q, with appropriate terminal or boundary conditions. Let us describe such conditions for problems of the classes A and B in Section 3. Boundary conditions for state constrained problems (class D, Section 3) will be described later in Section II.12. For class A, we have $Q = Q_0$. By (3.4) the terminal (Cauchy) data are

$$V(t_1,x) = \psi(x), \ x \in I\!\!R^n. \tag{5.5}$$

We now state a theorem which connects the dynamic programming equation to the control problem of minimizing (3.4).

Theorem 5.1. $(Q = Q_0)$. *Let* $W \in C^1(\overline{Q}_0)$ *satisfy* (5.3) *and* (5.5). *Then:*

$$W(t,x) \leq V(t,x), \ \forall (t,x) \in \overline{Q}_0. \tag{5.6}$$

Moreover, if there exists $u^*(\cdot) \in \mathcal{U}^0(t)$ *such that*

$$\begin{aligned} L(s,x^*(s),u^*(s)) + f(s,x^*(s),u^*(s)) \cdot D_x W(s,x^*(s)) \\ = -H(s,x^*(s),D_x W(s,x^*(s))) \end{aligned} \tag{5.7}$$

for almost all $s \in [t,t_1]$, *then* $u^*(\cdot)$ *is optimal for initial data* (t,x) *and* $W(t,x) = V(t,x)$.

In Theorem 5.1, $x^*(\cdot)$ denotes the solution to (3.2) with $u(\cdot)=u^*(\cdot)$, $x^*(t)=x$. Theorem 5.1 is called a *Verification Theorem*. Note that, by the definition (5.4) of H, (5.7) is equivalent to

$$(5.7') \qquad u^*(s) \in \arg\min_{v \in U}\{f(s,x^*(s),v) \cdot D_x W(s,x^*(s)) + L(s,x^*(s),v)\}.$$

Proof of Theorem 5.1. Consider any $u(\cdot) \in \mathcal{U}^0(t)$. Using multivariate calculus and the dynamic programming equation (5.3), we obtain

$$(5.8) \quad W(t_1,x(t_1)) = W(t,x) + \int_t^{t_1} [\frac{\partial}{\partial t} W(s,x(s)) + \dot{x}(s) \cdot D_x W(s,x(s)]ds$$

$$= W(t,x) + \int_t^{t_1} [\frac{\partial}{\partial t} W(s,x(s)) + f(s,x(s),u(s)) \cdot D_x W(s,x(s))]ds$$

$$\geq W(t,x) - \int_t^{t_1} L(s,x(s),u(s))ds.$$

By (5.5), $W(t_1,x(t_1)) = \psi(x(t_1))$. Hence

$$W(t,x) \leq J(t,x;u).$$

We get (5.6) by taking the infimum over $u(\cdot)$.

To prove the second assertion of the theorem, let $u^*(\cdot) \in \mathcal{U}^0(t)$ satisfy (5.7). We redo the calculation above with $u^*(\cdot)$. This yields (5.8) with an equality. Hence

$$(5.9) \qquad W(t,x) = J(t,x;u^*).$$

By combining this equality with (5.6), we conclude that $u^*(\cdot)$ is optimal at (t,x).

□

Remark 5.1. Condition (5.7) is necessary as well as sufficient. Indeed, from the proof of Theorem 5.1 and the definition (5.4) of H it is immediate that (5.7) holds for almost all s if $u^*(\cdot)$ is optimal.

We illustrate the use of the Verification Theorem 5.1 in an example.

Example 5.1. Consider the linear quadratic regulator problem (LQRP) described in Example 2.3. In this example, $O = I\!R^n, U = I\!R^m, \mathcal{U}(t,x) = \mathcal{U}^0(t)$, and

$$f(t,x,v) = A(t)x + B(t)v$$

$$(5.10) \qquad L(t,x,v) = x \cdot M(t)x + v \cdot N(t)v$$

$$\Psi(t,x) = \psi(x) = x \cdot Dx.$$

The dynamic programming equation (5.3′) becomes

$$(5.11) \qquad -\frac{\partial}{\partial t}V(t,x) + H(t,x,D_xV(t,x)) = 0,\ t_0 \le t < t_1, x \in I\!R^n.$$

The Hamiltonion $H(t,x,p)$ is given by

$$(5.12) \qquad \begin{aligned} H(t,x,p) &= \sup_v \{-f(t,x,v)\cdot p - L(t,x,v)\} \\ &= \tfrac{1}{4}N^{-1}(t)B'(t)p\cdot B'(t)p - A(t)x\cdot p - x\cdot M(t)x, \end{aligned}$$

where $B'(t)$ is the transpose of the matrix $B(t)$ and $N^{-1}(t)$ is the inverse of $N(t)$ which is assumed to be invertible. For later use, we note that the unique maximizer of (5.12) is

$$(5.13) \qquad v^* = -\frac{1}{2}N^{-1}(t)B'(t)p.$$

To use the Verification Theorem 5.1, first we have to solve (5.11) with the terminal condition

$$(5.14) \qquad V(t_1,x) = x\cdot Dx,\ x \in R^n.$$

We guess that the solution of (5.11) and (5.14) is a quadratic function in x. So, let

$$W(t,x) = x\cdot P(t)x$$

for some symmetric matrix $P(t)$. We substitute $W(t,x)$ into (5.11) to obtain

$$\begin{aligned} &-\frac{\partial}{\partial t}W(t,x) + H(t,x,D_xW(t,x)) \\ &\quad = -x\cdot\frac{\partial}{\partial t}P(t)x + N^{-1}(t)B'(t)P(t)x\cdot B'(t)P(t)x \\ &\qquad -2A(t)x\cdot P(t)x - x\cdot M(t)x \\ &\quad = x\cdot[-\frac{\partial}{\partial t}P(t) + P(t)B(t)N^{-1}(t)B'(t)P(t) \\ &\qquad -A(t)P(t) - P(t)A'(t) - M(t)]x. \end{aligned}$$

Hence W satisfies (5.11) provided that

$$(5.15) \qquad \begin{aligned} \frac{d}{dt}P(t) &= P(t)B(t)N^{-1}(t)B'(t)P(t) \\ &\quad -A(t)P(t) - P(t)A'(t) - M(t),\ t \in [0,t_1). \end{aligned}$$

The continuity of W at time t_1 yields that

$$(5.16) \qquad P(t_1) = D.$$

Equation (5.15) is called a *matrix Riccati equation.* It has been studied extensively. If we fix t_1, then (5.15)-(5.16) has a solution $P(t)$ backward in time on some maximal interval $t_{\min} < t \le t_1$, where either $t_{\min} = -\infty$ or $t_{\min} < \infty$. Let us use Theorem 5.1 to show that $V(t,x) = W(t,x)$ for $t_{\min} < t \le t_1$, and to find an explicit formula for the optimal $u^*(s)$. In view of (5.13), (5.7′) holds at any $s \in [t, t_1]$ if and only if

$$\begin{aligned} u^*(s) &= -\frac{1}{2} N^{-1}(s) B'(s) D_x W(s, x^*(s)) \\ &= -N^{-1}(s) B'(s) P(s) x^*(s). \end{aligned} \tag{5.17}$$

Now substitute (5.17) back into the state equation (2.6) to obtain

$$\frac{d}{ds} x^*(s) = [A(s) - B(s) N^{-1}(s) B'(s) P(s)] x^*(s).$$

This equation has a unique solution satisfying the initial condition $x^*(t) = x$. Thus there is a unique control $u^*(\cdot)$ satisfying (5.17). Theorem 5.1 then implies that $u^*(\cdot)$ is optimal at (t,x).

Notice that the optimal control $u^*(s)$ in (5.17) is a linear function of the state $x^*(s)$. The matrix $N^{-1}(s)B'(s)P(s)$ can be precomputed by solving the Riccati differential equation (5.15) with terminal data (5.16), without reference to the initial conditions for $x(s)$. This is one of the important aspects of the LQRP.

In the LQRP as formulated in Example 2.3, the matrices $M(s)$ and D are nonnegative definite and $N(s)$ is positive definite. This implies that $P(t)$ is nonnegative definite and that $t_{\min} = -\infty$. To see this, for $t_{\min} < t \le t_1$, $0 \le V(t,x) \le J(t,x;0)$. Since $V(t,x) = x \cdot P(t)x$, $P(t)$ is nonnegative definite and bounded on any finite interval, which excludes the possibility that $t_{\min} > -\infty$.

In Section VI.8 we will encounter a class of problems in which $M(s)$ is negative definite. Such problems are called LQRP problems with *indefinite sign.* In this case, $P(t)$ may not be nonnegative definite and $t_{\min}$ may be finite. The following example illustrates these possibilities.

Example 5.2. Let $n = 1, f(v) = v$, $L(x,v) = -x^2 + v^2$ and $D = 0$. The solution to (5.15)-(5.16) is $P(t) = -\tan(t_1 - t)$ if $t_1 - t < \frac{\pi}{2}$ and $t_{\min} = t_1 - \frac{\pi}{2}$.

Control until exit from $\overline{Q}$. Let us next consider the problem of control until the time τ of exit from a closed cylindrical region $\overline{Q}$ (class B, Section 3.) We first formulate appropriate boundary conditions for the dynamic programming equation (5.3). Then we outline a proof of a Verification Theorem (Theorem 5.2) similar to Theorem 5.1. When $t = t_1$, we have as in (5.5):

$$V(t_1, x) = \psi(x), \quad x \in \overline{O}. \tag{5.18}$$

Let us assume that (3.11) holds on the lateral boundary $[t_0, t_1) \times \partial O$. This implies that, for $(t,x) \in [t_0, t_1) \times \partial O$, one choice is to exit immediately from $\overline{Q}$ (thus, $\tau = t$). Therefore,

$$V(t,x) \leq g(t,x),\ (t,x) \in [t_0,t_1) \times \partial O \tag{5.19}$$

If it is optimal to exit immediately from $\overline{Q}$, then equality holds in (5.19). However, in many examples, there are points (t,x) of the lateral boundary for which there exists a control $u^0(\cdot)$ such that $J(t,x;u(\cdot)) < g(t,x)$. See Example II.2.3. At such points, strict inequality holds in (5.19). If (3.10) is assumed, in addition to (3.11), then $V(t,x) \geq 0$. Since $g \equiv 0$ when (3.10) holds, (5.19) implies that the lateral boundary condition $V(t,x) = 0$ holds for all $(t,x) \in [t_0,t_1) \times \partial O$, if both (3.10) and (3.11) hold. Note that we have not yet proved that the value function V is continuous on $\overline{Q}$. However, such a result will be proved later (Theorem II.10.2.). Boundary conditions are discussed further in Section II.13.

Theorem 5.2. *Let $W \in C^1(\overline{Q})$ satisfy* (5.3), (5.18) *and* (5.19). *Then*

$$W(t,x) \leq V(t,x) \textit{ for all } (t,x) \in \overline{Q}. \tag{5.20}$$

Moreover, suppose that there exists $u^*(\cdot) \in \mathcal{U}^0(t)$ such that (5.7) *holds for almost all $s \in [t,\tau^*]$ and $W(\tau^*,x^*(\tau^*)) = g(\tau^*,x^*(\tau^*))$ in case $\tau^* < t_1$. Then $u^*(\cdot)$ is optimal for initial data (t,x) and $W(t,x) = V(t,x)$.*

Here τ^* is the exit time of $(s,x^*(s))$ from $\overline{Q}$. The proof of Theorem 5.2 is almost the same as for Theorem 5.1. In (5.8) the integral is now from t to the exit time τ, and $W(\tau,x(\tau))$ is on the left side. By (5.18) and (5.19), $W(\tau,x(\tau)) \leq \Psi(\tau,x(\tau))$ with Ψ as in (3.6). This gives (5.20). The second half goes exactly as for Theorem 5.1.

Remark 5.2. An entirely similar Verification Theorem is true for the problem of control until the time τ' of exit from Q (rather from $\overline{Q}$.) In fact, since $(s,x(s)) \in Q$ for $t \leq s < \tau'$, the proof of Theorem 5.2 shows that it suffices in this case to assume $W \in C^1(Q) \cap C(\overline{Q})$ rather than $W \in C^1(\overline{Q})$. A situation where such a weaker assumption on W is convenient will arise in Example 7.3.

In Example 5.1 we constructed an admissible control by using the value function. To generalize the procedure, let W be as in Theorem 5.2 (or as in Theorem 5.1 in case $Q = Q_0$.) For $(t,x) \in \overline{Q}$ define a set-valued map $F^*(t,x)$ by

$$F^*(t,x) = \{f(t,x,v) : v \in v^*(t,x)\}$$

where $v^*(t,x)$ is another set-valued map

$$v^*(t,x) = \arg\min_{v \in U} \left[f(t,x,v) \cdot D_x W(t,x) + L(t,x,v)\right]. \tag{5.21}$$

We may now restate (5.7′) as $u^*(s) \in v^*(s,x^*(s))$. Substituting this into the state dynamics yields

$$\dot{x}^*(s) \in F^*(s,x^*(s)),\ s \in [t,\tau^*]. \tag{5.22}$$

Thus, we have the following corollary to Theorem 5.2.