

Applications of
Mathematics
*Stochastic Modelling
and Applied Probability*

24

Edited by

A.V. Balakrishnan
I. Karatzas
M. Yor

Advisory Board

P. Brémaud
E. Carlen
R. Dobrushin
W. Fleming
D. Geman
G. Grimmett
G. Papanicolaou
J. Scheinkman

Applications of Mathematics

- 1 Fleming/Rishel, **Deterministic and Stochastic Optimal Control** (1975)
- 2 Marchuk, **Methods of Numerical Mathematics**, Second Ed. (1982)
- 3 Balakrishnan, **Applied Functional Analysis**, Second Ed. (1981)
- 4 Borovkov, **Stochastic Processes in Queueing Theory** (1976)
- 5 Lipster/Shiryayev, **Statistics of Random Processes I: General Theory** (1977)
- 6 Lipster/Shiryayev, **Statistics of Random Processes II: Applications** (1978)
- 7 Vorob'ev, **Game Theory: Lectures for Economists and Systems Scientists** (1977)
- 8 Shiryaev, **Optimal Stopping Rules** (1978)
- 9 Ibragimov/Rozanov, **Gaussian Random Processes** (1978)
- 10 Wonham, **Linear Multivariable Control: A Geometric Approach**, Third Ed. (1985)
- 11 Hida, **Brownian Motion** (1980)
- 12 Hestenes, **Conjugate Direction Methods in Optimization** (1980)
- 13 Kallianpur, **Stochastic Filtering Theory** (1980)
- 14 Krylov, **Controlled Diffusion Processes** (1980)
- 15 Prabhu, **Stochastic Storage Processes: Queues, Insurance Risk, and Dams** (1980)
- 16 Ibragimov/Has'minskii, **Statistical Estimation: Asymptotic Theory** (1981)
- 17 Cesari, **Optimization: Theory and Applications** (1982)
- 18 Elliott, **Stochastic Calculus and Applications** (1982)
- 19 Marchuk/Shaidourov, **Difference Methods and Their Extrapolations** (1983)
- 20 Hijab, **Stabilization of Control Systems** (1986)
- 21 Protter, **Stochastic Intergration and Differential Equations** (1990)
- 22 Benveniste/Métivier/Priouret, **Adaptive Algorithms and Stochastic Approximations** (1990)
- 23 Kloeden/Platen, **Numerical Solution of Stochastic Differential Equations** (1992)
- 24 Kushner/Dupuis, **Numerical Methods for Stochastic Control Problems in Continuous Time** (1992)

Harold J. Kushner Paul G. Dupuis

Numerical Methods for Stochastic Control Problems in Continuous Time

With 40 Illustrations in 43 parts



Springer-Verlag
New York Berlin Heidelberg London Paris
Tokyo Hong Kong Barcelona Budapest

Harold J. Kushner
Division of Applied Mathematics
Brown University
Providence, RI 02912 USA

Paul G. Dupuis
Division of Applied Mathematics
Brown University
Providence, RI 02912 USA

Managing Editors

A.V. Balakrishnan
Electrical Engineering Department
University of California
Los Angeles, CA 90024 USA

I. Karatzas
Department of Statistics
Columbia University
New York, NY 10027 USA

M. Yor
Laboratoire de Probabilités
Université Pierre et Marie Curie
4 Place Jussieu, Tour 56
75230 Paris Cedex
France

Mathematics Subject Classifications (1991): 65U05, 60F17, 93E20

Library of Congress Cataloging-in-Publication Data
Kushner, Harold J. (Harold Joseph), 1933-

Numerical methods for stochastic control problems in continuous
time / Harold J. Kushner, Paul Dupuis.

p. cm. -- (Applications of mathematics, 24)

Includes bibliographical references and index.

ISBN-13: 978-1-4684-0443-2

1. Stochastic control theory. 2. Markov processes. 3. Numerical
analysis. I. Dupuis, Paul. II. Title. III. Series.

QA402.37.K87 1992

003'.76--dc20

92-10350

Printed on acid-free paper.

© 1992 Springer-Verlag New York, Inc.

Softcover reprint of the hardcover 1st edition 1992

All rights reserved. This work may not be translated in whole or in part without the written
permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York,
NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis.
Use in connection with any form of information storage and retrieval, electronic adaptation,
computer software, or by similar or dissimilar methodology now known or hereafter
developed in forbidden.

The use of general descriptive names, trade names, trademarks, etc., in this publication,
even if the former are not especially identified, is not to be taken as a sign that such names,
as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used
freely by anyone.

Production managed by Karen Phillips; manufacturing supervised by Robert Paella.
Photocomposed pages prepared from the author's TeX file.

9 8 7 6 5 4 3 2 1

ISBN-13: 978-1-4684-0443-2 e-ISBN-13: 978-1-4684-0441-8
DOI: 10.1007/978-1-4684-0441-8

To
Linda, Diana, and Nina
and
Suzanne, Alexander, and Nicole

Contents

Introduction	1
1 Review of Continuous Time Models	7
1.1 Martingales and Martingale Inequalities	8
1.2 Stochastic Integration	9
1.3 Stochastic Differential Equations: Diffusions	14
1.4 Reflected Diffusions	21
1.5 Processes with Jumps	27
2 Controlled Markov Chains	35
2.1 Recursive Equations for the Cost	36
2.2 Optimal Stopping Problems	42
2.3 Discounted Cost	47
2.4 Control to a Target Set and Contraction Mappings	49
2.5 Finite Time Control Problems	51
3 Dynamic Programming Equations	53
3.1 Functionals of Uncontrolled Processes	54
3.2 The Optimal Stopping Problem	59
3.3 Control Until a Target Set Is Reached	61
3.4 A Discounted Problem with a Target Set and Reflection	64
3.5 Average Cost Per Unit Time	65
4 The Markov Chain Approximation Method: Introduction	67
4.1 The Markov Chain Approximation Method	69
4.2 Continuous Time Interpolation and Approximating Cost Function	71
4.3 A Continuous Time Markov Chain Interpolation	74
4.4 A Random Walk Approximation to the Wiener Process	78
4.5 A Deterministic Discounted Problem	80
4.6 Deterministic Relaxed Controls	85
5 Construction of the Approximating Markov Chain	89
5.1 Finite Difference Type Approximations: One Dimensional Examples	90
5.2 Numerical Simplifications and Alternatives for Example 4	98

5.3	The General Finite Difference Method	106
5.4	A Direct Construction of the Approximating Markov Chain	112
5.5	Variable Grids	122
5.6	Jump Diffusion Processes	127
5.7	Approximations for Reflecting Boundaries	133
5.8	Dynamic Programming Equations	142
6	Computational Methods for Controlled Markov Chains	151
6.1	The Problem Formulation	152
6.2	Classical Iterative Methods: Approximation in Policy and Value Space	154
6.3	Error Bounds for Discounted Problems	162
6.4	Accelerated Jacobi and Gauss-Seidel Methods	164
6.5	Domain Decomposition and Implementation on Parallel Processors	169
6.6	A State Aggregation Method	172
6.7	Coarse Grid-Fine Grid Solutions	177
6.8	A Multigrid Method	179
6.9	Linear Programming Formulations and Constraints	186
7	The Ergodic Cost Problem: Formulations and Algorithms	193
7.1	The Control Problem for the Markov Chain: Formulation .	194
7.2	A Jacobi Type Iteration	197
7.3	Approximation in Policy Space	199
7.4	Numerical Methods for the Solution of (3.4)	201
7.5	The Control Problem for the Approximating Markov Chain	203
7.6	The Continuous Parameter Markov Chain Interpolation .	207
7.7	Computations for the Approximating Markov Chain	209
7.8	Boundary Costs and Controls	214
8	Heavy Traffic and Singular Control Problems: Examples and Markov Chain Approximations	217
8.1	Motivating Examples	218
8.2	The Heavy Traffic Problem: A Markov Chain Approximation	236
8.3	Singular Control: A Markov Chain Approximation	242
9	Weak Convergence and the Characterization of Processes	247
9.1	Weak Convergence	248
9.2	Criteria for Tightness in $D^k[0, \infty)$	252
9.3	Characterization of Processes	253
9.4	An Example	254
9.5	Relaxed Controls	264

10 Convergence Proofs	269
10.1 Limit Theorems and Approximations of Relaxed Controls	270
10.2 Existence of an Optimal Control: Absorbing Boundary	278
10.3 Approximating the Optimal Control	284
10.4 The Approximating Markov Chain: Weak Convergence	288
10.5 Convergence of the Costs: Discounted Cost and Absorbing Boundary	294
10.6 The Optimal Stopping Problem	298
11 Convergence for Reflecting Boundaries, Singular Control and Ergodic Cost Problems	303
11.1 The Reflecting Boundary Problem	303
11.2 The Singular Control Problem	315
11.3 The Ergodic Cost Problem	320
12 Finite Time Problems and Nonlinear Filtering	325
12.1 The Explicit Approximation Method: An Example	326
12.2 The General Explicit Approximation Method	329
12.3 The Implicit Approximation Method: An Example	330
12.4 The General Implicit Approximation Method	332
12.5 The Optimal Control Problem: Approximations and Dynamic Programming Equations	335
12.6 Methods of Solution, Decomposition and Convergence	337
12.7 Nonlinear Filtering	339
13 Problems from the Calculus of Variations	347
13.1 Problems of Interest	348
13.2 Numerical Schemes and Convergence for the Finite Time Problem	351
13.3 Problems with a Controlled Stopping Time	365
13.4 Problems with a Discontinuous Running Cost	395
14 The Viscosity Solution Approach to Proving Convergence of Numerical Schemes	411
14.1 Definitions and Some Properties of Viscosity Solutions	412
14.2 Numerical Schemes	417
14.3 Proof of Convergence	420
References	423
List of Symbols	433
Index	435

Introduction

This book is concerned with numerical methods for stochastic control and optimal stochastic control problems. The random process models of the controlled or uncontrolled stochastic systems are either diffusions or jump diffusions. Stochastic control is a very active area of research and new problem formulations and sometimes surprising applications appear regularly. We have chosen forms of the models which cover the great bulk of the formulations of the continuous time stochastic control problems which have appeared to date. The standard formats are covered, but much emphasis is given to the newer and less well known formulations. The controlled process might be either stopped or absorbed on leaving a constraint set or upon first hitting a target set, or it might be reflected or “projected” from the boundary of a constraining set. In some of the more recent applications of the reflecting boundary problem, for example the so-called heavy traffic approximation problems, the directions of reflection are actually discontinuous. In general, the control might be representable as a bounded function or it might be of the so-called impulsive or singular control types. Both the “drift” and the “variance” might be controlled. The cost functions might be any of the standard types: Discounted, stopped on first exit from a set, finite time, optimal stopping, average cost per unit time over the infinite time interval, and so forth. There might be separate costs when the process is on the boundary and when it is in the interior of the set of interest. In fact all of the standard cost functionals can be dealt with by the methods to be presented. There is a close connection between approximation methods for stochastic control and those for optimal nonlinear filtering, and approximation methods for the latter problem are also discussed.

The class of methods to be dealt with is referred to generically as the *Markov chain approximation method*. It is a powerful and widely usable set of ideas for numerical and other approximation problems for either controlled or uncontrolled stochastic processes, and it will be shown that it has important applications to deterministic problems as well. The initial development of the approximation method and the convergence proofs appeared in the first author’s 1977 book (which was subsequently translated into Russian). Since that time new classes of problems have arisen to which the original proofs could not be applied directly, the techniques of approximation and mathematical proof have been considerably streamlined, and also extended to cover a large part of the new problems of interest in contin-

uous time stochastic control. In addition, many new techniques for actually doing the computations have been developed.

The basic idea of the Markov chain approximation method is to approximate the original controlled process by an appropriate controlled Markov chain on a finite state space. One also needs to approximate the original cost function by one which is appropriate for the approximating chain. These approximations should be chosen such that a good numerical approximation to the associated control or optimal control problem can be obtained with a reasonable amount of computation. The criterion which must be satisfied by the process approximation is quite mild. It is essentially what we will call "local consistency." Loosely speaking, this means that from a local point of view, the conditional mean and covariance of the changes in state of the chain are proportional to the local mean drift and covariance for the original process. Such approximations are readily obtained by a variety of methods and are discussed extensively in Chapters 4 and 5. The numerical problem is then to solve the problem for the approximating controlled chain. Methods for doing this are covered in detail in Chapters 6 to 8, with the basic concepts being in Chapter 6. One needs to prove that the solutions to the problems with the approximating chain actually converge to the correct value as some approximation parameter goes to zero. One of the great advantages of the approach is that this can often be done by probabilistic methods which do not require the use of any of the analytical properties of the actual solution. This is particularly important since for many classes of problems, little is known about the analytical properties of the Bellman equations.

The book is written on two levels, so that the methods of actual approximation and practical use of the algorithms can be read without any involvement with the mathematics of the convergence proofs. An effort is made to motivate the development of the algorithms in terms of the properties of the original process of interest, but this is on a purely intuitive level, and the various properties which are used should be intuitively natural. Thus the book should be accessible to a reader with only a formal acquaintance with the properties of diffusion and jump diffusion processes. Indeed, one of the primary purposes of the book is the encouragement of the use and the development of the actual algorithms on a wide variety of practical problems. We note that the methods are not restricted to optimal control problems. They can be used for the calculation of approximations to a large class of functionals of processes of the diffusion or jump diffusion type. The probabilistic nature of the process of approximation and proof allows us to use our physical intuition concerning the original problem in all phases of the development. The reader should keep in mind that we approximate the control problem and not the formal dynamical equation for the optimal value function.

The proofs of convergence for the Markov chain approximation method are purely probabilistic. One finds a suitable continuous time interpolation

of the optimally controlled approximating chain, and shows that in the sense of weak or distributional convergence, there is a convergent subsequence whose limit is an optimally controlled process of the original diffusion or jump diffusion type, and with the original cost function and boundary data. The methods of proof are from the theory of weak convergence of probability measures. The weak convergence method provides a unifying approach for all the problems of interest.

We also note that the deterministic problem is a special case of what is of interest here. All of the algorithms and approximations can be used for deterministic problems where one wants a feedback control. The basic fact which is needed for the proofs is the local consistency referred to above. In fact, in Chapter 13, we see how it can be used for deterministic problems which arise in the calculus of variations and with possibly discontinuous data.

From a formal point of view, as is well known and also discussed in Chapter 3, the optimal value functions for many stochastic control problems with a diffusion process model can be represented as the solution to either a highly nonlinear and possibly degenerate partial differential equation of the elliptic or parabolic type with appropriate boundary data, or possibly as a variational inequality with similar degeneracies and boundary data. If the underlying process is of the controlled jump diffusion type then the partial differential equation will be replaced by a nonlinear partial integrodifferential equation. Because of this formal “PDE-type” structure, the current literature in numerical analysis offers many useful ideas concerning both the approximation of the original problem and the solution of the approximating problem, at least in the simpler cases. This influence is particularly strong at the level of the actual computational algorithms, as might be seen by scanning Chapter 6. But, except for a few special situations, the current literature is not adequate to deal with the convergence proofs, and of course the convergence proofs are the ultimate guide for the determination of the actual approximations which are to be used for the computation. This is particularly true for the cases where the formal dynamical equation for the optimal value function does not fit the classical PDE models with which the numerical analysis literature has concerned itself. These new cases would include problems involving discontinuous boundary data, as occurs in the so-called heavy traffic problem; problems with singular controls, controlled variance, ergodic problems, or generally for degenerate models. By degenerate, we mean that the covariance matrix of the diffusion part of the model is not strictly positive definite, and we note that this is a quite common case in applications. One might sometimes use the formal PDE or variational inequality as a guide to get useful approximations in certain cases, but the ultimate proofs do not use these structures or any of their properties.

Chapter 2 contains an outline of results in the control of Markov chains on a finite state space which will be needed later in the book. The approxi-

mations are developed and the numerical algorithms explained in Chapters 4 to 8 and in the first part of Chapter 12. The mathematical proofs are in Chapters 10-14 with the basic ideas being in Chapter 10. Chapters 1 and 9 contain a survey of the mathematical results which will be needed in Chapters 10-14. Chapter 3 contains a formal discussion of the continuous time control problem. It introduces some of the basic models for the cost function and illustrates the use of the principle of optimality to formally get the nonlinear PDEs which are satisfied by the optimal value function. The chapter is for motivational purposes only, since the results are formal and are not explicitly used except for further motivation. But having the formal PDEs is helpful when discussing the approximations, since it can be used to show us that we are on familiar ground. Chapters 2 and 4 to 8 can be read independently of Chapters 1, 3 and 9-14. In particular, the methods for getting the approximations and numerical algorithms in Chapters 4-6, and in the relevant “algorithmic” parts of Chapters 7, 8, 12 and 13 can be read without familiarity with the more theoretical chapters.

The Markov chain approximation method is introduced in Chapter 4. This is done via some simple problems, but all of the main issues which will need to be considered are introduced. A deterministic example is developed in detail, since that allows us to demonstrate the ideas with the least encumbrance. To show us that the intuition of numerical analysis remains useful, the role of finite difference approximations in getting the approximating chain is emphasized.

Chapter 5 contains an extensive exposition of methods for getting appropriate approximating chains. It will be seen that this is essentially an exercise in common sense, and that we have considerable freedom. The methods which are discussed are illustrative of essentially automatic and easy to use procedures, but they also illustrate the general principles which one ought to follow in the development of extensions. Chapter 6 develops the basic computational methods which are used to solve the approximating optimal control problems. Many of the standard approaches to the solution of large systems of linear equations appear, since the standard approximation in policy space method leads to the solution of a sequence of such linear problems. In Chapter 7, we extend the discussion of Chapter 6 for the average cost per unit time problem. Chapter 8 deals with the so-called heavy traffic and singular control problems. These classes of problems are rather recent arrivals on the stochastic control scene, but they model some very important applications, and the numerical analysis questions are particularly interesting. Due to this, we discuss a number of specific examples. Then the ideas of Chapters 5 and 6 are extended. Chapter 12 is concerned with problems which are of interest over a finite time interval only, as well as with the approximation problem for the optimal nonlinear filter. In the course of development of the techniques for the numerical problem, various side results appear which have a broader use in stochastic process theory; for example, the simple proof of existence of a solution to the stochastic

differential equation with a reflecting boundary in Chapter 11. The survey paper [80] contains some topics which are not covered here; e.g., the controlled variance problem. Chapter 13 considers a variety of deterministic problems that are not special cases of the stochastic problems dealt with previously. In Chapter 14 we briefly discuss the “viscosity solution” approach, which provides an alternative method of proof in some cases.

The advances developed in this book allow us to do effective computations on a wide variety of problems. Nevertheless, the entire field of numerical stochastic control is in its infancy and much more effort is needed on all phases of the area.

It is a pleasure to acknowledge the considerable help of Luiz Felipe Martins, John Oliensis, Felisa Vásquez-Abad and Jichuan Yang for their many helpful suggestions as well as for their preparing and testing of codes. We also thank John Oliensis for the figures in Chapter 13. This work has been supported for many years by the Air Force Office of Scientific Research, the National Science Foundation and the Army Research Office.

1

Review of Continuous Time Models

In this book we will consider methods of numerically computing the value function for certain classes of controlled continuous time stochastic and deterministic processes. The purpose of the present chapter is to provide an introduction to and some of the background material for controlled diffusions and controlled jump diffusions. These types of processes include many of the models that are commonly used. This section is only intended to serve as a review of the main ideas and for purposes of reference. Other models (e.g., singularly controlled diffusions) that are also of interest will be introduced and elaborated on in the appropriate later sections of the book.

Our main interest in the present chapter is in constructing and establishing certain properties of the processes. Chapter 9 will also deal with important background material, such as alternative characterizations of the processes and the theory of weak convergence. Section 1.1 presents the definitions and fundamental inequalities of martingales. In Section 1.2, we review integration with respect to the Wiener process and state the associated chain rule (Itô's formula). With the definition of the stochastic integral and the appropriate martingale estimates in hand, in Sections 1.3 and 1.4 we define what is meant by a solution of a stochastic differential equation and outline the proof of existence of solutions. The processes defined by the solutions of these stochastic differential equations will serve as our models of controlled continuous time processes with continuous sample paths. We also discuss the notion of uniqueness of solutions that will be suitable for our later work. We will first consider processes without control, and then indicate the extension to the controlled case. For purposes of numerical computation, it is usually necessary that the processes be constrained to take values in some bounded set G . Thus, the original problem may have to be modified to reflect this need. There are at least two considerations that must be kept in mind when redefining the problem statement. The most obvious requirement is that quantities one is interested in estimating (e.g., certain expectations) should not be greatly perturbed. The second requirement is that the resulting modification should be convenient with respect to computations. One way to achieve this bounding of the state space is to stop the process the first time it leaves some large but bounded

domain. At this time, a cost should be assigned that is approximately equal to the total of the future costs if the process were not stopped. The determination of an appropriate stopping cost can be difficult in practice. A second method involves constraining the process without actually stopping it. For diffusions, such a bounding of the state space can be achieved by imposing a reflecting boundary condition on the boundary of G . For processes involving jumps, one may simply project the process back into G in some convenient way whenever it leaves that set. Besides being useful for numerical purposes, such constrained or reflected processes are important as basic models for many problems in stochastic systems theory. Examples and references are given Chapter 8. A unified and natural method for constructing and analyzing such constrained processes can be based on use of the *Skorokhod Problem* (SP). In Section 1.4, we define a solution to the Skorokhod Problem in a setting general enough to cover all of the situations involving reflecting diffusions and projected processes that will appear in the book. We consider several illustrative examples, including a case that is convenient for the purpose of simply constraining a process to a bounded set. We conclude the chapter in Section 1.5 with a discussion of the analogous results for jump diffusions.

1.1 Martingales and Martingale Inequalities

Consider a probability space (Ω, \mathcal{F}, P) . A family of σ -algebras $\{\mathcal{F}_t, t \geq 0\}$ is called a *filtration* on this probability space if $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $0 \leq s \leq t$. Let $E_{\mathcal{F}_t}$ and $P_{\mathcal{F}_t}$ denote expectation and probability conditioned on the σ -algebra \mathcal{F}_t , respectively. Suppose $C^k[0, T]$ denotes the space of continuous functions mapping $[0, T]$ into \mathbb{R}^k and that $D^k[0, T]$ consists of those functions from $[0, T]$ into \mathbb{R}^k that are continuous from the right and have limits from the left. Let $C^k[0, \infty)$ and $D^k[0, \infty)$ denote the analogous path spaces for the interval $[0, \infty)$. These spaces may be metrized so they are complete separable metric spaces [12]. We will drop k from the notation when $k = 1$.

Consider a stochastic process $x(s \cdot)$ defined on (Ω, \mathcal{F}, P) and taking values in the path space $D[0, \infty)$. Then $x(\cdot)$ is said to be an \mathcal{F}_t -*martingale* if $x(t)$ is \mathcal{F}_t -measurable and $E|x(t)| < \infty$ for all $t \geq 0$, and if

$$E_{\mathcal{F}_t}x(t+s) = x(t) \text{ w.p.1 for all } t \geq 0 \text{ and all } s \geq 0. \quad (1.1)$$

If the particular filtration is obvious or unimportant, then we will suppress the prefix and refer to $x(\cdot)$ simply as a martingale. We will refer to a vector valued process as a vector valued martingale if each of its components is a martingale with respect to the same filtration.

The importance of martingales is in part due to the bounds and inequalities associated with them. Processes can often be decomposed as a sum of a bounded variation term plus a martingale (with respect to some conve-

nient filtration). When performing calculations involving the process (e.g., obtaining bounds on moments), the bounded variation term is often easy to handle. Thus, the decomposition is useful since estimates such as those presented below can be used to bound the martingale part of the process. The inequalities that we will find most useful in this book are the following. Let $x(\cdot)$ be an \mathcal{F}_t -martingale that has right continuous sample paths. (All of the martingales encountered in this book will be of this type.) Then, for any $c > 0$, $T \geq 0$ and $0 \leq t \leq T$,

$$P_{\mathcal{F}_t} \left\{ \sup_{t \leq s \leq T} |x(s)| \geq c \right\} \leq E_{\mathcal{F}_t} |x(T)|^2 / c^2 \text{ w.p.1}, \quad (1.2)$$

$$E_{\mathcal{F}_t} \sup_{t \leq s \leq T} |x(s)|^2 \leq 4E_{\mathcal{F}_t} |x(T)|^2 \text{ w.p.1}. \quad (1.3)$$

See [69].

A random variable $\tau : \Omega \rightarrow [0, \infty]$ is called an \mathcal{F}_t -stopping time if $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \in [0, \infty)$. If $x(\cdot)$ is an \mathcal{F}_t -martingale and τ is a uniformly bounded \mathcal{F}_t -stopping time, then the stopped process $x(t \wedge \tau)$ is also an \mathcal{F}_t -martingale. Thus, (1.2) and (1.3) also hold if we replace T by $T \wedge \tau$, where τ is any \mathcal{F}_t -stopping time.

If there exists a nondecreasing sequence $\{\tau_n, n = 1, 2, \dots\}$ of \mathcal{F}_t -stopping times such that $\tau_n \rightarrow \infty$ w.p.1 and such that for each n the stopped process $x(t \wedge \tau_n)$ is a martingale, then $x(\cdot)$ is called an \mathcal{F}_t -local martingale.

1.2 Stochastic Integration

We are concerned in this chapter with reviewing the construction and various properties of some of the standard models used for continuous time problems. One of the most important models for stochastic systems is the stochastic differential equation (SDE) of the form

$$x(t) = x(0) + \int_0^t b(x(s))ds + \int_0^t \sigma(x(s))dw(s).$$

Here $x(\cdot)$ is an \mathbb{R}^k -valued process with continuous sample paths, $w(\cdot)$ is an \mathbb{R}^n -valued process which serves as a “driving noise,” and $b(\cdot)$ and $\sigma(\cdot)$ are vector and matrix valued functions of appropriate dimensions. Such equations produce a rich class of models all defined in terms of a relatively simple model for the driving noise. The only quantity needing explanation in the expression for $x(\cdot)$ is the term $\int_0^t \sigma(x(s))dw(s)$, to which we now turn.

We will consider stochastic integrals with respect to two basic processes. The first process is the *Wiener process*. As is well known, the resulting stochastic integral and related theory of stochastic differential equations

(due to K. Itô) provide a very convenient family of models that are Markovian and possess continuous sample paths. In the beginning of the section, we define the Wiener process and recall some basic properties. We then review Itô's definition of integration with respect to the Wiener process and state the chain rule. In order to model processes involving jumps, we will make use of *Poisson random measures* as a driving term. The associated stochastic integral is, in a certain sense, easier to define than for the case of the Wiener process. This integral will be defined and the combined chain rule for both types of driving noise will be given in Section 1.5. If A is a collection of random variables defined on a probability space (Ω, \mathcal{F}, P) , then we use $\mathcal{F}(A)$ to denote the σ -algebra generated by A . If S is a topological space, then $\mathcal{B}(S)$ is used to denote the σ -algebra of Borel subsets of S .

Wiener Process. Let (Ω, \mathcal{F}, P) be a probability space and let $\{\mathcal{F}_t, t \geq 0\}$ be a filtration defined on it. A process $\{w(t), t \geq 0\}$ is called an \mathcal{F}_t -Wiener process if it satisfies the following conditions.

1. $w(0) = 0$ w.p.1.
2. $w(t)$ is \mathcal{F}_t -measurable and $\mathcal{F}(w(s) - w(t) : s \geq t)$ is independent of \mathcal{F}_t for all $t \geq 0$.
3. The increments $w(s) - w(t)$ are normally distributed with mean 0 and variance $\sigma^2 > 0$ for all $s \geq t \geq 0$.
4. The sample paths of $w(\cdot)$ are in $C^k[0, \infty)$.

For several constructions as well as an account of the detailed properties of the Wiener process we refer the reader to the book of Karatzas and Shreve [69]. If $\sigma = 1$, then the process $w(\cdot)$ is called a standard \mathcal{F}_t -Wiener process. If \mathcal{F}_t is simply $\mathcal{F}(w(s) : 0 \leq s \leq t)$, then the \mathcal{F}_t prefix is often suppressed and we refer to $w(\cdot)$ simply as a Wiener process. A finite collection of mutually independent \mathcal{F}_t -Wiener processes is called a vector valued \mathcal{F}_t -Wiener process.

A very important property of any \mathcal{F}_t -Wiener process is that it is also an \mathcal{F}_t -martingale. This property follows from parts (2) and (3) of the definition. Indeed, the Wiener process is a canonical example of a continuous sample path process that is both Markovian and a martingale. The fact that it has continuous sample paths and is also a martingale imply the sample paths of $w(\cdot)$ are of unbounded variation over any nontrivial time interval (w.p.1). This excludes defining $\int \sigma(t)dw(t)$ by any pathwise construction if we wish to allow a large class of integrands. Nonetheless, a useful integral may be defined in a straightforward manner if we properly restrict the class of allowed integrands. We will impose conditions on the integrand that will imply that the resulting integral is an \mathcal{F}_t -martingale when considered as a function of the upper limit of integration. Thus, we will be able to use the martingale estimates in the construction and applications of the integral.

Remark. In this book we will always assume the coefficients in the equations are bounded. This is not much of a restriction for our purposes because the state spaces of the processes will be bounded for numerical purposes. Because of this boundedness our definitions are somewhat simpler than is typical in thorough treatments of the theory of SDE.

Assumptions on the Integrand. A random process $f(\cdot)$ is said to be \mathcal{F}_t -adapted if $f(t)$ is \mathcal{F}_t -measurable for each $t \geq 0$. If $w(\cdot)$ is an \mathcal{F}_t -Wiener process and $f(\cdot)$ is \mathcal{F}_t -adapted, then $f(\cdot)$ is said to be *nonanticipative* with respect to $w(\cdot)$, since $f(u)$ and $w(s) - w(t)$ are independent whenever $0 \leq u \leq t \leq s$. A process $f(t, \omega)$ is called *measurable* if $\{(t, \omega) : f(t, \omega) \in A\}$ belongs to the product σ -algebra $\mathcal{B}([0, \infty)) \times \mathcal{F}$, where $\mathcal{B}([0, \infty))$ is the σ -algebra of Borel subsets of $[0, \infty)$. Let $\Sigma_b(T)$ denote the set of \mathcal{F}_t -adapted, measurable, real valued processes $\sigma(\cdot)$ which are uniformly bounded in $t \in [0, T]$ and $\omega \in \Omega$, and let Σ_b denote those processes defined on $[0, \infty)$ that are in $\Sigma_b(T)$ for each $T < \infty$. We say that a random process is a *simple function* if there exists a sequence of deterministic times $\{t_i, i = 0, 1, \dots\}$ such that $0 = t_0 < t_1 < \dots < t_i \rightarrow \infty$, and such that $\sigma(t) = \sigma(t_i)$ for $t \in [t_i, t_{i+1})$. The set of all simple functions in Σ_b will be denoted by Σ_b^* .

We are now in the position to define the integral with respect to a Wiener process. For full details concerning the arguments used below, the reader may consult the book of Karatzas and Shreve [69]. The integral is defined for an arbitrary integrand in Σ_b via an approximation argument. In general, the stochastic integral defined below will be unique only in the sense that any two versions will have sample paths that agree with probability one. We will follow the usual convention of identifying any process with the class of processes whose sample paths are identical with probability one and, therefore, omit the corresponding qualification in the arguments below.

Definition and Elementary Properties of $\int_0^t \sigma dw$ when $\sigma \in \Sigma_b$. Let $w(\cdot)$ be a standard \mathcal{F}_t -Wiener process and let $\sigma \in \Sigma_b^*$ be given. Because the sample paths of σ are piecewise constant, it is possible to define $\int \sigma dw$ (referred to hereafter as the integral of σ) in a simple way. Let $0 = t_0 < t_1 < \dots < t_i \rightarrow \infty$ be the partition associated to σ . Then we set

$$\int_0^t \sigma(u) dw(u) = \sum_{i=0}^{n-1} \sigma(t_i) [w(t_{i+1}) - w(t_i)] + \sigma(t_n) [w(t) - w(t_n)]$$

for $t \in [t_n, t_{n+1})$. It can be shown [69, Proposition 3.2.6] that for each $\sigma \in \Sigma_b$ there exist $\{\sigma_n, n \in \mathbb{N}\} \subset \Sigma_b^*$ such that for each $T \in [0, \infty)$,

$$\int_0^T E|\sigma_n(u) - \sigma(u)|^2 du \rightarrow 0 \tag{2.1}$$

as $n \rightarrow \infty$. It is then natural to define $\int_0^t \sigma(u)dw(u)$ as the limit (in some suitable sense) of the processes $\int_0^t \sigma_n(u)dw(u)$. Full details of such a program are in [69, Section 3.2]. Important properties of the resulting integral are that for any $0 \leq s < t < \infty$, any σ, σ_1 , and σ_2 in Σ_b^* , we have (w.p.1)

$$E_{\mathcal{F}_s} \int_0^t \sigma(u)dw(u) = \int_0^s \sigma(u)dw(u), \quad (2.2)$$

$$E_{\mathcal{F}_s} \left[\int_s^t \sigma(u)dw(u) \right]^2 = \int_s^t E_{\mathcal{F}_s} [\sigma(u)]^2 du, \quad (2.3)$$

$$\int_0^t \sigma_1(u)dw(u) + \int_0^t \sigma_2(u)dw(u) = \int_0^t [\sigma_1 + \sigma_2](u)dw(u). \quad (2.4)$$

These properties follow easily from the definition of the integral for $\sigma \in \Sigma_b^*$ and are extended to integrands in Σ_b by approximation.

Given an open subset U of some Euclidean space, we let $C^k(U)$ denote the set of all real valued functions on U that have continuous derivatives up to and including order k .

Itô's Formula. Let $f \in C^1(\mathbb{R})$ and let $x(t) = x(0) + \int_0^t b(s)ds$. The change of variable formula for the composed function $f(x(t))$ is of course

$$f(x(t)) - f(x(0)) = \int_0^t f_x(x(s))b(s)ds,$$

which we can write symbolically as

$$df(x(t)) = f_x(x(t))dx(t) = f_x(x(t))b(t)dt.$$

The analogous formula for functions of the form $h(t) = \int_0^t \sigma(s)dw(s)$, as well as variations, plays an equally important role in stochastic analysis. With the definition of the stochastic integral given above, it turns out that the change of variable formula takes a slightly more cumbersome form and is valid under more restrictive conditions than in the classical case. Nonetheless, it is still an extraordinarily powerful tool. Consider the more general form

$$x(t) = x(0) + \int_0^t b(s)ds + \int_0^t \sigma(s)dw(s). \quad (2.5)$$

We also use the notation

$$dx(s) = b(s)ds + \sigma(s)dw(s)$$

to express the relationship (2.5). We will not state Itô's formula under very general conditions, but only in the form needed later in the book. Thus,

we assume $b(\cdot)$ and $\sigma(\cdot)$ are in Σ_b . Then Itô's formula states that for any $f \in C^2(\mathbb{R}^k)$,

$$f(x(t)) - f(x(0)) = \int_0^t f_x(x(s))dx(s) + \frac{1}{2} \int_0^t f_{xx}(x(s))\sigma^2(s)ds,$$

where

$$\int_0^t f_x(x(s))dx(s) = \int_0^t f_x(x(s))b(s)ds + \int_0^t f_x(x(s))\sigma(s)dw(s).$$

We can write this relationship symbolically as

$$df(x(t)) = \left[f_x(x(t))b(t) + \frac{1}{2}f_{xx}(x(t))\sigma^2(t) \right] dt + f_x(x(t))\sigma(t)dw(t).$$

The Vector Cases of Stochastic Integration and Itô's Formula. All vectors are assumed to be column vectors and we use a prime as a superscript to denote transpose. Let $w(\cdot)$ be an n -dimensional vector valued \mathcal{F}_t -Wiener process. Let $\sigma(\cdot)$ be a process taking values in the space of $n \times k$ real matrices with the property that each entry $\sigma_{ij}(\cdot)$ is in Σ_b for $(i, j) \in \{1, \dots, n\} \times \{1, \dots, k\}$. We then define the integral of $\sigma(\cdot)$ with respect to $w(\cdot)$ to be the vector valued random process $h(\cdot) = (h_1(\cdot), \dots, h_k(\cdot))'$ given by

$$h_i(t) = \sum_{j=1}^n \int_0^t \sigma_{ij}(s)dw_j(s)$$

for $i \in \{1, \dots, k\}$. Suppose that $b(\cdot) = (b_1(\cdot), \dots, b_k(\cdot))'$ and that $b_i(\cdot) \in \Sigma_b$ for each $i \in \{1, \dots, k\}$. Define

$$x(t) = x(0) + \int_0^t b(s)ds + \int_0^t \sigma(s)dw(s). \quad (2.6)$$

We will sometimes use the differential notation

$$dx(s) = b(s)ds + \sigma(s)dw(s)$$

in lieu of (2.6). Then the *vector version* of Itô's formula is as follows. For any $f \in C^2(\mathbb{R}^k)$, let $f_x(\cdot)$ and $f_{xx}(\cdot)$ denote the gradient and Hessian matrix of f , respectively. Define $a(s) = \sigma(s)\sigma'(s)$. Then

$$f((x(t)) - f(x(0)) = \int_0^t f'_x(x(s))dx(s) + \frac{1}{2} \int_0^t \sum_{i,j=1}^k \text{tr}[f_{xx}(x(s))a(s)] ds, \quad (2.7)$$

where

$$\int_0^t f'_x(x(s))dx(s) = \int_0^t f'_x(x(s))b(s)ds + \int_0^t f'_x(x(s))\sigma(s)dw(s),$$

and where $\text{tr } B$ denotes the trace of any square matrix B .

1.3 Stochastic Differential Equations: Diffusions

Let $b(\cdot)$ and $\sigma(\cdot)$ be bounded measurable functions mapping \mathbb{R}^k into \mathbb{R}^k and into the space of real $k \times n$ matrices, respectively. We now return to the class of models introduced at the beginning of the previous section. Thus, we consider solutions to

$$x(t) = x(0) + \int_0^t b(x(s))ds + \int_0^t \sigma(x(s))dw(s). \quad (3.1)$$

This class of models is widely used in applications in diverse areas. They are characterized in terms of the *drift vector* $b(\cdot)$ and *diffusion matrix* $a(\cdot) = \sigma(\cdot)\sigma'(\cdot)$. Loosely speaking, such a process is a Markov process $x(\cdot)$ with continuous sample paths and the following “local properties.” Assume (for now) that $b(\cdot)$ and $\sigma(\cdot)$ are continuous, and let $\Delta t > 0$ be small. Then the “local mean drift” and “local covariance” satisfy

$$E [x(t + \Delta t) - x(t)|x(t)] \approx b(x(t))\Delta t, \quad (3.2)$$

$$\text{cov}[x(t + \Delta t) - x(t)|x(t)] \approx a(x(t))\Delta t. \quad (3.3)$$

The classical probabilistic tool used in the construction and analysis of solutions to (3.1) is the theory of stochastic differential equations due to K. Itô. We often write equations such as (3.1) in the symbolic form

$$dx(t) = b(x(t))dt + \sigma(x(t))dw(t). \quad (3.4)$$

Let $w(\cdot)$ be a vector valued \mathcal{F}_t -Wiener process and let $x(0)$ be a given \mathcal{F}_0 -measurable random vector. By a solution to the SDE described by (3.1) we mean a continuous \mathcal{F}_t -adapted process $x(\cdot)$ which satisfies (3.1) with probability one.

There are two important notions of the sense in which a solution to (3.4) can be said to *exist* and also two senses in which *uniqueness* is said to hold. The distinction will turn out to be important for our purposes.

Strong Existence. We say that strong existence holds if given a probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process $w(\cdot)$ and an \mathcal{F}_0 -measurable initial condition $x(0)$, then an \mathcal{F}_t -adapted process $x(\cdot)$ exists satisfying (3.1) for all $t \geq 0$.

Weak Existence. We say that weak existence holds if given any probability measure μ on \mathbb{R}^k there exists a probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process $w(\cdot)$, and an \mathcal{F}_t -adapted process $x(\cdot)$ satisfying (3.1) for all $t \geq 0$ as well as $P\{x(0) \in \Gamma\} = \mu(\Gamma)$.

Strong existence of a solution requires that the probability space, filtration, and Wiener process be given first and that the solution $x(\cdot)$ then be

found for the given data. Weak sense existence allows these objects to be constructed together with the process $x(\cdot)$. Clearly, strong existence implies weak existence.

Strong Uniqueness. Suppose that a fixed probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , and an \mathcal{F}_t -Wiener process $w(\cdot)$ are given. Let $x_i(\cdot), i = 1, 2$, solve (3.1) for the given Wiener process $w(\cdot)$. We say that strong uniqueness holds if

$$P\{x_1(0) = x_2(0)\} = 1 \Rightarrow P\{x_1(t) = x_2(t) \text{ for all } t \geq 0\} = 1.$$

Weak Uniqueness. Suppose we are given weak sense solutions

$$\{(\Omega_i, \mathcal{F}_i, P_i, \mathcal{F}_{i,t}, w_i(\cdot), x_i(\cdot))\}, i = 1, 2,$$

to (3.1). We say that weak uniqueness holds if equality of the distributions induced on \mathbb{R}^k by $x_i(0)$ under $P_i, i = 1, 2$, implies the equality of the distributions induced on $C^k[0, \infty)$ by $x_i(\cdot)$ under $P_i, i = 1, 2$.

Strong uniqueness is also referred to as *pathwise uniqueness*, whereas weak uniqueness is often called *uniqueness in the sense of probability law*. Two relationships between these concepts that are useful are the following: strong uniqueness implies weak uniqueness, and weak existence together with strong uniqueness imply strong existence [69, Chapter 5.3]. For the problems of interest in this book, namely, proving the convergence of numerical schemes, it will be seen that weak solutions that are unique in the weak sense are all that is actually needed.

Itô's Formula and the Differential Operator. Let $x(\cdot)$ be any solution to (3.1), and suppose that the coefficients $b(x(\cdot))$ and $\sigma(x(\cdot))$ satisfy the conditions assumed of $b(\cdot)$ and $\sigma(\cdot)$ in the last section. The link between the process $x(\cdot)$ and certain second order partial differential equations is provided by Itô's formula and the differential operator that appears therein. Let $a(x) = \sigma(x)\sigma'(x)$, and for any $f \in C^2(\mathbb{R}^k)$ let

$$(\mathcal{L}f)(x) = f'_x(x)b(x) + \frac{1}{2}\text{tr}[f_{xx}(x)a(x)]. \quad (3.5)$$

Then Itô's formula (2.5) states that

$$f(x(t)) = f(x(0)) + \int_0^t (\mathcal{L}f)(x(s))ds + \int_0^t f'_x(x(s))\sigma(x(s))dw(s). \quad (3.6)$$

We next discuss some basic results concerning solutions to SDE's. Let $\|\cdot\|$ denote the norm on the space of real $k \times n$ matrices given by $\|\sigma\| = \sum_{i=1}^k \sum_{j=1}^n \sigma_{ij}^2$.

Picard Iteration. The Picard iteration method is a classical approach to the construction of solutions to ordinary differential equations. Itô extended the method to construct and prove strong uniqueness of strong solutions to (3.1). Although the proof appears in many places, we include an outline here. This is done mainly to facilitate the discussion in Section 1.4 and later sections on reflecting diffusions, for which the details are not so easily available. In order to limit the discussion, here and elsewhere in the chapter we consider only the case where the functions $b(\cdot)$ and $\sigma(\cdot)$ are nonrandom and independent of time. The reader is referred to the large literature on stochastic differential equations for generalizations.

A3.1. *There exists $C \in (0, \infty)$ such that*

$$|b(x) - b(y)| \vee \|\sigma(x) - \sigma(y)\| \leq C|x - y|$$

for all $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^k$.

Theorem 3.1. *Assume (A3.1). Then for every deterministic initial condition $x(0)$, the SDE (3.1) has a strong solution that is unique in the strong (and therefore also in the weak) sense.*

The proof turns on the following estimate: let $z_i, i = 1, 2$ be continuous \mathcal{F}_t -adapted processes and define

$$y_i(t) = x(0) + \int_0^t b(z_i(s))ds + \int_0^t \sigma(z_i(s))dw(s), \quad i = 1, 2,$$

$\Delta y(t) = y_1(t) - y_2(t)$, and $\Delta z(t) = z_1(t) - z_2(t)$. Then for each $T \in (0, \infty)$ there exists $L \in (0, \infty)$ such that for all $0 \leq t \leq T$,

$$E \sup_{0 \leq s \leq t} |\Delta y(s)|^2 \leq L \int_0^t E \sup_{0 \leq r \leq s} |\Delta z(r)|^2 ds. \quad (3.7)$$

This follows directly from the Lipschitz continuity properties of b and σ , the martingale property of the stochastic integral, and the estimate (1.3). For more details, the reader may consult [69, Section 5.2]. From this estimate, one may readily prove strong existence and strong uniqueness, as we now show.

To prove uniqueness, let $x_1(\cdot)$ and $x_2(\cdot)$ both be solutions, and let $f(t) = E \sup_{0 \leq s \leq t} |x_1(s) - x_2(s)|^2$. By taking $x_i(\cdot) = y_i(\cdot) = z_i(\cdot), i = 1, 2$, and applying (3.7), we obtain

$$f(t) \leq L \int_0^t f(s)ds.$$

Then Gronwall's inequality [46] implies $f(t) = 0$ for all $t \in [0, T]$. Because T is arbitrary, this proves uniqueness in $C^k[0, \infty)$.

A solution to (3.1) can be constructed by a variation on the classical technique of Picard iteration. A sequence of processes $\{x_n(\cdot)\}$ is defined recursively by $x_0(t) = x(0)$, $t \geq 0$, and

$$x_{n+1}(t) = x(0) + \int_0^t b(x_n(s))ds + \int_0^t \sigma(x_n(s))dw(s).$$

By the way in which the processes were defined, the elements of this sequence are \mathcal{F}_t -adapted processes with continuous sample paths. Applying (3.7) for $n \geq 1$ with $y_1(\cdot) = x_{n+1}(\cdot)$, $y_2(\cdot) = z_1(\cdot) = x_n(\cdot)$, and $z_2(\cdot) = x_{n-1}(\cdot)$, we obtain

$$E \left[\sup_{0 \leq s \leq t} |x_{n+1}(s) - x_n(s)|^2 \right] \leq L \int_0^t E \left[\sup_{0 \leq r \leq s} |x_n(r) - x_{n-1}(r)|^2 \right] ds.$$

Iterating backward to $n = 0$ and evaluating the resulting integral yields

$$E \left[\sup_{0 \leq t \leq T} |x_{n+1}(t) - x_n(t)|^2 \right] \leq \frac{(LT)^n K}{n!},$$

where

$$K = E \left[\sup_{0 \leq t \leq T} \left(\int_0^t b(x(0))ds + \int_0^t \sigma(x(0))dw(s) \right) \right]^2$$

is finite by (1.3) and (2.3). By Chebyshev's inequality,

$$P \left\{ \sup_{0 \leq t \leq T} |x_{n+1}(t) - x_n(t)| \geq 2^{-n} \right\} \leq K \frac{(4LT)^n}{n!}.$$

We may therefore apply the Borel-Cantelli lemma to conclude the event

$$\sup_{0 \leq t \leq T} |x_{n+1}(t) - x_n(t)| \geq 2^{-n}$$

occurs infinitely with probability zero. Therefore, off a set N of zero probability, the sample paths of $x_n(\cdot)$ are a Cauchy sequence in $C^k[0, T]$. Let $x(\cdot, \omega)$ denote the limit of $x_n(\cdot, \omega)$ for $\omega \notin N$. Because T is arbitrary, we can assume that the convergence, in fact, takes place in $C^k[0, \infty)$. Clearly, $x(\cdot)$ is \mathcal{F}_t -adapted. Since the assumed continuity properties of $b(\cdot)$ and $\sigma(\cdot)$, (1.3), and (2.3) imply

$$\int_0^t b(x_n(s))ds + \int_0^t \sigma(x_n(s))dw(s) \rightarrow \int_0^t b(x(s))ds + \int_0^t \sigma(x(s))dw(s)$$

in $C^k[0, \infty)$ (w.p.1), we conclude that $x(\cdot)$ is a solution to the stochastic differential equation (3.4).

Solutions Via a Measure Transformation Method. We will briefly outline a very useful method for obtaining weak solutions. The principle

application is in cases where $b(\cdot)$ has less regularity than is needed for the Picard iteration technique. Let $w(\cdot)$ be an n -dimensional \mathcal{F}_t -Wiener process and let $z(\cdot)$ be an n -dimensional process with each component in Σ_b . Define

$$R(t) = \exp \left(\int_0^t z'(s)dw(s) - \frac{1}{2} \int_0^t |z(s)|^2 ds \right). \quad (3.8)$$

By Itô's formula

$$R(t) = 1 + \int_0^t R(s)z'(s)dw(s).$$

Because $z(\cdot)$ is bounded, $E|R(t)| < \infty$. Thus, the process $R(\cdot)$ is a martingale and therefore $ER(t) = 1$ for all $t \in [0, \infty)$.

Now fix $T \in (0, \infty)$, and define a probability measure \tilde{P}_T on (Ω, \mathcal{F}_T) by

$$\tilde{P}_T(A) = E[I_A R(T)] \text{ for } A \in \mathcal{F}_T. \quad (3.9)$$

The equality $ER(T) = 1$ guarantees that \tilde{P}_T is indeed a probability measure.

Theorem 3.2. (Girsanov) *Assume that $R(\cdot)$ defined by equation (3.8) is a martingale. Then on the interval $[0, T]$ the process*

$$\tilde{w}(t) = w(t) - \int_0^t z(s)ds \quad (3.10)$$

is an \mathcal{F}_T -Wiener process on the probability space $(\Omega, \mathcal{F}_T, \tilde{P}_T)$.

The typical use of Theorem 3.2 is in the following situation. Let $w(\cdot)$ be a k -dimensional \mathcal{F}_t -Wiener process. Let $w_1(\cdot)$ and $w_2(\cdot)$ denote the first n and last $k-n$ components of $w(\cdot)$, respectively. Let $\sigma_1(\cdot)$ be $n \times n$ matrix valued with the property that $\sigma_1^{-1}(x)$ is uniformly bounded in $x \in \mathbb{R}^k$. Consider the stochastic differential equation

$$\begin{aligned} dx_1 &= \sigma_1(x)dw_1 \\ dx_2 &= b_2(x)dt + \sigma_2(x)dw_2, \end{aligned} \quad (3.11)$$

where the dimensions of the $x_i(\cdot), i = 1, 2$, $b_2(\cdot)$, and $\sigma_2(\cdot)$ are all compatible. Assume that the drift vector and diffusion matrix of this equation satisfy the continuity and boundedness conditions assumed for the Picard iteration method, or any other set of conditions which guarantees the existence of a weak sense solution. Let $b_1(\cdot)$ be a bounded Borel measurable function. If we define $z_1(\cdot) = \sigma_1^{-1}(x(\cdot))b_1(x(\cdot))$, and $R(\cdot)$, \tilde{P}_T , and $\tilde{w}_1(\cdot)$ by (3.8), (3.9), and (3.10), respectively, then under \tilde{P}_T , $x(\cdot)$ solves

$$\begin{aligned} dx_1(t) &= b_1(x(t))dt + \sigma_1(x(t))d\tilde{w}_1(t) \\ dx_2(t) &= b_2(x(t))dt + \sigma_2(x(t))dw_2(t). \end{aligned} \quad (3.12)$$

We thereby obtain weak sense existence for such a class of equations with bounded measurable drift terms.

It is also possible to use Girsanov's theorem to prove weak sense uniqueness. For the details under general conditions, the reader can consult [61, 69]. However, the basic idea which is involved can be outlined for the example above. It must first be observed that the distribution of $x(\cdot)$ under P uniquely determines its distribution under \tilde{P}_T . Indeed, the uniform invertibility of $\sigma_1(\cdot)$ implies that the distribution of $x(\cdot)$ determines that of $w_1(\cdot)$, and therefore that of the Radon-Nikodym derivative $R(\cdot)$ as well. Next, consider two weak sense solutions to (3.12). Applying Girsanov's theorem to these two solutions we may *remove* the b_1 part of the drift term and produce two weak sense solutions to (3.11). Suppose the assumptions made on the coefficients appearing in (3.11) imply weak sense uniqueness (e.g., (A3.1)). Then the distribution of the two weak sense solutions to (3.11) must be identical, and by the preceding comments the same is true of the two weak sense solutions of (3.12).

Remark. Clearly, weak sense uniqueness is really a statement concerning the measure induced by the process on $C^k[0, T]$. Suppose that a weak sense uniqueness result holds for every deterministic initial condition. Then weak sense uniqueness must also follow for an arbitrary initial distribution.

Controlled Diffusions. Suppose that $w(\cdot)$ is a Wiener process on some probability space (Ω, \mathcal{F}, P) . Let \mathcal{U} be a compact subset of some Euclidean space, and let $u(\cdot)$ be a \mathcal{U} -valued, measurable process also defined on (Ω, \mathcal{F}, P) . We can topologize the set of such controls in any convenient way, e.g., the topology inherited from the L^1 metric. We say that the control $u(\cdot)$ is *nonanticipative* with respect to the Wiener process $w(\cdot)$ if there exists a filtration \mathcal{F}_t defined on (Ω, \mathcal{F}, P) such that $u(\cdot)$ is \mathcal{F}_t -adapted and $w(\cdot)$ is an \mathcal{F}_t -Wiener process. We say that $u(\cdot)$ is an *admissible control law* with respect to $w(\cdot)$, or that *the pair $(u(\cdot), w(\cdot))$ is admissible*, if $u(\cdot)$ is nonanticipative with respect to $w(\cdot)$.

In this section, we will consider controlled diffusions of the form

$$dx(t) = b(x(t), u(t))dt + \sigma(x(t))dw(t), \quad (3.13)$$

which is the differential representation of

$$x(t) = x(0) + \int_0^t b(x(s), u(s))ds + \int_0^t \sigma(x(s))dw(s).$$

In analogy with the uncontrolled case, we have the following definitions.

Strong Existence. We say that strong existence holds if given a probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process $w(\cdot)$, a control process $u(\cdot)$ that is \mathcal{F}_t -adapted and an \mathcal{F}_0 -measurable initial condition $x(0)$, then an \mathcal{F}_t -adapted process $x(\cdot)$ exists satisfying (3.13) for all $t \geq 0$.

Weak Existence. Suppose we are given probability distributions Λ and μ on the sample path space of the pair $(u(\cdot), w(\cdot))$ and on \mathbb{R}^k , respectively. We say that weak existence holds if there exists a probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process $w(\cdot)$, an \mathcal{F}_t -adapted control process $u(\cdot)$ (i.e., the pair $(u(\cdot), w(\cdot))$ is admissible), and an \mathcal{F}_t -adapted process $x(\cdot)$ such that Λ and μ are the distributions induced by $(u(\cdot), w(\cdot))$ and $x(0)$ under P , and such that $x(\cdot)$ satisfies (3.13) for all $t \geq 0$.

Strong Uniqueness. Suppose that a fixed probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process $w(\cdot)$, and a control process $u(\cdot)$ that is \mathcal{F}_t -adapted are given. Let $x_i(\cdot)$, $i = 1, 2$, solve (3.13) for the given Wiener process and control process. We say that strong uniqueness holds if

$$P\{x_1(0) = x_2(0)\} = 1 \Rightarrow P\{x_1(t) = x_2(t) \text{ for all } t \geq 0\} = 1.$$

Weak Uniqueness. Assume that we are given weak sense solutions

$$\{(\Omega_i, \mathcal{F}_i, P_i), \mathcal{F}_{i,t}, w_i(\cdot), u_i(\cdot), x_i(\cdot)\}, i = 1, 2,$$

to (3.13). We say that weak uniqueness holds if equality of the joint distributions of $(u_i(\cdot), w_i(\cdot), x_i(0))$ under P_i , $i = 1, 2$, implies the equality of the distributions of $(x_i(\cdot), u_i(\cdot), w_i(\cdot), x_i(0))$ under P_i , $i = 1, 2$.

Strong solutions that are unique in the strong sense and weak solutions that are unique in the weak sense may be constructed in exactly the same way as for the case of uncontrolled diffusions. For example, assume the following analogue of (A3.1).

A3.2. *There exists $C \in (0, \infty)$ such that*

$$|b(x, u) - b(y, u)| \vee \|\sigma(x) - \sigma(y)\| \leq C|x - y|$$

for all $x \in \mathbb{R}^k$, $y \in \mathbb{R}^k$, and $u \in \mathcal{U}$. Furthermore, the function $b : \mathbb{R}^k \times \mathcal{U} \rightarrow \mathbb{R}^k$ is measurable, and $\sup_{u \in \mathcal{U}} |b(0, u)| < \infty$.

Under this assumption the Picard iteration and Girsanov transformation methods may be used to get existence and uniqueness results analogous to those without control. The details are the same as those in the case with no control as long as the control process is \mathcal{F}_t -adapted and measurable, with the only changes in the proofs being notational.

For fixed $\alpha \in \mathcal{U}$, we define the differential operator \mathcal{L}^α by

$$(\mathcal{L}^\alpha f)(x) = f'_x(x)b(x, \alpha) + \frac{1}{2}\text{tr}[f_{xx}(x)\sigma(x)].$$

If $u(\cdot)$ is an admissible control process and if $x(\cdot)$ is the associated solution to (3.13), then Itô's formula continues to hold in the form

$$f(x(t)) = f(x(0)) + \int_0^t (\mathcal{L}^{u(s)} f)(x(s))ds + \int_0^t f'_x(x(s))\sigma(x(s))dw(s).$$

1.4 Reflected Diffusions

In this section we will review results concerning diffusion processes that are “instantaneously” reflected back into a closed domain G when the process tries to leave G . One of the reasons for the wide applicability of diffusion processes is that they serve as mathematically tractable approximations to the actual “physical” processes that are encountered in applications. For example, a diffusion process could be used to approximate a process defined by driving an ODE by some “wide band” stationary noise process. Other examples occur when a process that is originally defined in discrete time is suitably interpolated and scaled and then replaced by a diffusion. A large number of examples of both types may be found in [79].

An example of a process that can be well approximated by a diffusion process is that of a queueing process under certain assumptions (the so-called *heavy traffic* assumptions) and after a suitable rescaling of the space and time variables. Here, the non-negativity of the components of the original process imply a constraint on the state space of the approximating diffusion process. It turns out that the proper way to implement this constraint in the diffusion approximation is by adding what is known as a *reflecting boundary*. Reflecting diffusions often give the appropriate diffusion approximations for processes that are constrained in some way to remain in a given set.

In order to motivate the definition given below for a reflecting diffusion, we consider a simple example. Let $\{\xi_i, i < \infty\}$ be a sequence of independent and identically distributed (iid) \mathbb{R}^k -valued random variables. Assume that

$$0 = E\xi_i, \quad \sigma\sigma' = \text{cov } \xi_i,$$

where the prime denotes transpose, cov stands for covariance, and σ is a $k \times k$ matrix. Let $[a]$ denote the integer part of a . It is well known that the process

$$x^n(t) = n^{-1/2} \sum_{i=1}^{[t/n]} \xi_i + x$$

tends weakly (see Section 9.1 for the definition) to the solution of the SDE

$$dx(t) = \sigma dw(t), \quad x(0) = x,$$

as $n \rightarrow \infty$. Suppose that we consider a constrained version of this process. We will use a very simple constraint mechanism. However, the basic ideas carry over to more complicated constraint mechanisms, such as those in queueing systems. Let v be any vector with positive first component v_1 , and let $G = \{x : x_1 \geq 0\}$. For $x \notin G$, define $t_x = \inf\{t > 0 : x + tv \in G\}$ and $\pi(x) = x + t_x v$. For $x \in G$, let $\pi(x) = x$. Thus, $\pi(\cdot)$ is the projection onto G “along” v . For each $n < \infty$, we define processes $\{x_i^n, i < \infty\}$,

$\{y_i^n, i < \infty\}$, and $\{z_i^n, i < \infty\}$ by $(x_0^n, y_0^n, z_0^n) = (x, x, 0)$, and

$$\begin{aligned} x_{i+1}^n &= \pi(x_i^n + \xi_i), \\ y_{i+1}^n &= y_i^n + \xi_i, \\ z_{i+1}^n &= z_i^n + [\pi(x_i^n + \xi_i) - \xi_i]. \end{aligned}$$

Define the interpolations $x^n(t) = n^{-1/2}x_{[t/n]}^n$, $y^n(t) = n^{-1/2}y_{[t/n]}^n$, and $z^n(t) = n^{-1/2}z_{[t/n]}^n$. Then we can write

$$x^n(\cdot) = y^n(\cdot) + z^n(\cdot).$$

Because the process $y^n(\cdot)$ does not involve the constraint, its behavior is easy to determine. As remarked previously, $y^n(\cdot)$ will converge weakly to the solution to $dy(t) = \sigma dw(t)$, $y(0) = x$. The process $z^n(\cdot)$ plays the role of a bookkeeping device, by recording the effects of the projections in such a way that $x^n(\cdot)$ can be recovered from $y^n(\cdot)$ by adding $z^n(\cdot)$. The $z^n(\cdot)$ has three important properties. First, $z^n(\cdot)$ can change only at times t such that $x^n(t) \in \partial G$. Thus, the evolution of $x^n(\cdot)$ is the same as that of $y^n(\cdot)$ when $x^n(\cdot)$ is away from ∂G . Second, the direction of change of $z^n(\cdot)$ is determined by the constraint mechanism, and the last property is that the change in $z^n(\cdot)$ at any given time is the minimal amount needed to keep $x^n(\cdot)$ in G .

For the scaled queueing systems referred to at the beginning of this section, as well as many other constrained processes, there are analogous decompositions into an essentially “unconstrained” part for which it is relatively easy to find a diffusion approximation, plus a term that records the effects of the constraint. Examples can be found in Chapter 8.

The properties of $y^n(\cdot)$ and $z^n(\cdot)$ given in the example above suggest that the following properties should be expected of any diffusion $x(\cdot)$ claiming to approximate $x^n(\cdot)$ (for large n).

- Away from the constraining boundary ∂G that is inherited from the constraint on $x^n(\cdot)$, the process $x(\cdot)$ should behave as the ordinary diffusion approximation to $y^n(\cdot)$.
- When $x(\cdot)$ attempts to cross ∂G it should be returned to G by the action of a compensating process. The direction in which this compensating process “pushes” $x(\cdot)$ will be determined by the constraint mechanism and should be the minimal amount required to keep $x(\cdot)$ in G .

It will be seen below that reflecting diffusions satisfy these properties. Owing to the range of processes for which reflected diffusions are the appropriate diffusion approximation, we must deal with a wide variety of possible settings, e.g., domains with corners, oblique reflections, and multi-valued directions of reflection.

In addition to problems in which reflecting boundaries arise as the result of a diffusion approximation, there are problems in which they are used to bound the state space of a process for numerical purposes. See, in particular, Chapter 11.

Remark. As the discussion above suggests, the term “reflecting” is a bit of a misnomer. It may be that “constrained” diffusion with a few more terms describing the method of constraint would be more accurate. Nonetheless we will follow current convention and retain the adjective “reflecting.”

The approach taken in this book to the study of reflecting diffusions is based on the use of the *Skorokhod Problem*. This approach goes back to Skorokhod in the case of one dimension and has been used by several authors since then in a variety of contexts [24, 36, 35, 57, 92, 106, 117]. It is also possible to base an approach to the study of reflecting diffusions on the *submartingale problem* [115]. The submartingale problem gives an alternative method of characterizing processes with certain boundary behaviors. There are some processes for which the only existing proofs of weak uniqueness (which will be essential for our purposes) are based on the submartingale problem formulation. Examples are diffusions on domains with “sticky” boundary conditions [115] and certain classes of reflecting Brownian motion models [119]. We have chosen to use the Skorokhod Problem approach to reflecting diffusions because it is intuitively appealing and relatively simple to apply. However, the reader should be aware that the basic property we will need for the numerical methods developed in this book to apply to a given reflecting diffusion model is the weak sense uniqueness. We note that because the definition of a weak sense solution is independent of the method of characterization, the methods of generating convergent numerical schemes presented in Chapter 5 and elsewhere and the proofs of convergence are applicable in general, regardless of the particular method used to prove the uniqueness.

Let $G \subset \mathbb{R}^k$ be a closed set and assume that G is the closure of its interior. To each point $x \in \partial G$, we associate a set $r(x) \subset \{y \in \mathbb{R}^k : |y| = 1\}$ called the *directions of reflection*. Let $\psi \in C^k[0, \infty)$ be a given path that satisfies $\psi(0) \in G$. For a function of bounded variation η mapping $[0, \infty)$ to \mathbb{R}^k , we let $|\eta|(t)$ denote the total variation over the interval $[0, t]$, and let μ_η denote the measure on $[0, \infty)$ which is defined by the total variation. The precise definition of the Skorokhod Problem is as follows.

Definition 4.1 (Skorokhod Problem) *Let $\psi \in C^k[0, \infty)$ with $\psi(0) \in G$ be given. Then (ϕ, η) solves the Skorokhod Problem for ψ (with respect to G and r) if*

1. $\phi = \psi + \eta$, $\phi(0) = \psi(0)$,
2. $\phi(t) \in G$ for $t \in [0, \infty)$,

3. $|\eta|(t) < \infty$ for all $t < \infty$,
4. $|\eta|(t) = \int_{(0,t]} I_{\{\phi(s) \in \partial G\}} d|\eta|(s)$,
5. there exists measurable $\gamma : [0, T] \rightarrow \mathbb{R}^k$ such that $\gamma(s) \in r(\phi(s))$ (μ_η w.p.1) and $\eta(t) = \int_{(0,t]} \gamma(s) d|\eta|(s)$.

Although the definition appears abstract, it gives a very convenient mathematical tool for the construction and analysis of reflecting diffusions. The level of abstraction is in part due to the fact that we need to define a reflected version for each path in $C^k[0, \infty)$. It is possible to give a simpler definition if we restrict to only smooth paths, but this is not sufficient if we wish to deal with reflecting diffusions. Some comments on the definition are as follows. The function ϕ is to be viewed as the natural constrained version of ψ . Part 1 of the definition states that ϕ starts at the same point as ψ , and that ϕ will be obtained from ψ by adding a function η which will “push” ϕ in the proper direction at the proper time. Part 2 simply states that ϕ is constrained to G . Part 3 is useful in establishing a semimartingale decomposition for reflected diffusions and for obtaining estimates for such processes. Part 4 implies that η may only push ϕ while ϕ is on the boundary of G and that η does not change while ϕ is off the boundary. Thus, the evolution of ϕ duplicates that of ψ when ϕ is away from the boundary. Lastly, part 5 implies that η is only allowed push in the directions consistent with the current position of ϕ when it is on the boundary.

A comparison of these properties with the properties that one would expect of the sample paths of reflecting diffusions indicates a close connection. Our definition of a stochastic differential equation with reflection (SDER) is as follows. Consider a probability space (Ω, \mathcal{F}, P) on which is defined a filtration $\{\mathcal{F}_t, t \geq 0\}$. The definition will be given without a control. Extending the definition to include a control is straightforward. Let $b(\cdot)$ and $\sigma(\cdot)$ be functions of the appropriate dimensions and suppose that $\{w(t), t \geq 0\}$ is an r -dimensional \mathcal{F}_t -Wiener process.

Definition 4.2. (SDER) *The \mathcal{F}_t -adapted process $x(\cdot)$ is a solution to the SDER for the domain G , directions of reflection $r(\cdot)$, initial condition $x(0) \in G$ and Brownian motion $\{w(t), t \geq 0\}$, if $x(t) \in G$ for all $t \geq 0$ (w.p.1), and*

$$x(t) = x + \int_0^t b(x(s)) ds + \int_0^t \sigma(x(s)) dw(s) + z(t),$$

where

$$|z|(t) = \int_0^t I_{\{x(s) \in \partial G\}} d|z|(s) < \infty,$$

and where there exists measurable $\gamma(s) \in r(x(s))$ (μ_z w.p.1) such that

$$z(t) = \int_0^t \gamma(s) d|z|(s)$$

(w.p.1).

In other words, $(x(\cdot), z(\cdot))$ should solve (on a pathwise and w.p.1 basis) the SP for $\psi(\cdot) = x + \int_0^{\cdot} b(x(s))ds + \int_0^{\cdot} \sigma(x(s))dw(s)$.

As in the case of diffusions without boundaries, there are two senses in which solutions can be said to exist and also two senses in which they can be said to be unique. These definitions are simply the exact analogues of those for the case of no boundary. In general, the existence of strong or weak solutions to the SDER and the relevant uniqueness properties depend on regularity properties of G and $r(\cdot)$. Those approaches that are based on the Skorokhod Problem use these regularity properties to derive estimates for the mapping $\psi \rightarrow (\phi, \eta)$ defined by solving the Skorokhod problem. We next present a few very simple examples for which this mapping is Lipschitz continuous. This allows an elementary derivation of (3.7), after which the analysis proceeds just as in the case of unconstrained diffusions. The basic differences between these examples and the more general situations considered in [35, 92, 106, 117] are the more involved calculations required to get (3.7).

Example 4.3. (Anderson and Orey) [3] Let n be a given unit vector, and let G be the interior of $\{x : \langle x, n \rangle \leq 0\}$. Thus, n is the outward normal at all points of ∂G . Let r be a unit vector that satisfies $\langle r, n \rangle < 0$. We will take $r(x) = \{r\}$ for all $x \in \partial G$. Suppose that $\phi(\cdot)$ is given with $\phi(0) \in G$. If we define

$$\begin{aligned} |\eta|(t) &= - \left(0 \vee \sup_{0 \leq s \leq t} \langle \psi(s), n \rangle \right) / \langle r, n \rangle, \\ \eta(t) &= |\eta|(t)r, \\ \phi(t) &= \psi(t) + \eta(t), \end{aligned}$$

then $(\phi(\cdot), \eta(\cdot))$ solves the Skorokhod Problem for $\psi(\cdot)$. From these equations, it is easy to show that if $(\phi_i(\cdot), \eta_i(\cdot))$ solves the Skorokhod Problem for $\psi_i(\cdot)$, $i = 1, 2$, then

$$\sup_{0 \leq s \leq t} |\eta_1(s) - \eta_2(s)| \leq \left(\frac{1}{\langle r, n \rangle} \right) \sup_{0 \leq s \leq t} |\psi_1(s) - \psi_2(s)|, \quad (4.1)$$

and, therefore,

$$\sup_{0 \leq s \leq t} |\phi_1(s) - \phi_2(s)| \leq \left(1 + \frac{1}{\langle r, n \rangle} \right) \sup_{0 \leq s \leq t} |\psi_1(s) - \psi_2(s)|. \quad (4.2)$$

These estimates allow a very simple study of the corresponding SDER. Let $\phi(\cdot) = \Gamma(\psi(\cdot))$ denote the mapping $\psi(\cdot) \rightarrow \phi(\cdot)$, and let $z_i, i = 1, 2$, be measurable, \mathcal{F}_t -adapted processes with continuous sample paths. Define

$$\tilde{y}_i(t) = x(0) + \int_0^t b(z_i(s))ds + \int_0^t \sigma(z_i(s))dw(s)$$

and $y_i(t) = \Gamma(\tilde{y}_i)(t)$ for $t \geq 0$ and $i = 1, 2$. Obviously, the $\tilde{y}_i(\cdot)$ are measurable, \mathcal{F}_t -adapted processes with continuous sample paths. Equation (4.2) implies uniqueness of the pair (ϕ, η) given ψ . If $x \in G$ and $\psi(t) = x$ for all $t \geq 0$, then $(\phi(t), \eta(t)) = (x, 0)$ is the solution to the SP. This fact together with (4.2) imply the continuity of ϕ whenever ψ is continuous and $\phi = \Gamma(\psi)$. Thus, the mapping Γ is a Lipschitz continuous mapping of $C^k[0, \infty)$ into itself. Furthermore, uniqueness of the solution to the SP and our explicit representation of the solution imply that $\phi(t)$ is a measurable function of $\{\psi(s), s \in [0, t]\}$. Therefore, the processes $y_i(\cdot)$, $i = 1, 2$, are measurable, \mathcal{F}_t -adapted processes with continuous sample paths. Define $\Delta\tilde{y}(t) = \tilde{y}_1(t) - \tilde{y}_2(t)$ and also $\Delta y(\cdot)$ and $\Delta z(\cdot)$ in an analogous fashion. Then, under (A3.1), equation (3.7) gives

$$E \left[\sup_{0 \leq s \leq t} |\Delta\tilde{y}(s)|^2 \right] \leq L \int_0^t E \left[\sup_{0 \leq u \leq s} |\Delta z(u)|^2 \right] ds$$

for some $L \in (0, \infty)$. Using (4.2), we have

$$E \left[\sup_{0 \leq s \leq t} |\Delta y(s)|^2 \right] \leq L_1 \int_0^t E \left[\sup_{0 \leq u \leq s} |\Delta z(u)|^2 \right] ds,$$

where $L_1 = 4L$. Therefore, the existence, uniqueness, and other properties follow from the same arguments as those used for the unreflected versions in Section 1.4. Note that we should now define the sequence used in the existence proof by $x_0(t) = x(0)$,

$$\tilde{x}_{n+1}(t) = x(0) + \int_0^t b(\tilde{x}_n(s), s)ds + \int_0^t \sigma(\tilde{x}_n(s), s)dw(s),$$

$$x_{n+1}(t) = \Gamma(\tilde{x}_{n+1})(t), \quad z_{n+1}(t) = x_{n+1}(t) - \tilde{x}_{n+1}(t).$$

Then $(x_n(\cdot), z_n(\cdot))$ converges to a pair of processes $(x(\cdot), z(\cdot))$, where $x(\cdot)$ is a solution to the SDER. ■

Suppose that G possesses a smooth boundary with exterior normal $n(\cdot)$ and that $r(\cdot)$ is a smooth vector field satisfying $\langle r(x), n(x) \rangle < 0$ for all $x \in \partial G$. Then, by use of appropriate local coordinate systems and stopping times, the results of Example 4.3 imply strong existence and uniqueness for the SDER corresponding to G and $r(\cdot)$ [3].

The next example considers the choice of G and $r(\cdot)$ that is probably most natural when a reflecting boundary condition must be imposed simply for the purposes of bounding the state space.

Example 4.4. Let $d_i > c_i$, $i = 1, \dots, k$ be given, and define

$$G = \{x : c_i < x_i < d_i \text{ for } i = 1, 2, \dots, k\}.$$

Let $n(\cdot)$ denote the set of outward normals and suppose $r(\cdot) = -n(\cdot)$. Thus, the constraining action is applied along the direction of the inward normal.

Note that $r(\cdot)$ is multi-valued at any of the “corner” points of ∂G . It is proved in [36] that the solution mapping to the Skorokhod Problem is again Lipschitz continuous in the sense of equations (4.1) and (4.2), although the coefficients in this case are $k^{1/2}$ and $k^{1/2} + 1$, respectively. Therefore, SDER may be solved and uniqueness proved just as in Example 4.3. ■

Remarks. In general, when good estimates are available on the Skorokhod Problem analogues of the statements regarding existence, uniqueness, and so on that hold for the case unreflected processes carry over to the reflected case. It should also be noted that the Girsanov transformation method (see Section 1.3) applies here as well and in precisely the same way. Other examples of reflecting or constrained processes will appear in later chapters.

1.5 Processes with Jumps

In this section we will discuss Markov models with jumps. We are interested in models of the form

$$x(t) - x(0) = \int_0^t b(x(s))ds + \int_0^t \sigma(x(s))dw(s) + J(t),$$

where $b(\cdot)$ and $\sigma(\cdot)$ are as in the preceding sections and the $J(t)$ term produces the jumps. For the jump term we would like to specify (at least approximately) the probability that a jump occurs in any small time interval together with the distribution of any resulting jumps as functions of the past history of the process. Between jumps, the term $J(\cdot)$ is constant. In order to preserve the Markov property, the “jump intensity” at time t and distribution of any jumps at time t should depend only on $\lim_{s \uparrow t} x(s)$.

Let $\lambda(\cdot)$ be a function mapping \mathbb{R}^k into $[0, \infty)$, and let $\bar{\Pi}(x, dy)$ be a probability transition kernel on \mathbb{R}^k . For now assume that $\lambda(\cdot)$, $b(\cdot)$ and $a(\cdot)$ are all continuous and that $\bar{\Pi}(x, \cdot)$ is continuous in x in the topology of weak convergence. Let $\Delta t > 0$ be small and fix t . A rough description of the term $J(\cdot)$ that is consistent with the properties described above is as follows. Let $x(t-) = \lim_{s \uparrow t} x(s)$. With probability equal to $\lambda(x(t-))\Delta t + o(\Delta t)$, $J(\cdot)$ will jump once at some time in the interval $[t, t + \Delta t]$. The probability of two or more jumps is $o(\Delta t)$. Thus, $\lambda(\cdot)$ gives the overall jump rate. Given that a jump has occurred, its distribution will be given approximately by $\bar{\Pi}(x(t-), \cdot)$. Between jumps, the process $x(\cdot)$ behaves like a diffusion process with no jumps and with the local properties described by $b(\cdot)$ and $a(\cdot)$.

The general theory and treatments of various approaches to such processes can be found in [61, 64, 65]. For the purposes of this book, the processes may be constructed and the needed properties proved using rather simple arguments.

Construction of the Process. Under the assumptions we will make below, the weak sense existence and uniqueness properties for this class of jump diffusion processes are essentially determined by the corresponding properties of the analogous diffusion with no jumps. For example, assume that (A3.1) holds. This assumption is sufficient to guarantee the existence of a unique strong sense solution to the diffusion with no jumps. It will turn out to imply (under our assumptions only, and not in the general case as considered in [65]) the corresponding result for the jump diffusion. A construction of the process is as follows. Assume that we are given a filtration \mathcal{F}_t on a probability space (Ω, \mathcal{F}, P) , together with an \mathcal{F}_t -Wiener process $w(\cdot)$. Let $\Pi(\cdot)$ be a probability measure on the Borel subsets of \mathbb{R}^n that has compact support Γ . Suppose that on the same probability space there are defined mutually independent sequences of iid random variables $\{\tau_n, n < \infty\}$ and $\{\rho_n, n < \infty\}$, where the τ_n are exponentially distributed with mean $1/\lambda$ and the ρ_n have distribution $\Pi(\cdot)$. Assume also that these random variables are independent of $w(\cdot)$. Let $\nu_0 = 0$ and $\nu_{n+1} = \nu_n + \tau_n$. The ν_n will be the jump times of the process. Let $q : \mathbb{R}^k \times \mathbb{R}^n \rightarrow \mathbb{R}^k$ be a bounded measurable function. Starting with a given initial condition x , we construct a solution $x_1(\cdot)$ to

$$dx_1 = b(x_1)dt + \sigma(x_1)dw, \quad x_1(0) = x.$$

Then define

$$x(t) = \begin{cases} x_1(t), & t \in [0, \nu_1) \\ x_1(\nu_1) + q(x_1(\nu_1), \rho_1), & t = \nu_1. \end{cases}$$

We continue by repeating the process. We define

$$dx_2 = b(x_2)dt + \sigma(x_2)dw, \quad x_2(\nu_1) = x(\nu_1),$$

and then define

$$x(t) = \begin{cases} x_2(t), & t \in [\nu_1, \nu_2) \\ x_2(\nu_2) + q(x_2(\nu_2), \rho_2), & t = \nu_2, \end{cases}$$

and so on. The process thus constructed will be defined for all $t \geq 0$ since $\nu_n \rightarrow \infty$ as $n \rightarrow \infty$ (w.p.1). The mutual independence of the components used to construct the process implies the Markov property.

The process we have constructed satisfies the description given at the beginning of the section, with

$$\lambda(x) = \lambda \int_{\{\rho: q(x, \rho) \neq 0\}} \Pi(d\rho) \leq \lambda \tag{5.1}$$

and

$$\bar{\Pi}(x, A) = \int_{\{\rho: q(x, \rho) \in A, q(x, \rho) \neq 0\}} \Pi(d\rho). \tag{5.2}$$

Consider the following assumptions on $\lambda(\cdot)$ and $\bar{\Pi}(\cdot, \cdot)$.

A5.1. *The function $\lambda(x)$ is continuous and uniformly bounded, and the support of $\bar{\Pi}(x, \cdot)$ is contained in some compact set that is independent of x . Furthermore, $\bar{\Pi}(x, \cdot)$ is continuous as a mapping from the $x \in \mathbb{R}^k$ into the space of probability measures endowed with the topology of weak convergence.*

It can be shown [65] that if $\lambda(\cdot)$ and $\bar{\Pi}(\cdot, \cdot)$ satisfy (A5.1), then $\lambda < \infty$, a probability measure $\Pi(\cdot)$, and a bounded measurable function $q(\cdot, \cdot)$ can be found so that (5.1) and (5.2) hold. For convenience, we will consider λ , $\Pi(\cdot)$, and $q(\cdot, \cdot)$ as characterizing the jump part of a jump diffusion.

The same approach to constructing jump diffusions can be used if we weaken the assumptions on $b(\cdot)$ and $\sigma(\cdot)$. For example, the Girsanov transformation method (see Section 1.3) can be used exactly as before to construct a weak sense solution for cases in which the drift is discontinuous.

It will turn out to be useful to have a representation of these processes as solutions to a SDE, analogous to that of diffusions without jumps. In order to account for the jumps, an additional driving term is introduced in the form of a Poisson random measure. The definition we use is not the most general, but is sufficient for our purposes. By an integer valued measure we mean a measure that always yields an integer as the measure of any measurable set.

Poisson Random Measures. Assume that we are given a probability space (Ω, \mathcal{F}, P) on which a filtration \mathcal{F}_t is defined and a probability measure $\Pi(\cdot)$ on the Borel subsets of \mathbb{R}^n . Let \mathcal{F}_{t-} be the σ -algebra generated by the union of \mathcal{F}_s , $s < t$. Then an \mathcal{F}_t -Poisson random measure with intensity measure $h(dt dy) = \lambda dt \times \Pi(dy)$ is a measurable mapping $N(\cdot)$ from (Ω, \mathcal{F}, P) into the space of integer valued positive measures on $\mathbb{R}^+ \times \mathbb{R}^n$ with the following properties:

1. For every $t \geq 0$ and every Borel subset A of $[0, t] \times \mathbb{R}^n$, $N(A)$ is \mathcal{F}_t -measurable.
2. For every $t \geq 0$ and every Borel subset A of $[t, \infty) \times \mathbb{R}^n$, $N(A)$ is independent of \mathcal{F}_{t-} .
3. $E[N(A)] = h(A)$ for every Borel subset A of $\mathbb{R}^+ \times \mathbb{R}^n$.

A Poisson random measure is the counting measure of a Poisson point process on $\mathbb{R}^+ \times \mathbb{R}^n$. We use the “points” of a realization of the random measure to identify the times and the magnitudes of jumps appearing in the jump Markov process defined by the stochastic differential equation given below. It can be shown [65] that, for any sequence $\{A_i\}$ of pairwise disjoint Borel subsets of $\mathbb{R}^+ \times \mathbb{R}^n$, the random variables $N(A_i)$ are independent, and, furthermore, each $N(A_i)$ is a Poisson random variable with mean value $h(A_i)$. The Poisson random measure representation provides a convenient representation for the jump term, in that it provides the correct relation

between the filtration of the process and the desired properties of the times between jumps and jump distribution. Suppose that $\{(\nu_i, \rho_i), i < \infty\}$ are the point masses of the Poisson random measure. Conditioned on \mathcal{F}_{t-} , the random variable $\inf\{\nu_i - t : \nu_i \geq t\}$ is exponentially distributed with mean $1/\lambda$. Given that (ν, ρ) is a point mass of $N(\cdot)$, ρ is distributed according to $\Pi(\cdot)$ (w.p.1).

Let \mathcal{F}_t be a filtration on a probability space (Ω, \mathcal{F}, P) , and let $w(\cdot)$ and $N(\cdot)$ be an \mathcal{F}_t -adapted Wiener process and Poisson random measure, respectively. By a solution of the SDE

$$dx(t) = b(x(t))dt + \sigma(x(t))dw(t) + dJ(t), \quad (5.3)$$

$$J(t) = \int_{[0,t] \times \mathbb{R}^n} q(x(s-), \rho) N(ds d\rho),$$

together with the \mathcal{F}_0 -measurable initial condition $x(0)$, what is meant is an \mathcal{F}_t -adapted process $x(\cdot)$ with paths in $D^k[0, \infty)$ which satisfies the integrated form

$$x(t) = \int_0^t b(x(s))ds + \int_0^t \sigma(x(s))dw(s) + \int_{[0,t] \times \mathbb{R}^n} q(x(s-), \rho) N(ds d\rho). \quad (5.4)$$

In complete analogy with case of diffusions with no jumps, we have definitions of weak and strong existence, and weak and strong uniqueness. Because we will only use the weak sense properties in both cases, statements are only given for these cases.

Weak Existence. We say that weak existence holds if given any probability measure μ on \mathbb{R}^k , there exists a probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process $w(\cdot)$, an \mathcal{F}_t -Poisson random measure $N(\cdot)$ and an \mathcal{F}_t -adapted process $x(\cdot)$ satisfying (5.4) for all $t \geq 0$, as well as $P\{x(0) \in A\} = \mu(A)$.

Weak Uniqueness. Suppose we are given two weak sense solutions

$$\{(\Omega_i, \mathcal{F}_i, P_i), \mathcal{F}_{i,t}, w_i(\cdot), N_i(\cdot), x_i(\cdot)\}, i = 1, 2,$$

to (5.4). We say that weak uniqueness holds if equality of the distributions induced on \mathbb{R}^k by $x_i(0)$ under $P_i, i = 1, 2$, implies equality of the distributions induced on $D^k[0, \infty)$ by $x_i(\cdot)$ under $P_i, i = 1, 2$.

Existence and Uniqueness of Solutions. As remarked previously, the weak sense existence and uniqueness properties of these jump diffusions basically follow from the corresponding properties in the case of no jumps. Weak sense existence has been discussed already. Weak sense uniqueness will also hold for the jump diffusion if it holds for the corresponding case

without jumps [this assertion is true under (A5.1), but not under more general assumptions that are often used, e.g., [65]]. This follows under our assumptions because with probability one, the jump times are isolated and tend to infinity. Let x_1 and x_2 be weak sense solutions to (5.4) that start at x (we neglect specifying the probability spaces, the Wiener processes, etc.). Let ν_1 and ν_2 be the time of the first atoms for the Poisson random measures corresponding to x_1 and x_2 , respectively. [This may or may not equal the first jump time, since there may be points (x, ρ) for which $q(x, \rho) = 0$.] Let ρ_1 and ρ_2 identify the spatial coordinates of these atoms. Because the two Poisson random measures have the same intensity measure, the distributions of ν_1 and ν_2 are the same, and the same is true of ρ_1 and ρ_2 . By the weak sense uniqueness, and the independence of the Wiener process and the Poisson random measure, the distributions induced by x_1 and x_2 up to and including the times ν_1 and ν_2 are the same. Repeating the procedure, we obtain equivalence of distributions up to the time of occurrence of the second atom, and so on. The fact that the occurrence times of the atoms tend to infinity (w.p.1) implies the equivalence of the distributions on $[0, \infty)$.

Itô's Formula. The following formula follows directly from the definition of a solution to (5.4) and the corresponding formula for the case of diffusions with no jumps.

$$f(x(t)) = f(x(0)) + \int_0^t (\mathcal{L}f)(x(s))ds + \int_0^t f'_x(x(s))\sigma(x(s))dw(s) + J_f(t),$$

where $a(x) = \sigma(x)\sigma'(x)$ as usual,

$$(\mathcal{L}f)(x) = f'_x(x)b(x) + \frac{1}{2} \text{tr}[f_{xx}(x)a(x)] + \lambda(x) \int_{\Gamma} [f(x+y) - f(x)] \bar{\Pi}(x, dy), \quad (5.5)$$

and

$$\begin{aligned} J_f(t) &= \sum_{s \leq t} [f(x(s)) - f(x(s-))] \\ &\quad - \int_0^t \lambda(x(s)) \int_{\Gamma} [f(x(s)+y) - f(x(s))] \bar{\Pi}(x(s), dy) ds. \end{aligned}$$

Since $J_f(t)$ is a martingale,

$$Ef(x(t)) = Ef(x(0)) + E \int_0^t \mathcal{L}f(x(s))ds.$$

Constrained Diffusions With Jumps. We can define a reflected jump diffusion that is analogous to the reflected diffusion. Let the domain in

which the process is to stay be G . We need only define what happens if a jump takes the process out of G .

Suppose that a jump diffusion with components λ , $\Pi(\cdot)$, and $q(\cdot, \cdot)$ is given. Suppose also that the desired behavior for the reflected jump diffusion is the following. Whenever a jump would take the process out of G to a point y , then the process is to be “instantaneously” returned to a point $c(y) \in G$. Assume that $c(\cdot)$ is measurable. This desired behavior of the process can be achieved by replacing $q(\cdot, \cdot)$ by $\bar{q}(\cdot, \cdot)$ defined by

$$c(x + q(x, \rho)) = x + \bar{q}(x, \rho).$$

Thus, without loss, we can assume that the following convention is always in effect. *Whenever we are dealing with a reflected jump diffusion, we will assume that $q(\cdot, \cdot)$ has been chosen so that*

$$x + q(x, \rho) \in G \text{ for all } x \in G \text{ and all } \rho.$$

Controlled Diffusions with Jumps. In analogy with Section 1.3, we have the following definition. Let $u(\cdot)$ be a control process taking values in \mathcal{U} , let $w(\cdot)$ be a Wiener process, let $N(\cdot)$ be a Poisson random measure, and assume they are all defined on a common probability space and that $w(\cdot)$ and $N(\cdot)$ are independent. We say that the control $u(\cdot)$ is *nonanticipative* with respect to the pair $(w(\cdot), N(\cdot))$ if there exists a filtration \mathcal{F}_t such that $u(\cdot)$ is \mathcal{F}_t -measurable, $w(\cdot)$ is an \mathcal{F}_t -Wiener process, and $N(\cdot)$ is an \mathcal{F}_t -Poisson random measure. We say that $u(\cdot)$ is an *admissible control law with respect to $(w(\cdot), N(\cdot))$* or that *the triple $(u(\cdot), w(\cdot), N(\cdot))$ is admissible* if $u(\cdot)$ is nonanticipative with respect to $(w(\cdot), N(\cdot))$.

We consider controlled SDE's of the form

$$dx(t) = b(x(t), u(t))dt + \sigma(x(t))dw(t) + dJ(t), \quad (5.6)$$

$$J(t) = \int_{[0,t] \times \mathbb{R}^n} q(x(s-), \rho)N(ds d\rho),$$

where, as usual, we interpret this equation via its integrated form. We have the following definitions.

Weak Existence. Suppose we are given probability distributions Λ and μ on the sample space of the triple $(u(\cdot), w(\cdot), N(\cdot))$ and \mathbb{R}^k , respectively. We say that weak existence holds if there exists a probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process $w(\cdot)$, an \mathcal{F}_t -Poisson random measure $N(\cdot)$, an \mathcal{F}_t -adapted control process $u(\cdot)$ [i.e., the triple $(u(\cdot), w(\cdot), N(\cdot))$ is admissible], and an \mathcal{F}_t -adapted process $x(\cdot)$, such that Λ and μ are the distributions induced by $(u(\cdot), w(\cdot), N(\cdot))$ and $x(0)$ under P , and such that $x(\cdot)$ satisfies (5.6) for all $t \geq 0$.

Weak Uniqueness. Assume we are given two weak sense solutions

$$\{(\Omega_i, \mathcal{F}_i, P_i), \mathcal{F}_{i,t}, w_i(\cdot), N_i(\cdot), u_i(\cdot), x_i(\cdot)\}, i = 1, 2,$$

to (5.6). We say that weak uniqueness holds if equality of the joint distributions of $(u_i(\cdot), w_i(\cdot), N_i(\cdot), x_i(0))$ under $P_i, i = 1, 2$, implies the equality of the distributions of $(x_i(\cdot), u_i(\cdot), w_i(\cdot), N_i(\cdot), x_i(0))$ under $P_i, i = 1, 2$.

By arguing in the same way as for the case without control, one can show that if there is weak sense uniqueness for the case of diffusions with arbitrary \mathcal{U} -valued admissible control laws and initial conditions, then there is weak sense uniqueness here as well.

For fixed $\alpha \in \mathcal{U}$, we define the differential operator for solutions of (5.6) by

$$\begin{aligned} (\mathcal{L}^\alpha f)(x) &= f'_x(x)b(x, \alpha) + \frac{1}{2}\text{tr}[f_{xx}(x)a(x)] \\ &\quad + \lambda(x) \int_\Gamma [f(x+y) - f(x)]\bar{\Pi}(x, dy). \end{aligned}$$

For a control process $u(\cdot)$, define the operator by $\mathcal{L}^{u(\cdot)}$. If $u(\cdot)$ is an admissible control law, then Itô's formula for functions of the controlled process (5.6) continues to hold with this definition of the differential operator.

2

Controlled Markov Chains

The main computational techniques in this book require the approximation of an original controlled processes in continuous time by appropriately chosen controlled finite state Markov chains. In this chapter, we will define some of the canonical control problems for the Markov chain models which will be used in the sequel as “approximating processes.” The cost functions will be defined. The functional equations which are satisfied by these cost functions for fixed controls, as well as the functional equations satisfied by the optimal cost functions (the dynamic programming or Bellman equation), will be obtained by exploiting the Markov property and the uniqueness of their solutions is shown, under appropriate conditions. These are the equations which will have to be solved in order to get the required approximate solutions to the original control or optimal control problem. The simplest case, where there is no control or where the control is fixed, is dealt with in Section 2.1, and the recursive equations satisfied by the cost functionals are obtained. A similar method is used to get the recursive equations for the optimal value functions for the controlled problems. The optimal stopping problem is treated in Section 2.2. This is a relatively simple control problem, because the only decision to be made is the choice of the moment at which the process is to be stopped. This problem will illustrate the basic ideas of dynamic programming for Markov chains and introduce the fundamental *principle of optimality* in a simple way. Section 2.3 concerns the general discounted cost problem. Section 2.4 deals with the optimization problem when the control stops at the first moment of reaching a target or stopping set. The basic concept of *contraction map* is introduced and its role in the solution of the functional equations for the costs is emphasized. Section 2.5 gives the results for the case where the process is of interest over a finite time only. The chapter contains only a brief outline. Further information concerning controlled or uncontrolled Markov chain models can be found in the standard references [10, 44, 70, 74, 100, 121, 125].

2.1 Recursive Equations for the Cost

Let $\{\xi_n, n < \infty\}$ be a Markov chain with the time independent transition probabilities $p(x, y) = P\{\xi_{n+1} = y | \xi_n = x\}$ on a finite state space S . It is sufficient for our purposes that S be finite, because that will be the case for the Markov chain models which will be used for the approximations of the continuous time control problems in the later chapters. Let E_x denote the expectation of functionals of the chain, given that the initial condition is x . We will define the cost functionals and give the equations which they satisfy for many of the cases which will be of interest later in the book.

2.1.1 STOPPING ON FIRST EXIT FROM A GIVEN SET

Let $\partial S \subset S$ be a given subset of the state space such that we stop the chain at the moment $N = \min\{n : \xi_n \in \partial S\}$ of first reaching ∂S . We use the “boundary” notation ∂S to denote this set, since later on this set will be a “discretization” of the boundary of the state space of the processes which we will be approximating, and it is useful to introduce the notation at this point.

Suppose that

$$E_x N < \infty, \text{ for all } x \in S - \partial S. \quad (1.1)$$

For given functions $c(\cdot)$ and $g(\cdot)$, define the total cost until stopping by

$$W(x) = E_x \left[\sum_{n=0}^{N-1} c(\xi_n) + g(\xi_N) \right]. \quad (1.2)$$

$c(\cdot)$ is called a “running” cost or cost “rate” and $g(\cdot)$ is a “stopping” or “boundary” cost. It follows from the definitions that $W(x) = g(x)$ for $x \in \partial S$. A functional equation for $W(\cdot)$ can be derived by exploiting the Markov property of $\{\xi_n, n < \infty\}$ and rewriting (1.2) as follows: For $x \in S - \partial S$,

$$\begin{aligned} W(x) &= c(x) + E_x \left\{ E_x \left[\sum_{n=1}^{N-1} c(\xi_n) + g(\xi_N) | \xi_1 \right] \right\} \\ &= c(x) + E_x W(\xi_1) \\ &= c(x) + \sum_{y \in S} p(x, y) W(y). \end{aligned} \quad (1.3a)$$

For $x \in \partial S$, we have

$$W(x) = g(x) \quad (1.3b).$$

A Vector Form of (1.3). It is often useful to write (1.3) in a vector form. There are two convenient methods for doing this, depending on whether the states in the “stopping” or boundary set ∂S are included or not. Let $|S|$

denote the number of elements in the set S . For the first method, define the vector $\tilde{C} = \{\tilde{C}(x), x \in S\}$ by $\tilde{C}(x) = c(x)$ for $x \in S - \partial S$ and $\tilde{C}(x) = g(x)$ for $x \in \partial S$. Define the vector of costs $\tilde{W} = \{W(x), x \in S\}$. All vectors are column vectors unless otherwise mentioned. Because our interest in the chain stops when it first enters the stopping set ∂S , let us “kill” it at that point, and define the “killed” transition probability matrix $\tilde{R} = \{\tilde{r}(x, y); x, y \in S\}$, where $\tilde{r}(x, y) = p(x, y)$ for $x \in S - \partial S$ and $\tilde{r}(x, y) = 0$ for $x \in \partial S$. Then (1.3) can be written as the $|S|$ dimensional equation:

$$\tilde{W} = \tilde{R}\tilde{W} + \tilde{C}. \quad (1.4)$$

By the finiteness (with probability one) of the stopping times which is implied by (1.1), we have $\tilde{R}^n \rightarrow 0$. Thus, we can interpret \tilde{R} as a one step Markov transition matrix for a finite state chain all of whose states are eventually killed. The probability that the chain is “killed” when the state is x is $1 - \sum_y \tilde{r}(x, y)$.

Equation (1.4) has a unique solution and its components are given by (1.2). To see this, simply iterate (1.4) n times to get

$$\tilde{W} = \tilde{R}^n \tilde{W} + \sum_{i=0}^{n-1} \tilde{R}^i \tilde{C}.$$

Now let $n \rightarrow \infty$. The first term on the right side goes to zero and the limit of the second term is just the vector of costs with components defined by (1.2).

For an alternative and often preferred method of writing the cost function in vector form, one eliminates the states in the boundary set ∂S and uses the reduced state space $S - \partial S$, as follows: Define $r(x, y) = p(x, y)$ for $x, y \in S - \partial S$ and define the transition matrix $R = \{r(x, y); x, y \in S - \partial S\}$. Set $W = \{W(x), x \in S - \partial S\}$, and define the vector of cost rates $C = \{C(x), x \in S - \partial S\}$ by $C(x) = c(x) + \sum_{y \in \partial S} p(x, y)g(y)$ for $x \in S - \partial S$. Then we get the $|S - \partial S|$ -dimensional equation to which (1.2) is the unique solution, for $x \in S - \partial S$:

$$W = RW + C. \quad (1.5)$$

2.1.2 DISCOUNTED COST

Let $\beta > 0$, and define the total discounted cost

$$W(x) = E_x \sum_{n=0}^{\infty} e^{-\beta n} c(\xi_n). \quad (1.6)$$

Here there is no explicit “stopping” or “boundary” set. Later, we will allow the discount factor β to depend on the current state. Following the procedure used for (1.2), we can get a functional equation for the $W(x)$ of (1.6) by using the Markov property and writing:

$$\begin{aligned}
W(x) &= E_x \left\{ E_x e^{-\beta} \left[\sum_{n=1}^{\infty} e^{-\beta(n-1)} c(\xi_n) | \xi_1 \right] \right\} + c(x) \\
&= e^{-\beta} E_x W(\xi_1) + c(x) \\
&= e^{-\beta} \sum_y p(x, y) W(y) + c(x).
\end{aligned} \tag{1.7}$$

Next, let us modify (1.6) by introducing an explicit stopping set ∂S as in Subsection 2.1.1, and suppose that the accumulation of the running cost stops at the time N of first entrance into ∂S , and a boundary or stopping cost $g(\cdot)$ is added as in (1.2) above. That is, the total cost is

$$W(x) = E_x \left[\sum_{n=0}^{N-1} e^{-\beta n} c(\xi_n) + e^{-\beta N} g(\xi_N) \right]. \tag{1.8}$$

Then the functional equation for (1.8) is

$$W(x) = \begin{cases} e^{-\beta} E_x W(\xi_1) + c(x), & x \in S - \partial S \\ g(x), & x \in \partial S. \end{cases} \tag{1.9}$$

Define the “discounted” and degenerate transition matrix $R = \{r(x, y), x, y \in S - \partial S\}$ by $r(x, y) = e^{-\beta} p(x, y)$. Define the cost vector $C = \{C(x), x \in S - \partial S\}$ by $C(x) = c(x) + e^{-\beta} \sum_{y \in \partial S} p(x, y) g(y)$ for $x \in S - \partial S$. Define the vector $W = \{W(x), x \in S - \partial S\}$. Then we can write (1.9) in the vector form

$$W = RW + C. \tag{1.10}$$

Equation (1.10) has a unique solution whether or not $E_x N < \infty$ for all x , due to the fact that the discounting implies that $R^n \rightarrow 0$.

State Dependent Discount Factor. In the Markov chain problems which arise as the numerical approximations to the discounted cost continuous time control problems of interest in this book, it is often the case that the discount factor depends on the current state (and possibly on the current control action), even if the discount factor in the original continuous time problem was not state dependent. The costs (1.6) or (1.8) are easily modified to account for this possibility, and in the rest of this chapter, we will use a state (or state and control, if appropriate) dependent discount factor. For $\beta(x) > 0$, let $e^{-\beta(x)}$ be the discount factor when the state is x . The strict positivity of $\beta(x)$ will be dropped below for the so-called instantaneous reflecting states. The appropriate modification of the cost used in (1.8) is

$$W(x) = E_x \sum_{n=0}^{N-1} \exp \left[- \sum_{i=0}^{n-1} \beta(\xi_i) \right] c(\xi_n) + E_x \exp \left[- \sum_{i=0}^{N-1} \beta(\xi_i) \right] g(\xi_N). \quad (1.11)$$

Also, (1.9) and (1.10) continue to hold, but with β replaced by $\beta(x)$ in the line that calculates $W(x)$.

A Special Case: A Reflecting Boundary. The term “reflecting boundary” is used in a loose sense, because there is no special geometric structure assumed on the state space here. But the terminology and formulas will be useful in later chapters when the Markov chains are obtained as “approximating” processes for the “physical” continuous time controlled processes. These will be defined on some set in a Euclidean space, and S will be a “discretization” of that set. These continuous parameter processes might have a reflecting boundary, and the so-called “reflecting set” (to be called ∂S^+ below) for the chain will be an appropriate “discretization” of the reflecting boundary for the original continuous time problem. For the continuous time problem, the reflection is assumed to be “instantaneous.” (Recall the model of the Skorokhod problem in Chapter 1). The transition probability and cost for the approximating chains will “imitate” this behavior. This is the reason that we do not discount the time spent at the so-called reflecting states of the chain. When there is an absorbing or stopping boundary ∂S , as well as a reflection set ∂S^+ , it is always assumed that the two are disjoint. An appropriate modification of the discounted cost function will be defined in the next two paragraphs.

The Discounted Cost Function If There Is a Reflecting Boundary. A *reflecting set* $\partial S^+ \subset S$ is any selected set which satisfies

$$P_x\{\xi_n \in \partial S^+, \text{ all } n < \infty\} = 0, \text{ and } \beta(x) = 0 \text{ for all } x \in \partial S^+. \quad (1.12)$$

The above equation guarantees that that the chain cannot get stuck on the reflecting boundary. Suppose that the cost function $W(x)$ of interest is still (1.11) but with $\beta(x) > 0$ for $x \in S - \partial S^+$. Then the recursive equation for $W(x)$ is

$$W(x) = \begin{cases} e^{-\beta(x)} E_x W(\xi_1) + c(x), & x \in S - \partial S - \partial S^+ \\ g(x), & x \in \partial S \\ E_x W(\xi_1) + c(x), & x \in \partial S^+. \end{cases} \quad (1.13)$$

Equation (1.13) has a unique solution due to the discounting and (1.12).

Let us put the cost equation (1.13) into vector form. Define the discounted substochastic matrix $R = \{r(x, y); x, y \in S - \partial S\}$ by

$$r(x, y) = \begin{cases} e^{-\beta(x)} p(x, y), & x \in S - \partial S - \partial S^+ \\ p(x, y), & x \in \partial S^+. \end{cases} \quad (1.14)$$

Define the cost rate vector C by

$$C(x) = \begin{cases} c(x) + e^{-\beta(x)} \sum_{y \in \partial S} p(x, y)g(y), & x \in S - \partial S - \partial S^+ \\ c(x) + \sum_{y \in \partial S} p(x, y)g(y), & x \in \partial S^+. \end{cases} \quad (1.15)$$

Then (1.10) holds with these new definitions. Because $R^n \xrightarrow{n} 0$, the solution to (1.10) is still unique.

2.1.3 AVERAGE COST PER UNIT TIME

Suppose that the state space S is a single connected aperiodic class: That is, for large n , the n -step transition probabilities satisfy $p^n(x, y) > 0$ for all x, y . Then there is a unique invariant measure $\pi = \{\pi(x), x \in S\}$ which we write as a *row vector*, and which satisfies the equation $\pi = \pi P$ [70]. Define the “stationary” cost value γ and “stationary expectation” operator E_π by

$$\gamma \equiv E_\pi c(\xi) = \sum_{x \in S} \pi(x)c(x). \quad (1.16)$$

By the ergodic theorem for Markov chains [20, 70], for each x

$$\gamma = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n E_x c(\xi_i).$$

There is an auxiliary function $W(\cdot)$ such that the pair $(W(\cdot), \gamma)$ satisfies the equation [10, 74, 100]

$$W(x) + \gamma = E_x W(\xi_1) + c(x). \quad (1.17)$$

In vector form, (1.17) is (W is an $|S|$ -dimensional vector here)

$$W + \mathbf{e}\gamma = PW + C, \quad (1.18)$$

where $\mathbf{e} = (1, \dots, 1)$ is the column vector all of whose components are unity, $P = \{p(x, y); x, y \in S\}$, and $C = \{c(x); x \in S\}$. On the other hand, suppose that (1.18) holds for some pair (W, γ) . Premultiply each side of (1.18) by π and use the fact that $\pi = \pi P$ to get (1.16). One choice for W is

$$\sum_{i=0}^{\infty} P^i(C - \mathbf{e}\gamma),$$

which is well defined because $P^i C \rightarrow \mathbf{e}\gamma$ at a geometric rate under our conditions.

An alternative way of showing that the γ in (1.17) equals the value in (1.16) involves iterating (1.17) n times to get

$$W = P^n W + \sum_{i=0}^{n-1} P^i (C - e\gamma).$$

Now divide by n and let $n \rightarrow \infty$ to get (1.16) again. The function $W(\cdot)$ in (1.18) is not unique. If any vector of the form $k\mathbf{e}$ is added to $W(\cdot)$, where k is a real number, then (1.18) still holds for the new value. We will return to a discussion of the ergodic cost problem for Markov chains in Chapter 7, where additional conditions for the existence of solutions to (1.17) are given.

2.1.4 STOPPING AT A GIVEN TERMINAL TIME

Consider the case where the interest in the chain stops at either a given nonrandom time M or at the first time N that the chain enters a selected stopping set $\partial S \subset S$, whichever comes first. Let $E_{x,n}$ denote expectation of functionals of the chain $\{\xi_i, i \geq n\}$, given that $\xi_n = x$. For given functions $c(\cdot)$ and $g(\cdot)$, let the cost starting at time $n \leq M$ be defined by

$$W(x, n) = E_{x,n} \left[\sum_{i=n}^{(N \wedge M)-1} c(\xi_i) + g(\xi_{N \wedge M}, N \wedge M) \right]. \quad (1.19)$$

Using the Markov property, as done in (1.3) or (1.7), for $n < M$ and $x \in S - \partial S$ we can write (1.19) as

$$\begin{aligned} W(x, n) &= c(x) + E_{x,n} \left\{ E_{x,n} \left[\sum_{i=n+1}^{(N \wedge M)-1} c(\xi_i) + g(\xi_{N \wedge M}, N \wedge M) \middle| \xi_{n+1} \right] \right\} \\ &= E_{x,n} W(\xi_{n+1}, n+1) + c(x), \end{aligned} \quad (1.20)$$

where we use ξ_{n+1} to denote the state at absolute time $n+1$. Note that we are not using the “shifting initial time” terminology which is frequently used (e.g., as in [40]) when working with Markov chains, but rather an absolute time scale. The boundary conditions are

$$W(x, n) = g(x, n), \quad x \in \partial S \text{ or } n = M. \quad (1.21)$$

Define $W(n) = \{W(x, n), x \in S - \partial S\}$. Then in vector form, (1.20) is

$$W(n) = RW(n+1) + C, \quad n < M, \quad (1.22)$$

where R and C are the same as above (1.5) and the terminal boundary condition is $W(M) = \{g(x, M), x \in S - \partial S\}$.

2.2 Optimal Stopping Problems

One of the the simplest control problems is the optimal stopping problem, where the only two possible control actions at any time are to stop the process or to let it continue (if it has not yet been stopped). As in Section 2.1, let $\{\xi_n, n < \infty\}$ be a Markov chain on a finite state space S with time independent transition probabilities $p(x, y)$. The cost of control is the expectation of some function of the path followed until the stopping time. A quite complete development under general conditions appears in [109].

Examples occur in the theory of hypothesis testing, where one wishes to decide which one of several alternative hypotheses is the true one on the basis of a sequence of sample observations. At the end of the sampling, a decision is made concerning which hypothesis is the true one. Under each given hypothesis, it is assumed that the components of the sequence of samples are mutually independent. The basic question concerns the number of samples to take, and the decision to continue sampling will depend on the values of the already available samples. There is a cost assigned to each sample taken, and also a cost associated with an incorrect final decision concerning which of the two hypothesis is the true one. For this example, after each new sample is taken, the conditional probability that each alternative hypothesis is the true one is recomputed, and a decision is made on whether to continue sampling or not. Sampling continues until the additional information obtained concerning the true hypothesis is not worth the additional cost of sampling. Other examples of optimal stopping problems occur in the timing of the buying or selling of assets and in the theory of reliability and maintenance, and additional examples are given in the references on stochastic control.

Definition. Let N denote a random variable with values in the set $[0, \infty]$. N will be the time at which the chain is stopped. If the chain is never stopped for some sample path, then $N = \infty$ on that path. In the sequential sampling problem above, the value $N = \infty$ would occur if sampling never stopped. We say that N is an *admissible stopping time* or simply a *stopping time* for the chain $\{\xi_n, n < \infty\}$ if it “does not depend the future.” More precisely, N is an admissible stopping time if for any $n, m > n$, and function $F(\cdot)$, we have

$$E[F(\xi_i, n < i \leq m) | \xi_i, i \leq n, NI_{\{N \leq n\}}] = E[F(\xi_i, n < i \leq m) | \xi_n].$$

Loosely speaking, for any n , the event $\{N = n\}$ might depend only on the past and current values of the state $\{\xi_i, i \leq n\}$, or it might depend on other quantities as well, as long as the above “Markov property” is preserved.

2.2.1 DISCOUNTED COST

A natural extension of the discounted cost function (1.8) can be associated with an admissible stopping time. In particular, let $\beta > 0$ and define

$$W(x, N) = E_x \left[\sum_{n=0}^{N-1} e^{-\beta n} c(\xi_n) + e^{-\beta N} g(\xi_N) \right], \quad (2.1)$$

for some given functions $c(\cdot)$ and $g(\cdot)$. In Subsection 2.1.2, N was the first entrance time into the a priori selected set ∂S . If the state dependent discount factor $\beta(x) > 0$ is used, then replace (2.1) by

$$W(x, N) = E_x \sum_{n=0}^{N-1} \exp \left[- \sum_{i=0}^{n-1} \beta(\xi_i) \right] c(\xi_n) + E_x \exp \left[- \sum_{i=0}^{N-1} \beta(\xi_i) \right] g(\xi_N). \quad (2.2)$$

Before writing the functional equation for the optimal cost, let us consider the special case where N is a stopping time which is a priori defined to be the first entrance time of the chain into a selected set and use the Markov property to get the functional equation for the associated cost. In particular, suppose that there is a set $S_0 \subset S$ such that $N = \min\{n : \xi_n \in S_0\}$. Then the functional equation (1.9) holds for the cost (2.2). In particular,

$$W(x, N) = \begin{cases} e^{-\beta(x)} E_x W(\xi_1, N) + c(x), & x \in S - S_0 \\ g(x), & x \in S_0. \end{cases} \quad (2.3)$$

We use both notations ∂S and S_0 because ∂S will be a set in the state space which will approximate a “boundary,” whereas S_0 is an unknown set, which is to be determined by solving the optimal stopping problem.

A Recursive Equation for the Optimal Cost. Define the infimum of the costs

$$V(x) = \inf_N W(x, N),$$

where the infimum is over all the admissible stopping times for the chain. Owing to the discounting, $V(x)$ is bounded. We now describe the fundamental procedure for obtaining a functional equation for $V(x)$. Suppose that the current time is n , the current state is $\xi_n = x$, and that the process has not yet been stopped. We need to decide whether to stop (and attain an immediate and final cost $g(x)$) or to continue. If we allow the process to continue, whether or not that decision is optimal, then there is an immediately realized cost $c(x)$ as well as costs to be realized in the future. If we allow the process to continue, then the next state is ξ_{n+1} . Suppose that we continue, but follow an optimal decision policy from the next time on. Then, by the definition of $V(x)$, the optimal discounted cost as seen from the next time is $V(\xi_{n+1})$. Its mean value, discounted to the present, is

$$e^{-\beta(x)} E[V(\xi_{n+1}) | \xi_n = x] = e^{-\beta(x)} \sum_y V(y) p(x, y) = e^{-\beta(x)} E_x V(\xi_1).$$

Thus, if we continue at the current time and then act optimally from the next time on, the total discounted cost, *as seen from the present time*, is $e^{-\beta(x)} E_x V(\xi_1) + c(x)$.

We still need to choose the decision to be taken at the present time. The optimal decision to be taken now is the decision which attains the minimum of the costs over the two possibilities, which are: (1) stop now; (2) continue and then act optimally in the future. Hence, the optimal cost must satisfy the equation

$$V(x) = \min[g(x), e^{-\beta(x)} E_x V(\xi_1) + c(x)]. \quad (2.4)$$

If the two terms in the bracket in (2.4) are equal at some x , then at that x it does not matter whether we stop or continue. The term in (2.4) which is the minimum tells us what the optimal action is: The set $S_0 = \{x : V(x) = g(x)\}$ is known as the *stopping set*. For $x \notin S_0$, (2.4) implies that we should continue; otherwise, we should stop.

The Principle of Optimality. The method just used for the derivation of (2.4) is known as the *principle of optimality*. It is the usual method for getting the functional equations which are satisfied by *optimal value functions* for control problems for Markov process models when the control at any time can depend on the state of the process at that time. It will also be used in a formal way in Chapter 3 to get partial differential equations which are formally satisfied by the value functions for the continuous time problems. Let the current state be x . The principle of optimality basically asserts that *whatever* we do now at state x , in order to get the optimal value function (least cost), we will have to use an *optimal policy* from the next step on, from *whatever* position the system finds itself in at that time. The distribution of the next value of the state of the chain depends on the current control action and state. The total expected cost from the present time on, given the current state and control action, is just the cost attained at the present time plus the future expected cost (with discounting used, if appropriate). Then, the current value of the control must be selected to minimize the sum of these costs. Other examples of the use of the principle of optimality appear in the next section.

Rewriting (2.4) in Terms of a Control. Equation (2.4) is known as the *dynamic programming equation* or *Bellman equation*. For the discounted cost problem of this subsection, it has a unique solution as will be shown below (2.6). It is easier to see this if we rewrite the equation in terms of a controlled transition function and then put it into vector form. The control here is trivial, but the notation allows us to get a convenient expression, and to set the stage for the more complicated problems of the next section. Define the control space $\mathcal{U} = \{0, 1\}$, with generic value α . The control value $\alpha = 0$ is used to denote that we continue the process, and the control value

$\alpha = 1$ is used to denote that we stop the process. Define the controlled transition function $p(x, y|\alpha)$ and cost rate function $c(x, \alpha)$ by

$$p(x, y|\alpha) = \begin{cases} p(x, y), & \alpha = 0 \\ 0, & \alpha = 1, \end{cases}$$

$$c(x, \alpha) = \begin{cases} c(x), & \alpha = 0 \\ g(x), & \alpha = 1. \end{cases}$$

If there is a \mathcal{U} -valued function $u(\cdot)$ on S such that the transition probabilities are $p(x, y|u(x))$, then the control is said to be *pure Markov*. Such a control is a *feedback control* which does not depend on time. For such a control, the stopping set is $S_0 = \{x : u(x) = 1\}$. Let u_n denote the decision at time n . Then $u_n = 0$, for $n < N$, and $u_N = 1$. If there are numbers p_x such that for all n

$$P\{u_n = 1 | \xi_i, i \leq n, u_i = 0, i < n\} = p_{\xi_n},$$

where $p_x \in (0, 1)$ for some x , then the stopping rule or control is said to be *randomized Markov*. If the control or (equivalently) the stopping decision is determined by a pure Markov control or decision function $u(\cdot)$, then we write the associated cost as $W(x, u)$ instead of as $W(x, N)$.

We next write the expression for the minimum cost in vector notation. For a feedback control $u(\cdot)$, define the cost rate vector $C(u) = \{c(x, u(x)), x \in S\}$, the controlled discounted transition function

$$r(x, y|u(x)) = e^{-\beta(x)} p(x, y|u(x)),$$

and the degenerate transition matrix $R(u) = \{r(x, y|u(x)); x, y \in S\}$. Let $W(u) = \{W(x, u), x \in S\}$ denote the vector of total discounted costs (2.2) under the feedback control $u(\cdot)$, and let $V = \{V(x), x \in S\}$ denote the vector of least costs. Then we can write (2.4) in the following $|S|$ -dimensional vector form:

$$V = \min_{u(x) \in \mathcal{U}} [R(u)V + C(u)]. \quad (2.5)$$

In (2.5) and in future minimizations of vector valued functions, the minimum is taken line by line. That is, in the x^{th} line of (2.5), we minimize over all the values of $u(x) \in \mathcal{U}$.

Existence of a Solution to (2.5) and an Optimal Feedback Control. Let $\bar{u}(\cdot)$ denote some particular control at which the minimum in (2.5) is taken on. We will show that it is an optimal feedback control. Letting $u(\cdot)$ be any other feedback control, the minimizing operation in (2.5) yields

$$V = R(\bar{u})V + C(\bar{u}) \leq R(u)V + C(u), \quad (2.6)$$

where the inequality is *component by component*. Iterating (2.6) and using the fact that (due to the discounting)

$$R^n(\bar{u}) \rightarrow 0, \quad R^n(u) \rightarrow 0, \quad (2.7)$$

we have

$$V = \sum_{n=0}^{\infty} R^n(\bar{u})C(\bar{u}) = W(\bar{u}) \leq \sum_{n=0}^{\infty} R^n(u)C(u) = W(u), \quad (2.8)$$

where the inequality is component by component. Thus, V is indeed the minimal (over all feedback controls) cost vector, $\bar{u}(\cdot)$ is an optimal decision or control function, and the solution to (2.5) is unique because the minimal cost is unique. Via a similar argument, it can be seen that for any V_0 , the iteration

$$V_{n+1} = \min_{u(x) \in \mathcal{U}} [R(u)V_n + C(u)]$$

converges to a solution to (2.5). The initial condition is irrelevant due to the discounting.

A Note on the “Contraction” (2.7). Note the critical role that (2.7) played. If the n -step discounted transition probabilities did not go to zero, we could not have obtained the uniqueness and possibly not even the interpretation of the solution to (2.5) as the minimum cost vector. A similar property is needed for the general control problem with a Markov chain model and will be dealt with in Section 2.4 in more detail. Such properties are also very useful in dealing with the convergence proofs for numerical methods for solving equations such as (2.5). A proof of convergence in a setting where the contraction property is absent is given in Subsection 13.3.3.

Including an Obligatory Stopping Set. There is a variation of the optimal stopping problem where, irrespective of wherever else we might decide to stop, we *must* stop on first reaching some a priori selected set $\partial S \in S$, and where we attain a cost of $g(x)$ if we first enter ∂S at the point x . Then (2.4) still holds for $x \in S - \partial S$, but we have the boundary condition $V(x) = g(x)$, $x \in \partial S$.

Remark. The above discussion showed only that the minimizing control in (2.4) or (2.5) is the optimal control/stopping decision function with respect to the class of comparison stopping rules which are also determined by first entrance times into sets in the state space or, equivalently, by feedback control laws, and that $V(x)$ is the least cost only in this class. But it is also true that the optimality of $\bar{u}(\cdot)$ in this class implies optimality with respect to all admissible stopping times. The proof is essentially the same as used above for the pure Markov rules, except that the alternative $u(\xi_n)$

are replaced by a sequence of appropriate “admissible” decision variables $\{u_n\}$ and the details are omitted [74, 110].

2.2.2 UNDISCOUNTED COST

In the absence of discounting [i.e., $\beta(x) = 0$], (2.2) need not have a well defined meaning without some additional conditions. The essential condition used in the arguments of the last subsection was (2.7), where $R(u)$ was the “effective” transition matrix for the controlled chain. The condition (2.7) held due to the discounting, irrespective of the values of $p(x, y)$ or of the choice of control function.

Alternative Conditions. Define $c_0 = \min_x c(x)$, and suppose that $c_0 > 0$. Then we need only consider stopping times N which satisfy

$$E_x N \leq \frac{2 \sup |g(y)|}{c_0}. \quad (2.9)$$

To see this, suppose that N is an admissible stopping time which violates (2.9). Then $c_0 E_x N > 2 \sup_y |g(y)|$ and the associated cost satisfies

$$W(x, N) \geq c_0 E_x N + E_x g(\xi_N) > 2 \sup |g(y)| + E_x g(\xi_N) > g(x).$$

Thus, it would have been preferable to stop at the initial time rather than at N . Hence, we can assume (2.9). This implies that $V(x)$ is finite. Also, $R^n(u) \rightarrow 0$ for all pure Markov $u(\cdot)$ for which (2.9) holds. By this result and the principle of optimality, $V(\cdot)$ satisfies

$$V(x) = \min[g(x), E_x V(\xi_1) + c(x)]. \quad (2.10)$$

An obligatory stopping set can also be added and the comments made in the paragraph just above Subsection 2.2.2 hold for this case too.

2.3 Discounted Cost

Terminology. In this section, we treat the general discounted cost problem for a Markov chain model. The principle of optimality introduced in the last section will be used to obtain a functional equation for the minimal cost. Let \mathcal{U} , the control action set, be a compact set in some topological space, with generic variable α . The actual space is unimportant, although in applications it is generally either a finite set or a subspace of a Euclidean space. We say that $\{\xi_n, n < \infty\}$ is a *controlled Markov chain* on the finite state space S if the transition probabilities are functions of a control variable which takes values in the set \mathcal{U} . We use α for the generic value of this variable. The “controlled” transition probabilities will be written as

$p(x, y|\alpha)$, where α will usually depend on x . For reasons of notational simplicity, the controlled transition probabilities will not depend explicitly on time. At any time n , the control action is a random variable, which we denote by u_n . Let $u = (u_0, u_1, \dots)$ denote the sequence of \mathcal{U} -valued random variables which are the control actions at times $0, 1, \dots$. We say that u is *admissible* if the Markov property continues to hold under use of the sequence u , namely, that

$$P\{\xi_{n+1} = y|\xi_i, u_i, i \leq n\} = P\{\xi_{n+1} = y|\xi_n, u_n\} = p(\xi_n, y|u_n).$$

Let E_x^α denote the expectation of functionals of ξ_1 given $\xi_0 = x, u_0 = \alpha$. For an admissible control sequence u , let E_x^u denote the expectation of functionals of the chain $\{\xi_n, n < \infty\}$ given that the initial condition is x and which uses the transition probabilities $p(z, y|u_n)$ at time n if $\xi_n = z$. There is an abuse of notation concerning the use of u which is convenient and which should not cause any confusion. If there is a function $u(\cdot)$ such that $u_n = u(\xi_n)$, then we refer to the control as a *feedback* or *pure Markov* policy and use the notation u to refer to both the function and the sequence of actual control actions.

In the last section, we allowed the discount factor to depend on the state. Throughout the rest of the chapter, we allow it to depend on both the state and control, and suppose that $\beta(\cdot)$ is a non-negative function of x and α and is continuous in α for each x . We will also suppose that the cost function $c(x, \cdot)$ and the transition probability $p(x, y|\cdot)$ are continuous functions of the control parameter $\alpha \in \mathcal{U}$ for each x, y . Now, suppose that $\beta(x, \alpha) > 0$ for each x, α . For an admissible control sequence u , define the cost

$$W(x, u) = E_x^u \sum_{n=0}^{\infty} \exp \left[- \sum_{i=0}^{n-1} \beta(\xi_i, u_i) \right] c(\xi_n, u_n). \quad (3.1)$$

The modifications which are needed if we are obliged to stop when first reaching a selected “boundary” set ∂S will be stated below.

Let $V(x)$ denote the infimum of the costs $W(x, u)$ over all admissible control sequences. $V(x)$ is finite for each x due to the discounting and satisfies the dynamic programming equation

$$V(x) = \min_{\alpha \in \mathcal{U}} \left[e^{-\beta(x, \alpha)} E_x^\alpha V(\xi_1) + c(x, \alpha) \right]. \quad (3.2)$$

Equation (3.2) is easy to derive via use of the principle of optimality, as follows: Given the current state x , we use a control action α . This value of α must be chosen in an optimal way. However, for any choice of α , there is a current “running” cost of $c(x, \alpha)$. Suppose that from the next step on, we choose the control value in an optimal manner. Then whatever the next state and current control action, the expected value (as seen from the present state and time) of the discounted *future* cost is $e^{-\beta(x, \alpha)} E_x^\alpha V(\xi_1)$. The total discounted cost as seen from the present time, and given that the

current control action is α , is $c(x, \alpha) + e^{-\beta(x, \alpha)} E_x^\alpha V(\xi_1)$. Now we need to choose the minimizing value of α , from which (3.2) follows.

There is a function $\bar{u}(x)$ which attains the minimum in (3.2). Reasoning as in the paragraph below (2.4) shows that the solution to (3.2) is unique and that $\bar{u}(\cdot)$ is indeed an optimal control with respect to all admissible control sequences. We will repeat some of the details. The argument to follow is essentially a verification of the principle of optimality. Let $u(\cdot)$ be any feedback control. Define the discounted transition matrix $R(u) = \{e^{-\beta(x, u(x))} p(x, y|u(x)); x, y \in S\}$. Define the cost rate vector $C(u) = \{c(x, u(x)), x \in S\}$, and define the cost vector $W(u)$ and minimum cost vector V as done in connection with (2.4). Then the vector version of (3.2) is

$$V = \min_{u(x) \in \mathcal{U}} [R(u)V + C(u)]. \quad (3.3)$$

Then, as in Section 2.2, (3.3) implies (where the inequality is for each component)

$$V = R(\bar{u})V + C(\bar{u}) \leq R(u)V + C(u), \quad (3.4)$$

where $\bar{u}(x)$ is the minimizing control in (3.2) or in the x^{th} line of (3.3) and $u(\cdot)$ is any other feedback control. As in the last section, the discounting implies (2.7) and that (2.8) holds, from which follows both the uniqueness of the solution to (3.3) and the fact that the solution is the minimum value function over all feedback control alternatives. A similar proof shows that that $\bar{u}(\cdot)$ is optimal with respect to all admissible control sequences.

Adding a Stopping Set. Suppose that there is a chosen “boundary” set ∂S such that we must stop on first contact with it, and with stopping cost $g(x)$ if first contact is at x . Let N be the first time of contact. Then (3.1) is modified as:

$$\begin{aligned} W(x, u) &= E_x^u \sum_{n=0}^{N-1} \exp \left[- \sum_{i=0}^{n-1} \beta(\xi_i, u_i) \right] c(\xi_n, u_n) \\ &\quad + E_x^u \exp \left[- \sum_{i=0}^{N-1} \beta(\xi_i, u_i) \right] g(\xi_N). \end{aligned} \quad (3.5)$$

Then the dynamic programming equation is (3.2) for $x \notin \partial S$, and with the boundary condition $V(x) = g(x)$, $x \in \partial S$. All of the previous results hold here also, and a “reflecting boundary” ∂S^+ can be added as in Section 2.1.

2.4 Control to a Target Set and Contraction Mappings

If $\beta(x, \alpha) > 0$, then the discount factor $e^{-\beta(x, \alpha)}$ guarantees that (3.2) and (3.3) have a unique solution which is the minimum value function. In the

absence of discounting, one needs to be more careful. Property (2.7) is required for the candidate $\bar{u}(\cdot)$ for the optimal feedback control and for appropriate comparison controls $u(\cdot)$, and some condition which guarantees this is needed. We will describe the general requirement. Let $\{\xi_n, n < \infty\}$ be a controlled Markov chain as in the last section. Let $\partial S \in S$ be a selected stopping or boundary set, so that we are obliged to stop the control process at time N , the moment of first entrance into ∂S . For an admissible control sequence u , define the undiscounted cost

$$W(x, u) = E_x^u \sum_{n=0}^{N-1} c(\xi_n, u_n) + E_x^u g(\xi_N). \quad (4.1)$$

Let us define the optimal cost $V(x) = \inf_u W(x, u)$, where the infimum is over all admissible control sequences. In the absence of discounting, neither the function $W(x, u)$ in (4.1) nor the $V(x)$ need be well defined or finite. Let us proceed formally for the moment so that we can see what is required. By a formal use of the principle of optimality, the dynamic programming equation for the cost function (4.1) is

$$V(x) = \begin{cases} \inf_{\alpha \in \mathcal{U}} [E_x^\alpha V(\xi_1) + c(x, \alpha)], & x \in S - \partial S \\ g(x), & x \in \partial S. \end{cases} \quad (4.2)$$

Next, working as in the previous sections, we put (4.2) into vector form. For a feedback control $u(\cdot)$, define the reduced transition matrix $R(u) = \{r(x, y|u(x)), x, y \in S - \partial S\}$, where

$$r(x, y|\alpha) = p(x, y|\alpha), \quad x \in S - \partial S. \quad (4.3)$$

Define the cost rate vector $C(u)$ with components

$$C(x, u) = c(x, u(x)) + \sum_{y \in \partial S} p(x, y|u(x))g(y), \quad x \in S - \partial S.$$

Then, defining $V = \{V(x), x \in S - \partial S\}$, we can write (4.2) as

$$V = \min_{u(x) \in \mathcal{U}} [R(u)V + C(u)]. \quad (4.4)$$

In (4.4), the minimization is over each component separately, as usual.

Definition. Let R be a square matrix. We say that R is a *contraction* if $R^n \xrightarrow{n} 0$.

Note that the $R(u)$ which appeared in (2.5) and (3.3) are contractions for all feedback $u(\cdot)$. If $R(u)$ is a contraction for each feedback control $u(\cdot)$, then the arguments of the previous two sections can be repeated to show that the solution to (4.2) or (4.4) is unique, that it is the optimal cost, and

that the minimizing feedback control is an optimal control with respect to all admissible controls.

A Sufficient Condition for a Contraction, and Discussion. In the applications of interest in this book, it is often the case that $R(u)$ is a contraction for each feedback control $u(\cdot)$. A useful and frequently used alternative condition is the following:

A4.1. *There are $c_0 > 0$, such that $c(x, \alpha) \geq c_0$, and a feedback control $u_0(\cdot)$ such that $R(u_0)$ is a contraction.*

In this case, the positivity of the running cost rate $c(\cdot)$ implies that $W(x, u)$ might be unbounded for some controls. But the existence of $u_0(\cdot)$ in (A4.1) guarantees that there is a feedback control $\bar{u}(\cdot)$ which is optimal with respect to all admissible controls and such that $R(\bar{u})$ is a contraction and that (4.4) holds. The proofs of such assertions can be found in many places in the literature on control problems with Markov chain models [10, 74, 100, 104].

It is important to recall that $R(u)$, being a contraction, is equivalent to $E_x^u N < \infty$ for all x because the state space is finite, where N is defined above (4.1). This latter interpretation is useful because it can often be checked by inspection of the transition matrices, for particular controls $u(\cdot)$. In particular, for each $x \notin \partial S$ there needs to be a chain of states of positive probability under $u(\cdot)$ and which leads to some state in ∂S .

The concept of contraction plays a fundamental role in the proofs of convergence of the numerical methods used in Chapter 6.

2.5 Finite Time Control Problems

The controlled analogue of the cost function (1.19) in Subsection 2.1.4 is

$$W(x, u, n) = E_{x,n}^u \left[\sum_{i=n}^{(N \wedge M)-1} c(\xi_i, u_i) + g(\xi_{N \wedge M}, N \wedge M) \right], \quad (5.1)$$

where we use the notation $E_{x,n}^u$ to denote the expectation given that $\xi_n = x$ and that the control sequence $\{u_n, \dots\}$ is used. As in Subsection 2.1.4, $N = \min\{n : \xi_n \in \partial S\}$. Letting $V(x, n)$ denote the infimum of $W(x, n, u)$ over all admissible controls and using the principle of optimality, we get that the dynamic programming equation is

$$V(x, n) = \min_{\alpha \in \mathcal{U}} [E_{x,n}^\alpha V(\xi_{n+1}, n+1) + c(x, \alpha)], \quad (5.2)$$

for $x \in S - \partial S$, $n < M$. The boundary conditions are the same as in (1.21), namely, $W(x, n) = g(x, n)$ for $x \in \partial S$ or $n = M$.

3

Dynamic Programming Equations

In this chapter we define many of the standard control problems whose numerical solutions will concern us in the subsequent chapters. Other, less familiar control problems will be discussed separately in later chapters. We will first define cost functionals for uncontrolled processes, and then formally discuss the partial differential equations which they satisfy. Then the cost functionals for the controlled problems will be stated and the partial differential equations for the optimal cost formally derived. These partial differential equations are generally known as Bellman equations or dynamic programming equations. The main tool in the derivations is Itô's formula.

It should be noted that not only are our derivations of the equations formal, but, in general, the equations themselves have only a formal meaning, and little is known concerning existence, uniqueness, etc. (An exception to this is the so-called “viscosity solution” method of Chapter 14.) One of the reasons we present the equations is because their forms are very suggestive of good numerical methods. However, when we deal with the convergence of the numerical methods, our aim will be to show that the value functions to which our approximations converge are optimal value functions, and we do this by identifying an optimal controlled process with that value function. Thus, the Bellman equations never enter into the convergence proof, and all the analysis is carried out in terms of the cost functionals. It is not necessary that the reader be familiar with the PDE’s to understand or use the algorithms or to understand the convergence proofs.

As remarked above, our interest in the formally derived PDE, boundary conditions, and so on, is mainly due to the fact that they suggest useful numerical schemes. With this in mind it makes little sense to maintain any pretense regarding mathematical rigor in these derivations. The “validation” for any derived equation will come in the form of a rigorous convergence proof for numerical schemes suggested by this equation. Thus, the formal derivations themselves are not used in any direct way. Our motivation for including them is to provide a guide to similar formal derivations that might be useful for less standard or more novel stochastic control problems. For a more rigorous development of the dynamic programming equations we refer the reader to [46] and [47].

3.1 Functionals of Uncontrolled Processes

In this section we will work with the uncontrolled processes

$$dx = b(x)dt + \sigma(x)dw, \quad (1.1)$$

$$dx = b(x)dt + \sigma(x)dw + dJ, \quad (1.2)$$

where the jump term J has jump rate $\lambda(x)$ and jump distribution $\bar{\Pi}(x, dy)$. Processes (1.1) and (1.2) were discussed in Sections 1.3 and 1.5, respectively. For the finite time problem, it is often of interest to introduce a time variable explicitly and rewrite (1.1) as

$$dx = b(x, t)dt + \sigma(x, t)dw \quad (1.3)$$

and with a corresponding change in (1.2). Time variations can be included for all the cost functionals, except for the average cost per unit time case. The reader should have no difficulty in filling in the details, which we leave out in order to keep the development within reasonable bounds. The functions $k(\cdot)$, $g(\cdot)$, $b(\cdot)$, $\sigma(\cdot)$ are all assumed to be bounded and continuous. To begin, we consider several examples (Subsections 3.1.1 to 3.1.5) involving diffusions with no jump term.

3.1.1 COST UNTIL A TARGET SET IS REACHED

Let G be a compact set with a smooth boundary ∂G . Let G^0 denote the interior of G , and suppose that G is the closure of its interior. Define the stopping time τ by $\tau = \min\{t : x(t) \notin G^0\}$. As is customary, we make the convention that if a stopping time is not defined, then its value is set to infinity. Thus, if $x(t) \in G^0$ for all $t < \infty$, then $\tau = \infty$. Suppose that for all $x \in G$ we have $E_x \tau < \infty$. [A discussion of elementary conditions on the process components that are sufficient to ensure this property can be found in [69, Lemma 5.7.2]. For example, it is sufficient to assume there is i such that $a_{ii}(x) > 0$ for all x .] Define the cost functional

$$W(x) = E_x \left[\int_0^\tau k(x(s))ds + g(x(\tau)) \right].$$

Because the process is stopped upon hitting ∂G , we refer to the associated boundary condition as *absorbing*. From a formal point of view, $W(\cdot)$ satisfies the equation

$$\mathcal{L}W(x) + k(x) = 0, \quad x \in G^0, \quad (1.4)$$

where \mathcal{L} is the differential operator of (1.1). A formal derivation of (1.4) is as follows. Under broad conditions we have $P_x\{\tau \leq \Delta\}/\Delta \rightarrow 0$ as $\Delta \rightarrow 0$ for $x \in G^0$. Suppose that $W(\cdot)$ is bounded and in $C^2(G^0)$. For $\Delta > 0$ we

may write

$$\begin{aligned} W(x) &= E_x \left\{ \int_0^{\Delta \wedge \tau} k(x(s))ds + g(x(\tau)) + \left[\int_{\Delta}^{\tau} k(x(s))ds \right] I_{\{\tau > \Delta\}} \right\} \\ &= E_x \left\{ \int_0^{\Delta \wedge \tau} k(x(s))ds + g(x(\tau))I_{\{\tau \leq \Delta\}} + W(x(\Delta))I_{\{\tau > \Delta\}} \right\}, \end{aligned}$$

where the second inequality follows from the Markov property and the definition of $W(\cdot)$. It follows that

$$\frac{1}{\Delta} E_x \left[W(x(\Delta)) - W(x) + \int_0^{\Delta} k(x(s))ds \right] = \frac{1}{\Delta} E_x h(\tau, \Delta) I_{\{\tau \leq \Delta\}}, \quad (1.5)$$

where

$$h(\tau, \Delta) = W(x(\Delta)) - \int_{\tau \wedge \Delta}^{\Delta} k(x(s))ds - g(x(\tau))$$

is bounded uniformly in ω and Δ . Therefore, under the condition $P_x\{\tau \leq \Delta\}/\Delta \rightarrow 0$ as $\Delta \rightarrow 0$, the right hand side of the last equation tends to zero as $\Delta \rightarrow 0$. Applying Itô's formula to the left hand side of (1.5) and sending $\Delta \rightarrow 0$, we formally obtain (1.4).

The boundary conditions are not so obvious, since not all the points on the boundary are necessarily reachable by the process. Only a few comments will be made on this point. We define a *regular point* of the ∂G to be any point $x \in \partial G$ such that, for all $\delta > 0$,

$$\lim_{y \rightarrow x, y \in G^0} P_y\{\tau > \delta\} = 0.$$

Thus, a regular point is a point such that if the process starts nearby, then exit is virtually assured in an arbitrarily small time and in an arbitrarily small neighborhood. The reader is referred to [69] for further discussion of the distinction between regular points and those points which are not regular.

Now suppose that $x \in \partial G$ is regular. Then $E_y \int_0^{\tau} k(x(s))ds \rightarrow 0$ as $y \rightarrow x$. Furthermore, the continuity of g implies that $E_y g(x(\tau)) \rightarrow g(x)$ as $y \rightarrow x$. Combining, we have $W(y) \rightarrow g(x)$ as $y \rightarrow x$. Thus, the correct boundary condition is $W(x) = g(x)$ for regular points x of ∂G .

The Verification Theorem. Suppose that (1.4) holds for $W(\cdot)$ bounded and smooth inside G , and that the probability that $x(\cdot)$ exits G^0 through the set of regular points is unity for each initial condition $x \in G^0$. Then by Itô's formula, for $t < \infty$ we have

$$W(x(t \wedge \tau)) = W(x) + \int_0^{t \wedge \tau} \mathcal{L}W(x(s))ds + \int_0^{t \wedge \tau} W'_x(x(s))\sigma(x(s))dw(s),$$

which by (1.4) implies

$$E_x W(x(t \wedge \tau)) = W(x) - E_x \int_0^{t \wedge \tau} k(x(s))ds.$$

Now let $t \rightarrow \infty$, and recall our assumption $E_x \tau < \infty$. Then $W(x(t \wedge \tau)) \rightarrow W(x(\tau)) = g(x(\tau))$ and $\int_0^{t \wedge \tau} k(x(s))ds \rightarrow \int_0^\tau k(x(s))ds$. By rearranging terms and applying the dominated convergence theorem, we obtain

$$W(x) = E_x \left[\int_0^\tau k(x(s))ds + g(x(\tau)) \right],$$

which proves that $W(x)$ is the cost.

3.1.2 THE DISCOUNTED COST

For $\beta > 0$, define the discounted cost function with no boundary conditions:

$$W(x) = E_x \int_0^\infty e^{-\beta s} k(x(s))ds. \quad (1.6)$$

If $W(x)$ is sufficiently smooth, then one can easily show it satisfies the equation

$$\mathcal{L}W(x) - \beta W(x) + k(x) = 0, \quad x \in \mathbb{R}^k. \quad (1.7)$$

The idea is essentially as follows. Following the argument of the previous subsection, for $\Delta > 0$ we write

$$\begin{aligned} W(x) &= E_x \int_0^\Delta e^{-\beta t} k(x(t))dt + E_x \int_\Delta^\infty e^{-\beta t} k(x(t))dt \\ &= E_x \int_0^\Delta e^{-\beta t} k(x(t))dt \\ &\quad + E_x e^{-\beta \Delta} \left\{ E_{x(\Delta)} \left[\int_\Delta^\infty e^{-\beta(t-\Delta)} k(x(t))dt \right] \right\} \\ &= E_x \int_0^\Delta e^{-\beta t} k(x(t))dt + E_x e^{-\beta \Delta} W(x(\Delta)). \end{aligned}$$

From this we get

$$\frac{1}{\Delta} E_x \left[e^{-\beta \Delta} W(x(\Delta)) - W(x) + \int_0^\Delta e^{-\beta t} k(x(t))dt \right] = 0,$$

or

$$\begin{aligned} 0 &= E_x \frac{W(x(\Delta)) - W(x)}{\Delta} + \frac{e^{-\beta \Delta} - 1}{\Delta} E_x W(x(\Delta)) \\ &\quad + E_x \frac{1}{\Delta} \int_0^\Delta e^{-\beta t} k(x(t))dt. \end{aligned}$$

Using Itô's formula and then sending $\Delta \rightarrow 0$ formally yields (1.7). An argument analogous to that of the Verification Theorem of Subsection 3.1.1 shows that if $W(\cdot)$ is in $C^2(\mathbb{R}^k)$ and is bounded and (1.7) holds, then it is actually the cost (1.6).

A target set G can be introduced as in the case of Subsection 3.1.1. Defining G and τ as in that subsection and using

$$W(x) = E_x \int_0^\tau e^{-\beta s} k(x(s)) ds + E_x e^{-\beta \tau} g(x(\tau)), \quad (1.8)$$

we have formally that $W(\cdot)$ satisfies (1.7) for $x \in G^0$, while $W(x) = g(x)$ holds at regular points of the boundary.

3.1.3 A REFLECTING BOUNDARY

We now consider the problem where the process is reflected rather than absorbed on hitting the boundary of the set G . A special case will be dealt with and even that will be treated in a casual manner. Suppose that the boundary ∂G is smooth; in particular, that it is continuously differentiable. Let $n(x)$ denote the outward normal to the boundary at the point x . The reflection direction is denoted by $r(x)$, where $r : \mathbb{R}^k \rightarrow \mathbb{R}^k$ is a continuous function. The reflection will be “instantaneous” and will be strictly inward in the sense that there is a $\delta > 0$ such that $n(x)'r(x) \leq -\delta$ on ∂G . Our process model is the stochastic differential equation with reflection of Section 1.4 with the domain and boundary data as above. See Example 1.4.3.

Any of the cost functionals used above can be adapted to the reflected process. If the cost functional is (1.6), then the equation satisfied by the cost functional is (1.7) with the boundary condition

$$W_x(x)'r(x) = 0 \quad (1.9)$$

on an appropriate subset of ∂G .

We can give the following intuitive interpretation of the boundary condition. Using the Skorokhod Problem formulation of the reflected process, it is possible to show that for the purpose of approximately computing (1.6) the reflected processes can be replaced by a process constructed as follows. Since the reflection is to be “instantaneous,” we can consider a process $x(\cdot)$ such that upon hitting the boundary at the point x , $x(\cdot)$ is instantly projected a distance $r(x)\delta$, where $\delta > 0$ is small. Thus, the cost functional at x and that at $x + r(x)\delta$ are the same for the approximation to the reflected process, which suggests (1.9).

Part of the boundary can be absorbing and part reflecting. In order to contain the discussion, we describe only the special case where there is an outer boundary that is reflecting and an inner boundary that is absorbing. Let G_1 and G be compact sets, each the closure of its interior, and with $G_1 \subset G^0$. The process will be reflected instantaneously to the interior of G

on hitting its boundary, and it will be absorbed on hitting the boundary of G_1 . Define $\tau = \inf\{t : x(t) \in G_1\}$, and consider the cost functional (1.8). Then $W(\cdot)$ formally satisfies (1.7) but with the boundary condition $W(x) = g(x)$ on the boundary of G_1 and $W_x(x)'r(x) = 0$ on the boundary of G .

3.1.4 THE AVERAGE COST PER UNIT TIME

When systems operate over long time periods, an appropriate cost functional is the average cost per unit time. As in the previous subsections, we work formally. Suppose that the limit

$$\gamma = \lim_{t \rightarrow \infty} \frac{E_x \int_0^t k(x(s))ds}{t} \quad (1.10)$$

exists. One can formally derive a PDE which yields γ , but we only discuss a verification theorem. Suppose that there is a smooth function $W(\cdot)$ and a constant γ which satisfy the equation

$$\mathcal{L}W(x) = \gamma - k(x). \quad (1.11)$$

Using relation (1.11) and Itô's formula yields

$$E_x W(x(t)) - W(x) = E_x \int_0^t [\gamma - k(x(s))] ds.$$

If

$$\frac{E_x W(x(t))}{t} \rightarrow 0,$$

then γ satisfies (1.10).

For numerical purposes it is generally necessary to work with compact state spaces, and if the problem is not given a priori on a compact set, then the state space must be “compactified” just so that a numerical solution can be obtained. A typical way of doing this is to introduce a reflecting boundary. There are many methods that are appropriate and one can even have more complex boundary processes (such as “sticky” boundaries, or movement along the boundary). One tries to choose a boundary condition such that the essential features of the original problem are preserved, and the modified problem can be solved or well approximated numerically. For the ergodic cost problem, we can introduce a reflecting boundary G as in the last subsection. Then (1.11) holds inside G and the boundary condition (1.9) holds on ∂G .

3.1.5 THE COST OVER A FIXED FINITE TIME INTERVAL

Here the process is of interest over the finite interval $[0, T]$ only, and the process stops on hitting the boundary of the region G . Let $g(\cdot)$ be a continuous and bounded function on $\mathbb{R}^k \times [0, T]$. The process will stop if it

hits the boundary of G before time T . If this occurs at time $s < T$ and the exit point is x , then the penalty will be $g(x, s)$. If the process does not exit G before T , then the cost will be $g(x(T), T)$. It is, therefore, natural to set the problem up with a terminal value imposed at time $t = T$. The problem may be rewritten as an initial value problem if desired. Then the cost, starting at point $x \in G^0$ at time $t \leq T$, is

$$W(x, T) = E_{x,t} \left[\int_t^{T \wedge \tau} k(x(s)) ds + g(x(T \wedge \tau), T \wedge \tau) \right].$$

From a formal point of view, it can be shown that $W(\cdot)$ satisfies the PDE

$$\frac{\partial W(x, t)}{\partial t} + \mathcal{L}W(x, t) + k(x) = 0$$

for $x \in G^0, t < T$, together with $W(x, T) = g(x, T)$. We also have $W(y, t) \rightarrow g(x, t)$ as $y \rightarrow x \in \partial G$ for regular points x and

$$E_{x,t}W(x(\tau \wedge T), \tau \wedge T) = E_{x,t}g(x(\tau \wedge T), \tau \wedge T).$$

A reflecting boundary and discounting can be readily added.

3.1.6 A JUMP DIFFUSION EXAMPLE

We consider the setup of Subsection 3.1.1 for the case of a diffusion with jumps. Recall that G is assumed to be a compact set with smooth boundary and the property that G is the closure of its interior, and that $\tau = \inf\{t : x(t) \in \mathbb{R}^k \setminus G^0\}$. Let $x(\cdot)$ satisfy (1.2), with the specified jump rate and distribution. Once again, define the cost by

$$W(x) = E_x \left[\int_0^\tau k(x(s)) ds + g(x(\tau)) \right].$$

We can then formally derive (exactly as in Subsection 3.1.1) the equation

$$\mathcal{L}W(x) + k(x) = 0, \quad x \in G^0,$$

where \mathcal{L} is defined by (1.5.5).

For the jump diffusion, we obviously need to specify more than simply a boundary condition because since at the time of exit the process may be far removed from the set G . The condition becomes $W(x) = g(x)$ for $x \notin G^0$.

3.2 The Optimal Stopping Problem

Perhaps the simplest stochastic control problem is the optimal stopping problem, where the only action to be taken is to decide when to stop

the process. The model is the diffusion (1.1) or jump diffusion (1.2), and \mathcal{F}_t denotes the underlying filtration. We would like to consider rules for stopping that are suitably nonanticipative. For later use as well as for use here, we give the following definition.

Definition 2.1. *We say that a non-negative random variable τ is an admissible stopping time if it is a \mathcal{F}_t -stopping time. The stopping time is said to be pure Markov if there is a Borel B set in the state space such that $\tau = \inf\{t : x(t) \in B\}$.*

One can show that the pure Markov stopping times are admissible. (Actually, for the pure Markov stopping times to be admissible for all Borel sets B requires an additional technical assumption on the filtration, namely, right continuity [42]. In keeping with the spirit of this chapter we will not worry about this issue.) Let $k(\cdot)$ and $g(\cdot)$ be bounded and continuous real valued functions, with $\inf k(x) \geq k_0 > 0$. For an admissible stopping time, define the cost

$$W(x, \tau) = E_x \left[\int_0^\tau k(x(t)) dt + g(x(\tau)) \right] \quad (2.1)$$

and the optimal cost, where the infimum is taken over all admissible stopping times:

$$V(x) = \inf_{\tau} W(x, \tau).$$

It is well known that the infimum is the same if it is taken over pure Markov stopping times. Indeed, it is the optimal pure Markov stopping set that we seek with the numerical methods. A further simplification, owing to the positive lower bound on k , is that we need to consider only stopping times whose mean value satisfies

$$E_x \tau \leq \frac{2 \sup_x |g(x)|}{k_0}.$$

This follows because if the mean value is larger, we will do better to stop at $t = 0$.

Formal Derivation of the Bellman Equation. Let B denote the optimal stopping set; i.e., the process stops when the set B is reached or entered for the first time. Then $V(x) \leq g(x)$ and $V(x) = g(x)$ only on B . The equation satisfied by the optimal cost is

$$\begin{cases} \mathcal{L}V(x) + k(x) = 0, & x \notin B \\ V(x) = g(x), & x \in B, \end{cases} \quad (2.2)$$

where the set B is part of the solution. We will give a simple derivation. The derivation assumes familiarity with the principle of optimality, which is discussed in Chapter 2 for the optimal stopping problem. Suppose that

we restrict the times to be multiples of a small $\Delta > 0$. Then at each time $n\Delta$, we have the choice of stopping or continuing. Given the current state, the additional cost that we pay for immediate stopping is $g(x(n\Delta))$. Heuristically, the usual dynamic programming argument tells us that the additional cost paid for continuing and using the optimal decisions in all future steps is $E_{x(n\Delta)}[V(x((n+1)\Delta)) + \Delta k(x(n\Delta))]$. Thus, heuristically,

$$V(x) = \min [g(x), E_x V(x(\Delta)) + \Delta k(x)].$$

Now subtract $V(x)$ from both sides:

$$\min [g(x) - V(x), E_x V(x(\Delta)) - V(x) + \Delta k(x)] = 0.$$

Recalling that for $x \notin B$ we have $V(x) < g(x)$, we see that for $x \notin B$ the minimum must be taken on by the second term. If we divide by Δ , we obtain

$$\frac{1}{\Delta} [E_x V(x(\Delta)) - V(x)] + k(x) = 0.$$

If we apply Itô's formula, use the assumption that $V(\cdot)$ is smooth, and then send Δ to zero, the result follows. Depending on whether the boundary is reflecting or absorbing, the appropriate boundary conditions are added.

3.3 Control Until a Target Set Is Reached

The model will be the continuously controlled versions of (1.1) and (1.2):

$$dx = b(x, u)dt + \sigma(x)dw, \quad (3.1)$$

$$dx = b(x, u)dt + \sigma(x)dw + dJ, \quad (3.2)$$

where the jump term J is uncontrolled and satisfies the conditions in Section 1.5. Let us start by considering just (3.1). We will assume that the control takes values in a compact set \mathcal{U} . Recall that the control is *admissible* if it is a \mathcal{F}_t -adapted, measurable and \mathcal{U} -valued process. If the control can be written as a function of the current state and time, then we say that it is a *pure Markov control*. In this section we will consider pure Markov controls that depend only on the state. Define the covariance matrix $a(\cdot)$ and differential operator \mathcal{L}^α : $a(x) = \sigma(x)\sigma'(x)$, and for $f(\cdot)$ in $C^2(\mathbb{R}^k)$,

$$\mathcal{L}^\alpha f(x) = f_x(x)'b(x, \alpha) + \frac{1}{2}\text{tr}[f_{xx}(x)a(x)], \quad (3.3)$$

where

$$\text{tr}[f_{xx}(x)a(x)] = \sum_{ij} f_{x_i x_j}(x)a_{ij}(x).$$

We take G to be a target set as in Section 3.1, and $\tau = \inf\{t : x(t) \in \partial G\}$.

For an admissible control $u(\cdot)$, the cost functional is

$$W(x, u) = E_x^u \left[\int_0^\tau k(x(s), u(s))ds + g(x(\tau)) \right]. \quad (3.4)$$

Define

$$V(x) = \inf W(x, u),$$

where the infimum is over the admissible controls.

We now apply a formal dynamic programming argument to derive the PDE which is satisfied by the optimal value function $V(\cdot)$. Suppose that $V(\cdot)$ is as smooth as necessary for the following calculations to be valid. Suppose that there is an optimal control $\bar{u}(\cdot)$ which is pure Markov. Let $\Delta > 0$, and let α be any value in \mathcal{U} . Define $\tilde{u}(\cdot)$ to be the control process that uses the feedback control $\bar{u}(\cdot)$ for $t \geq \Delta$ and uses the control identically equal to α for $t < \Delta$. Define the process $\tilde{x}(\cdot)$ to be the process which corresponds to use of the control $\tilde{u}(\cdot)$. Let $\tilde{\tau}$ denote the time that the target set is reached under this composite control. Let $x(\cdot)$ and τ denote the solution and escape time under the optimal control $\bar{u}(\cdot)$. By definition, we have

$$V(x) = E_x^{\bar{u}} \left[\int_0^\tau k(x(s), \bar{u}(x(s)))ds + g(x(\tau)) \right].$$

The optimality of $V(\cdot)$ implies

$$\begin{aligned} V(x) &\leq E_x^{\tilde{u}} \left[\int_0^\tau k(x(s), \tilde{u}(s))ds + g(x(\tau)) \right] \\ &= E_x^{\tilde{u}} \left[\int_0^{\tau \wedge \Delta} k(x(s), \alpha)ds + g(x(\tau))I_{\{\tau < \Delta\}} \right] \\ &\quad + E_x^{\tilde{u}} \left[\int_\Delta^\tau k(x(s), \bar{u}(x(s)))ds + g(x(\tau)) \right] I_{\{\tau \geq \Delta\}}. \end{aligned}$$

By the Markov property, the definition of $\tilde{u}(\cdot)$, and the optimality of $\bar{u}(\cdot)$, the inequality above may be rewritten as

$$V(x) \leq E_x^{\bar{u}} \left[\int_0^{\tau \wedge \Delta} k(x(s), \alpha)ds + g(x(\tau))I_{\{\tau < \Delta\}} + V(x(\Delta))I_{\{\tau \geq \Delta\}} \right].$$

Therefore

$$\frac{1}{\Delta} E_x^{\bar{u}} \left[V(x(\Delta)) - V(x) + \int_0^\Delta k(x(s), \alpha)ds \right] \geq \frac{1}{\Delta} E_x^{\bar{u}} h(\tau, \Delta, \tilde{u}) I_{\{\tau < \Delta\}}, \quad (3.5)$$

where

$$h(\tau, \Delta, \tilde{u}) = V(x(\Delta)) - \int_{\tau \wedge \Delta}^\Delta k(x(s), \alpha)ds - g(x(\tau))$$

is bounded uniformly in ω and Δ . If we assume (as in Subsection 3.1.1) the condition $P_x^{\bar{u}}\{\tau < \Delta\}/\Delta \rightarrow 0$ as $\Delta \rightarrow 0$, then the right hand side of (3.5) tends to zero as $\Delta \rightarrow 0$. Therefore, taking this limit yields that, for any value of α in \mathcal{U} ,

$$\mathcal{L}^\alpha V(x) + k(x, \alpha) \geq 0.$$

Suppose in the calculations above that α is replaced by $\bar{u}(x(s))$ on $[0, \Delta)$, and that $\bar{u}(\cdot)$ is continuous at x . Then the analogue of (3.5) holds with the inequality replaced by an equality. We then formally obtain the equation

$$\mathcal{L}^{\bar{u}(x)}V(x) + k(x, \bar{u}(x)) = 0.$$

It follows that

$$\begin{cases} \inf_{\alpha \in \mathcal{U}} [\mathcal{L}^\alpha V(x) + k(x, \alpha)] = 0, & x \in G^0 \\ V(x) = g(x), & x \in \partial G. \end{cases} \quad (3.6)$$

It should also be noted that

$$E_x^u V(x(t \wedge \tau)) \rightarrow E_x^u g(x(\tau))$$

for admissible u .

A Verification Theorem. Suppose that there is a bounded function $V(\cdot)$ which is in $C_b^2(G^0)$ and a feedback control $\bar{u}(x)$ such that (3.6) holds with the infimum taken on at $\alpha = \bar{u}(x)$, and that $E_x \tau < \infty$ holds for admissible controls for which the cost is bounded. [To simplify the discussion, we will assume that $b(\cdot, \cdot)$, $\sigma(\cdot)$, and the control $u(\cdot)$ are all Lipschitz. This allows all solutions considered here to be defined on a common fixed probability space (see Section 1.3). This condition will not be assumed in the convergence proofs later in the book.] Then it can be shown that $V(x)$ is indeed the optimal cost and $\bar{u}(\cdot)$ an optimal control. We will give an outline of the proof.

Let $\tilde{u}(\cdot)$ be an admissible control with $\tilde{x}(\cdot)$ the associated solution to (3.1). By the minimization in (3.6),

$$0 = \mathcal{L}^{\bar{u}(x)}V(x) + k(x, \bar{u}(x)),$$

$$0 \leq \mathcal{L}^{\tilde{u}(t)}V(\tilde{x}(t)) + k(\tilde{x}(t), \tilde{u}(t))$$

for all values of $x \in G^0$, $t \geq 0$, and ω . Let τ and $\tilde{\tau}$ denote the escape time under the controls $\bar{u}(\cdot)$ and $\tilde{u}(\cdot)$, respectively. Then by Itô's formula we can write

$$\begin{aligned} -E_x^{\bar{u}}V(x(t \wedge \tau)) + V(x) &= -E_x^{\bar{u}} \int_0^{t \wedge \tau} \mathcal{L}^{\bar{u}(x(s))}V(x(s))ds \\ &= \int_0^{t \wedge \tau} k(x(s), \bar{u}(x(s)))ds \end{aligned}$$

and

$$\begin{aligned} -E_x^{\tilde{u}} V(\tilde{x}(t \wedge \tilde{\tau})) + V(x) &= -E_x^{\tilde{u}} \int_0^{t \wedge \tilde{\tau}} \mathcal{L}^{\tilde{u}(s)} V(\tilde{x}(s)) ds \\ &\leq \int_0^{t \wedge \tilde{\tau}} k(\tilde{x}(s), \tilde{u}(s)) ds. \end{aligned}$$

Suppose that $E_x^{\tilde{u}} V(\tilde{x}(t \wedge \tilde{\tau})) \rightarrow E_x^{\tilde{u}} g(\tilde{x}(\tilde{\tau}))$ and $E_x^{\tilde{u}} V(x(t \wedge \tau)) \rightarrow E_x^{\tilde{u}} g(x(\tau))$ as $t \rightarrow \infty$. Then the equations just given imply

$$W(x, \bar{u}) = E_x^{\bar{u}} \left[\int_0^\tau k(x(s), \bar{u}(x(s))) ds + g(x(\tau)) \right] = V(x)$$

and

$$V(x) \leq E_x^{\tilde{u}} \left[\int_0^{\tilde{\tau}} k(\tilde{x}(s), \tilde{u}(s)) ds + g(\tilde{x}(\tilde{\tau})) \right] = W(x, \tilde{u}).$$

Hence, the minimizing control $\bar{u}(\cdot)$ is optimal.

There are numerous combinations of this framework of controlled drift term with the cost functions and processes which are discussed in this chapter. For example, an absorbing or target set can be added to the optimal stopping problem, and one can have both a choice over the stopping time and a continuously running cost. For example, consider a problem with the cost (3.4), but where the stopping time τ can be chosen to be any admissible stopping time which is no greater than the first time that the set G^0 is exited. Then the Bellman equation is

$$\begin{cases} \inf_{\alpha \in \mathcal{U}} [\mathcal{L}^\alpha V(x) + k(x, \alpha)] = 0, & x \in G^0, x \notin B, \\ V(x) = g(x), & x \in B, \end{cases}$$

with the absorbing boundary condition on ∂G and with the set B part of the solution.

3.4 A Discounted Problem with a Target Set and Reflection

Let us add a discount factor to the problem discussed in Section 3. The model can be either (1.1) or (1.2). The cost is now

$$W(x, u) = E_x^u \left[\int_0^\tau e^{-\beta s} k(x(s), u(s)) ds + e^{-\beta \tau} g(x(\tau)) \right]. \quad (4.1)$$

Then the Bellman equation for the minimum cost is

$$\inf_{\alpha} [\mathcal{L}^\alpha V(x) - \beta V(x) + k(x, a)] = 0, \quad x \in G^0, \quad (4.2)$$

together with the absorbing boundary condition.

As another alteration, let the “local discount rate” depend on the state. In particular, define $A(t) = \exp - \int_0^t \beta(x(s))ds$ for some bounded, continuous, and non-negative function $\beta(\cdot)$. Let the cost be (4.1), with βt replaced by $A(t)$. Then the Bellman equation for the optimal cost is just (4.2), with the β replaced by $\beta(x)$.

Let the cost be (4.1), and let the process be the reflected form of (1.1), with reflection set ∂G and reflection direction $r(x)$ as in Section 1.3. Then the Bellman equation is

$$\inf_{\alpha \in \mathcal{U}} [\mathcal{L}^\alpha V(x) - \beta V(x) + k(x, \alpha)] = 0, \quad x \in G^0$$

with boundary condition

$$V_x(x)'r(x) = 0.$$

3.5 Average Cost Per Unit Time

For the average cost per unit time problem to be computable from a numerical analysis point of view, it is usually necessary that the state space be compact. Thus, we are essentially confined to some sort of reflection problem. There could be some more general boundary process and this will be discussed in the chapter dealing with the convergence of the approximation for the reflected case. For the purposes of this chapter, we simply state the Bellman equation for an ergodic cost function and with a reflected diffusion. Let the domain G and directions of reflection $r(\cdot)$ be as in Subsection 3.1.3. Suppose that there is a smooth function $V(\cdot)$ and a constant γ which satisfy the equation

$$\begin{aligned} \inf_{\alpha \in \mathcal{U}} [\mathcal{L}^\alpha V(x) - \gamma + k(x, \alpha)] &= 0, & x \in G^0, \\ V_x(x)'r(x) &= 0, & x \in \partial G. \end{aligned}$$

Let the infimum be taken on by a feedback function $\bar{u}(\cdot)$ under which the reflected diffusion is well defined. Then one can show that $\bar{u}(\cdot)$ is optimal with respect to any admissible control $u(\cdot)$ for which

$$E_x^u V(x(t))/t \rightarrow 0$$

as $t \rightarrow \infty$. The proof of a verification theorem for this problem combines the ideas of Subsection 3.1.4 and Section 3.3.

4

The Markov Chain Approximation Method: Introduction

The main purpose of the book is the development of numerical methods for the solution of control or optimal control problems, or for the computation of functionals of the stochastic processes of interest, of the type described in Chapters 3, 7, 8, 9, 12, and 13. It was shown in Chapter 3 that the cost or optimal cost functionals can be the (at least formal) solutions to certain nonlinear partial differential equations. It is tempting to try to solve for or approximate the various cost functions and optimal controls by dealing directly with the appropriate PDE's, and numerically approximating their solutions. A basic impediment is that the PDE's often have only a formal meaning, and standard methods of numerical analysis might not be usable to prove convergence of the numerical methods. For many problems of interest, one cannot even write down a partial differential equation. The Bellman equation might be replaced by a system of "variational inequalities," or the proper form might not be known. Optimal stochastic control problems occur in an enormous variety of forms. As time goes on, we learn more about the analytical methods which can be used to describe and analyze the various optimal cost functions, but even then it seems that many important classes of problems are still not covered and new models appear which need even further analysis. The optimal stochastic control or stochastic modelling problem usually starts with a physical model, which guides the formulation of the precise stochastic process model to be used in the analysis. One would like numerical methods which are able to conveniently exploit the intuition contained in the physical model.

The general methods developed in this book can be applied to a very broad class of stochastic and deterministic control problems, as well as to problems of optimal filtering and optimal filtering combined with control, or for the computation of a large class of functionals of diffusion or jump diffusion processes. The methods are quite intuitive. They do not require an understanding of the analytical properties of the equation of the model or of the Bellman equation for the optimal cost function. They have been used with success on most of the usual problems of stochastic control. Some common forms of the methods do reduce to standard finite element or finite difference methods. But then, owing to the degeneracies of the operators or

to non-standard boundary conditions or controls, it seems that the standard methods of proof of numerical analysis often cannot be used.

In this chapter, we first describe the basic idea, which is quite simple. The procedure will be illustrated via a “canonical” problem, and other classes of problems will be treated in subsequent chapters.

The basic idea is the following. We approximate the original (controlled or not) problem with a simpler (controlled or not) stochastic process model and associated cost function for which the desired computation can be carried out. In particular, the approximating process is a (controlled or not) Markov chain on a finite state space. This state space is a “discretization” of the original state space of the problem. There are many methods for the numerical solution of such Markov chain problems (see, for example, Chapter 6). The approximating Markov chain is chosen such that certain “local” properties of the approximating chain are “similar” to those of the original controlled process. A cost function for the Markov chain model which is an appropriate analogue to that for the original model is then found. One chooses a Markov chain model for which the computation is reasonable. It turns out that the procedure can be used almost automatically, in that there are standard methods which can be used to construct the chains and cost functions. Two classes of methods, discussed in the next chapter, illustrate the possibilities.

The approximating chain is parametrized by a parameter (say, analogous to a finite difference interval or to a “finite element size” in classical numerical analysis), such that as the parameter goes to, say, zero, the “local properties” of the chain resemble more and more closely those of the original process. By local properties, we mean, essentially, the mean and mean square change per step, under any control, as well as the mean reflection direction for the problem with a reflecting boundary. Under very broad conditions, one can prove that the sequence of optimal cost functions for the sequence of approximating chains converges to that for the underlying original process as the approximation parameter goes to zero. The proofs are purely probabilistic: We never need to appeal to regularity properties of or even explicitly use the Bellman equation, whether it is formal or not. In addition, one can take advantage of knowledge of or intuition concerning the physical process. The optimal value function for the approximating chain is an optimal value function for a controlled process and cost criterion which are very close to the originally given ones. The convergence is analogous to the convergence of a sequence of finite difference or finite element approximations to a PDE as the approximation interval goes to zero.

In order to illustrate the idea in a simple form, we first show how to use it on a controlled Wiener process in Section 4.4. In Section 4.5, we show how to use it on a deterministic problem for which current alternative methods seem to be less intuitive, and even more complex. It will be shown that one form of the approach is equivalent to a finite element approximation to the

Bellman equation for that deterministic problem, and that a convergence proof can be readily obtained using simple probabilistic methods, even though the original problem is not probabilistic.

In the previous chapter, we defined an admissible control for several types of stochastic problems. It turns out that for convergence analysis, neither that definition of admissible control nor its deterministic analog are always adequate, and one has to enlarge the class of allowed controls, to the so-called “relaxed controls.” The class is enlarged in such a way that the infimum of the value functions over the enlarged class is the same as that over the original class, so that the optimal value function does not change when working with the larger class of controls. The enlarged class of controls is used for analytical purposes only. The numerical procedures will always give feedback controls. The discussion of the convergence of the numerical method applied to the deterministic problem gives us the opportunity to introduce the notion of deterministic relaxed control, and to discuss its proper role. This will be developed further together with the stochastic relaxed control in Chapter 9. There is some overlap between this chapter and Chapter 5, but that overlap concerns fundamental points.

The general Markov chain approximation method is outlined in Section 4.1. The actual process which approximates the original controlled diffusion or jump diffusion is a continuous time parameter interpolation of the Markov chain. Two useful interpolations are discussed in Sections 4.2 and 4.3. The process discussed in Section 4.3 is a continuous time parameter Markov chain and will be the one most used in the sequel. The ideas of Sections 4.1 to 4.3 will be used heavily in the rest of the book. The material in Sections 4.4 to 4.6 illustrates their application in special simple problems.

4.1 The Markov Chain Approximation Method

In this section, we will describe and motivate the basic type of Markov chain approximation that will be used and show that it is quite natural. For illustrative purposes, consider the diffusion process model:

$$dx(t) = b(x(t), u(t))dt + \sigma(x(t))dw. \quad (1.1)$$

Let G be a compact set which is the closure of its interior G^0 . For $\beta > 0$, we consider the discounted cost

$$W(x, u) = E_x^u \int_0^\tau e^{-\beta t} k(x(t), u(t))dt + E_x^u e^{-\beta \tau} g(x(\tau)), \quad (1.2)$$

where $\tau = \inf\{t : x(t) \notin G^0\}$, the first escape time of $x(\cdot)$ from G^0 . Define

$$V(x) = \inf_u W(x, u),$$

where the infimum is over all admissible controls. Recall (Section 1.3) that an admissible control $u(\cdot)$ is a measurable process which is nonanticipative with respect to $w(\cdot)$, and $u(t)$ takes values in \mathcal{U} , a compact set. The development will be entirely formal, because we are concerned with motivation only. But in order to be certain that we are on solid ground, let us suppose here that the diffusion is well defined for any admissible control. In particular, suppose that:

A1.1. *$b(\cdot)$ and $\sigma(\cdot)$ are bounded, continuous, and Lipschitz continuous in x , uniformly in u . Both $k(\cdot)$ and $g(\cdot)$ are bounded and continuous.*

The conditions will be weakened in Chapters 9-12.

Approximating the Process: General Remarks. The methods to be employed are not based on the analytical expressions (e.g., the formal PDE's of Chapter 3) for the functions whose values we wish to compute. Rather, they are based on approximating the basic controlled process (1.1) by a simpler controlled process, for which the evaluation of either the cost function for a fixed control or of the optimal cost can be done with an acceptable amount of computational work. If the approximating controlled process is close to the original process $x(\cdot)$ in an appropriate statistical sense and the associated cost function is close to (1.2), then we would expect that the value of the cost function for the approximating process for a fixed control (or to its optimal value over all controls) will be close to the cost function $W(x, u)$ for a similar control [or to its optimal value $V(x)$, respectively]. The approximating processes will be piecewise constant. Let $h > 0$ be a scalar approximation parameter. The basis of the approximation is a discrete time parameter finite state controlled Markov chain $\{\xi_n^h, n < \infty\}$ whose "local properties" are "consistent" with those of (1.1). The continuous time parameter approximating process will be a piecewise constant interpolation of this chain, with appropriately chosen interpolation intervals. The next chapter is devoted to convenient ways of getting the approximating chains.

A Markov Chain Approximation: Terminology. For each $h > 0$, let $\{\xi_n^h, n < \infty\}$ be a controlled discrete parameter Markov chain on a discrete state space $S_h \in \mathbb{R}^r$ with transition probabilities denoted by $p^h(x, y|\alpha)$. The α is the control parameter and takes values in the compact set \mathcal{U} . We use u_n^h to denote the random variable which is the actual control action for the chain at discrete time n . We now define some terms which will allow us to relate the chain to the diffusion (1.1). In Chapter 5, we will show how all of these quantities can be readily calculated. Suppose that we have an "interpolation interval" $\Delta t^h(x, \alpha) > 0$, and define $\Delta t_n^h = \Delta t^h(\xi_n^h, u_n^h)$. Let $\sup_{x, \alpha} \Delta t^h(x, \alpha) \rightarrow 0$ as $h \rightarrow 0$, but $\inf_{x, \alpha} \Delta t^h(x, \alpha) > 0$ for each $h > 0$. This positivity will be relaxed when considering systems with singular con-

trols or instantaneously reflecting boundaries. Let G_h^0 denote the components of the state space which are interior to the set $G : G_h^0 = S_h \cap G^0$. Thus, G_h^0 is the *finite state space* of the chain until it escapes from G^0 .

Define the difference $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$. Let $E_{x,n}^{h,\alpha}$ denote the conditional expectation given $\{\xi_i^h, u_i^h, i \leq n, \xi_n^h = x, u_n^h = \alpha\}$. Suppose that the chain obeys the following “local consistency” conditions, which also defines the functions $b_h(\cdot)$ and $a_h(\cdot)$:

$$\begin{aligned} E_{x,n}^{h,\alpha} \Delta\xi_n^h &\equiv b_h(x, \alpha) \Delta t^h(x, \alpha) = b(x, \alpha) \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)), \\ E_{x,n}^{h,\alpha} [\Delta\xi_n^h - E_{x,n}^{h,\alpha} \Delta\xi_n^h] [\Delta\xi_n^h - E_{x,n}^{h,\alpha} \Delta\xi_n^h]' &\equiv a_h(x) \Delta t^h(x, \alpha) \\ &= a(x) \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)), \end{aligned} \quad (1.3)$$

$$\begin{aligned} a(x) &= \sigma(x)\sigma'(x), \\ \sup_{n,\omega} |\xi_{n+1}^h - \xi_n^h| &\xrightarrow{h} 0. \end{aligned}$$

Note that the chain has the “local properties” of the diffusion process (1.1) in the following sense: By Section 1.3, letting $x(0) = x$ and $u(t) = \alpha$ on the interval $[0, \delta]$ in (1.1) gives us

$$\begin{aligned} E_x(x(\delta) - x) &= b(x, \alpha)\delta + o(\delta), \\ E_x[x(\delta) - x][x(\delta) - x]' &= a(x)\delta + o(\delta). \end{aligned} \quad (1.4)$$

The local consistency (1.3) is essentially all that is required of the approximating chain, except for the analogous considerations which will be needed when dealing with jump diffusions, reflecting boundaries, or singular controls.

Following the terminology in Chapter 2, we say that a control policy $u^h = \{u_n^h, n < \infty\}$ for the chain is *admissible* if the chain has the Markov property under that policy. In particular, the policy is admissible if

$$P\{\xi_{n+1}^h = y | \xi_i^h, u_i^h, i \leq n\} = p^h(\xi_n^h, y | u_n^h). \quad (1.5)$$

Let $E_x^{u^h}$ denote the expectation, given that $\xi_0^h = x$ and that either an admissible control sequence $u^h = \{u_n^h, n < \infty\}$ or a feedback control denoted by $u^h(\cdot)$ is used, according to the case.

4.2 Continuous Time Interpolation and Approximating Cost Function

The chain $\{\xi_n^h, n < \infty\}$ is a discrete time parameter process. In order to approximate the continuous time parameter process $x(\cdot)$, we will need to use an appropriate continuous time interpolation of the chain. Owing to

the properties of the “interpolation interval” $\Delta t^h(x, \alpha)$, a natural interpolating time scale might be obvious. There are two interpolations which are useful, the first [to be denoted by $\xi^h(\cdot)$] uses interpolation intervals which are deterministic functions of the current state and control value. The second interpolation [to be denoted by $\psi^h(\cdot)$] is actually a *continuous parameter Markov process*. The first was used in much of the convergence analysis in past works [80, 76]. The second was initially introduced in [82]. It allows some simplifications in the proofs and will be used in most of the convergence analysis of this book. In this section, we will define $\xi^h(\cdot)$ and discuss appropriate analogs [(2.2) and (2.3) below] of the cost function (1.2) for the chain, and then write the dynamic programming equations for the resulting control problem. In the next section, the second interpolation $\psi^h(\cdot)$ is defined and its properties discussed.

Let $\{u_n^h, n < \infty\}$ be an admissible control for the chain and define the interpolated time $t_n^h = \sum_0^{n-1} \Delta t_i^h$. Define the *continuous parameter interpolations* $\xi^h(\cdot)$ and $u^h(\cdot)$ by:

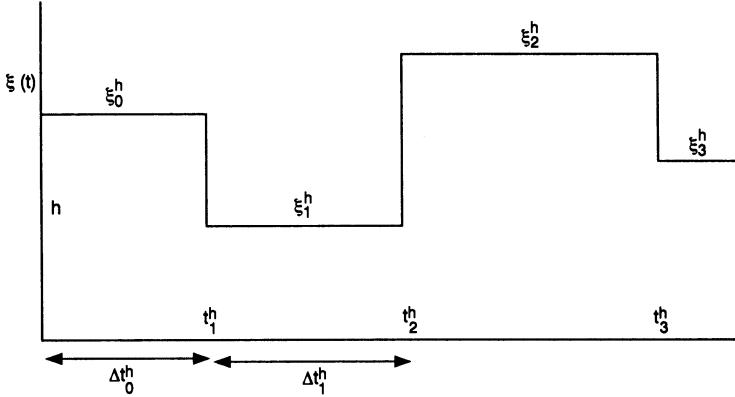
$$\xi^h(t) = \xi_n^h, \quad u^h(t) = u_n^h, \quad t \in [t_n^h, t_{n+1}^h]. \quad (2.1)$$

The construction of $\xi^h(\cdot)$ is illustrated in Figure 4.1.

The interpolated process defined by (2.1) is an approximation to the diffusion (1.1) in the sense that the “local properties” (1.3) hold. The interpolation intervals $\Delta t^h(x, \alpha)$ can always be chosen to be constant if we wish (see Section 5.2), but we might find that restrictive. For example, if the local velocity $b(\cdot)$ is large at some value of x , then we might want to use a smaller interpolation interval there. Also, the numerical procedures converge faster when we take advantage of the added flexibility allowed by the variable interpolation intervals. The interpolated process $\xi^h(\cdot)$ is piecewise constant. Given the value of the current state and control action, the current interval is known. The interpolation intervals are obtained automatically when the transition functions $p^h(x, y|\alpha)$ are constructed. See Chapter 5 and Sections 4.4 and 4.5 below. Let N_h denote the first time that $\{\xi_n^h, n < \infty\}$ leaves G^0 . Then, the first exit time of $\xi^h(\cdot)$ from G^0 is defined by $\tau_h = t_{N_h}^h$. There are several natural cost functionals for the chain which approximate (1.2), depending on how time is discounted on the intervals of constancy $[t_n^h, t_{n+1}^h]$. If the discounting is constant on this interval then we can use the approximation

$$W_1^h(x, u^h) = E_x^{u^h} \sum_{n=0}^{N_h-1} e^{-\beta t_n^h} k(\xi_n^h, u_n^h) \Delta t_n^h + E_x^{u^h} e^{-\beta \tau_h} g(\xi_{N_h}^h). \quad (2.2)$$

If future time is to be “continuously discounted,” then we can use the

Figure 4.1. Construction of the interpolation $\xi^h(\cdot)$.

approximation

$$\begin{aligned}
 W_2^h(x, u^h) &= E_x^{u^h} \sum_{n=0}^{N_h-1} k(\xi_n^h, u_n^h) \int_{t_n^h}^{t_{n+1}^h} e^{-\beta t} dt + E_x^{u^h} e^{-\beta \tau_h} g(\xi^h(\tau_h)) \\
 &= E_x^{u^h} \int_0^{\tau_h} e^{-\beta t} k(\xi^h(t), u^h(t)) dt + E_x^{u^h} e^{-\beta \tau_h} g(\xi^h(\tau_h)).
 \end{aligned} \tag{2.3}$$

We have

$$W_2^h(x, u^h) - W_1^h(x, u^h) = E_x^{u^h} \sum_{n=0}^{N_h-1} \int_{t_n^h}^{t_{n+1}^h} (e^{-\beta t} - e^{-\beta t_n^h}) k(\xi_n^h, u_n^h) dt,$$

which goes to zero as $h \rightarrow 0$ because $\sup_{x, \alpha} \Delta t^h(x, \alpha) \xrightarrow{h \rightarrow 0} 0$. Let us define $V_i^h(x) = \inf_u W_i^h(x, u)$, where the infimum is over all admissible controls. The cost functions (2.2) and (2.3) both approximate (1.2). We clearly have

$$|V_1^h(x) - V_2^h(x)| \xrightarrow{h \rightarrow 0} 0.$$

The dynamic programming equation (2.3.2) for cost function (2.2) is

$$V_1^h(x) = \begin{cases} \min_{\alpha \in \mathcal{U}} \left[\sum_y e^{-\beta \Delta t^h(x, \alpha)} p^h(x, y | \alpha) V_1^h(y) \right. \\ \quad \left. + k(x, \alpha) \Delta t^h(x, \alpha) \right], & x \in G_h^0, \\ g(x), & x \notin G_h^0. \end{cases} \tag{2.4}$$

The dynamic programming equation for $V_2^h(x)$ is the same, except that the

$\Delta t^h(x, \alpha)$ coefficient of $k(x, \alpha)$ is replaced by

$$\int_0^{\Delta t^h(x, \alpha)} e^{-\beta s} ds = [1 - e^{-\beta \Delta t^h(x, \alpha)}] / \beta = \Delta t^h(x, \alpha) + O((\Delta t^h(x, \alpha))^2). \quad (2.5)$$

A third possibility for the approximation of the discount factor appears in (3.7), and is based on the continuous parameter Markov chain interpolation. The difference between the solutions of (2.4) and (3.7) goes to zero as $h \rightarrow 0$.

Discussion. The similarity of the cost functions (2.2) and (2.3) to (1.2) and the similarity of the local properties of the interpolation $\xi^h(\cdot)$ to those of the original controlled diffusion $x(\cdot)$ suggest that the $V_i^h(x)$ might be good approximations to $V(x)$ for small values of h . This turns out to be true. Any sequence $\xi^h(\cdot)$ has a subsequence which converges in an appropriate sense to a controlled diffusion of the type (1.1). This will be dealt with in Chapters 9 and 10. Suppose that $\tilde{u}^h(x)$ is the optimal control for the chain $\{\xi_n^h, n < \infty\}$ with cost function (say) (2.2), and suppose that the associated sequence $\xi^h(\cdot)$ converges to a limit diffusion $x(\cdot)$ with admissible control $\tilde{u}(\cdot)$. Under quite broad conditions, the sequence τ_h of times that the chains first exit G_h^0 will also converge to the time that the limit process $x(\cdot)$ first exits G^0 . If this is the case, then the cost functionals $V_i^h(x)$ for the sequence of chains will converge to the cost functional $W(x, \tilde{u})$ for the limit process. Because $V(x)$ is the optimal value function, we have that $W(x, \tilde{u}) \geq V(x)$ and, hence, $\liminf_h V_i^h(x) \geq V(x)$. The reverse inequality will be proved by another approximation procedure, which uses the optimality of the cost functionals $V_i^h(x)$ for the controlled chain. For the mathematical proof of the convergence, we might need to extend the class of allowed controls to a class of so-called “relaxed controls,” but the infimum of the cost function over the original class of controls and that over the new class of controls are equal. A good example of the entire procedure is given in Sections 4.5 and 4.6 below for a simple deterministic problem.

The Markov chain approximation method thus seems to be quite straightforward: (a) get a locally consistent chain; (b) get a suitable approximation to the original cost function for the chain.

4.3 A Continuous Time Markov Chain Interpolation

By the construction of the process $\xi^h(\cdot)$ in the last section, its “holding times” or interpolation intervals are $\Delta t^h(\xi_n^h, u_n^h)$. Once the control and state at the start of the n^{th} interval are known, the width of the interval is known. It is sometimes more convenient for the proofs of convergence

to use a continuous parameter interpolation of $\{\xi_n^h, n < \infty\}$ that is a Markov process itself. We now show how to construct this new continuous parameter interpolation, which will be denoted by $\psi^h(\cdot)$. Define $\tau_0^h = 0$, let $\{\tau_n^h, n < \infty\}$ denote the moments of change of $\psi^h(\cdot)$, and set $\Delta\tau_n^h = \tau_{n+1}^h - \tau_n^h$. Define $\psi^h(\cdot)$ at the moments of change by

$$\psi^h(\tau_n^h) = \xi_n^h, \quad n < \infty. \quad (3.1)$$

An alternative definition is

$$\psi^h(t) = \sum_{i: \tau_{i+1}^h \leq t} \Delta\xi_i^h + \xi_0^h. \quad (3.1')$$

We need only define the distribution of $\Delta\tau_n^h$, conditioned on $(\xi_n^h = x, u_n^h = \alpha)$. This will be an exponential distribution with mean value $\Delta t^h(x, \alpha)$. We next write this more precisely.

Using the same notation as we did in the last section, define $u^h(\cdot)$ by $u^h(t) = u_n^h$ for $t \in [\tau_n^h, \tau_{n+1}^h]$. Define the holding times by their conditional probability distribution:

$$P\{\Delta\tau_n^h \leq t | \xi_i^h, u_i^h, \tau_i^h, i \leq n; \xi_n^h = x, u_n^h = \alpha\} = 1 - \exp[-t/\Delta t^h(x, \alpha)].$$

The mean holding times satisfy

$$E_{x,n}^{h,\alpha} \Delta\tau_n^h = \Delta t^h(x, \alpha), \quad (3.2)$$

which is just the holding interval for the interpolation $\xi^h(\cdot)$. Let $P_{x,n}^{h,\alpha}$ denote the conditional probability defined by the conditional expectation $E_{x,n}^{h,\alpha}$. Let $E_{x,t}^{h,\alpha}$ (with associated conditional probability $P_{x,t}^{h,\alpha}$) denote the expectation given the data

$$\{\psi^h(s), u^h(s), s \leq t; \tau_n^h : \tau_n^h \leq t; \psi^h(t) = x, u^h(t) = \alpha\}.$$

Local Properties of $\psi^h(\cdot)$ and a Convention Concerning “Zero” Jumps. For the controlled Markov chain $\{\xi_n^h, n < \infty\}$, it is possible that the transition probability $p^h(x, x|\alpha)$ be positive for some values of x and α . In this case, the *actual sample value* of $\psi^h(\cdot)$ might not change with probability one at each time τ_n^h . For notational consistency, we allow “zero” jumps for $\psi^h(\cdot)$. That is, the jump times for $\psi^h(\cdot)$ are *defined* to be the times $\{\tau_n^h\}$ whether or not $\psi^h(\cdot)$ actually changes value at all of those times.

By definition, for each h ,

$$P_{x,t}^{h,\alpha}\{\text{jump on } [t, t + \delta]\} = \frac{\delta}{\Delta t^h(x, \alpha)} + o(\delta).$$

For $\delta > 0$, define the increment $\Delta\psi^h(t) = \psi^h(t + \delta) - \psi^h(t)$. The local properties of $\psi^h(\cdot)$ follow from (1.3) and the above definitions and are

$$P_{x,t}^{h,\alpha} \{ \psi^h(t+\delta) = y, \text{ jump on } [t, t+\delta) \} = \frac{\delta}{\Delta t^h(x, \alpha)} p^h(x, y | \alpha) + o(\delta), \quad (3.3)$$

$$\begin{aligned} E_{x,t}^{h,\alpha} \Delta \psi^h(t) &= P_{x,t}^{h,\alpha} \{ \text{jump on } [t, t+\delta) \} \sum_y p^h(x, y | \alpha) (y - x) \\ &= P_{x,t}^{h,\alpha} \{ \text{jump on } [t, t+\delta) \} b_h(x, \alpha) \Delta t^h(x, \alpha) \\ &= \frac{\delta}{\Delta t^h(x, \alpha)} b(x, \alpha) \Delta t^h(x, \alpha) + \delta o(\Delta t^h(x, \alpha)) + o(\delta), \end{aligned} \quad (3.4)$$

$$E_{x,t}^{h,\alpha} [\Delta \psi^h(t)] [\Delta \psi^h(t)]' = a(x) \delta + \delta o(\Delta t^h(x, \alpha)) + o(\delta). \quad (3.5)$$

These will be extended to cover the jump diffusion process in Chapter 5. We define an *admissible control* for the process $\psi^h(\cdot)$ as it was defined for the chain $\{\xi_n^h, n < \infty\}$. It is any \mathcal{U} -valued process which is constant on the intervals $[\tau_n^h, \tau_{n+1}^h)$ and for which the imbedded chain $\{\xi_n^h, n < \infty\}$ has the Markov property (1.5). Thus, there is a complete equivalence between the control models with the discrete and continuous parameter chains.

Abusing terminology, let us *reuse* the symbols τ_h for the escape times from G^0 as in the last section, but here it is for the process $\psi^h(\cdot)$. A natural analogue of the cost function (1.2) is

$$\begin{aligned} W^h(x, u^h) &= E_x^{u^h} \sum_{n=0}^{N_h-1} k(\xi_n^h, u_n^h) e^{-\beta \tau_n^h} \int_{\tau_n^h}^{\tau_{n+1}^h} e^{-\beta t} dt + E_x^{u^h} e^{-\beta \tau_h} g(\psi^h(\tau_h)) \\ &= E_x^{u^h} \int_0^{\tau_h} e^{-\beta t} k(\psi^h(t), u^h(t)) dt + E_x^{u^h} e^{-\beta \tau_h} g(\psi^h(\tau_h)), \end{aligned} \quad (3.6)$$

where the symbol u^h denotes the use of admissible $u^h(\cdot)$.

The Dynamic Programming Equation for Cost (3.6) and Process $\psi^h(\cdot)$. The dynamic programming equation is the same as (2.4) except for a slight difference in the discount factor. Note that

$$E_{x,n}^{h,\alpha} \int_0^{\Delta \tau_n^h} e^{-\beta s} ds = \frac{\Delta t^h(x, \alpha)}{1 + \beta \Delta t^h(x, \alpha)}$$

and

$$E_{x,n}^{h,\alpha} e^{-\beta \Delta \tau_n^h} = \frac{1}{1 + \beta \Delta t^h(x, \alpha)}.$$

Thus the integral on the right hand side of (3.6) can be written as

$$E_x^{u^h} \sum_{n=0}^{N_h-1} e^{-\beta \tau_n^h} k(\xi_n^h, u_n^h) \frac{\Delta \tau_n^h}{1 + \beta \Delta t^h(x, \alpha)}.$$

Consequently, the *effective average discount factor* from time τ_n^h to time τ_{n+1}^h , given that $\xi_n^h = x$ and $u_n^h = \alpha$, is

$$\frac{1}{1 + \beta \Delta t^h(x, \alpha)} = \exp[-\beta \Delta t^h(x, \alpha)](1 + O(\Delta t^h(x, \alpha))).$$

Define $V^h(x) = \inf_u W^h(x, u)$, where the infimum is over all admissible control sequences. Then the dynamic programming equation for the controlled chain $\{\xi_n^h, n < \infty\}$ and cost (3.6) is

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \left[\frac{1}{1 + \beta \Delta t^h(x, \alpha)} \sum_y p^h(x, y | \alpha) V^h(y) + k(x, \alpha) \frac{\Delta t^h(x, \alpha)}{1 + \beta \Delta t^h(x, \alpha)} \right], \quad (3.7)$$

for $x \in G_h^0$ and with the boundary condition $V^h(x) = g(x)$, $x \notin G_h^0$. We have $|V^h(x) - V_i^h(x)| \xrightarrow{h} 0$, and any of the above dynamic programming equations can be used with the same asymptotic results. We will return to this equation in connection with the example in Section 4.5.

An Alternative Derivation of (3.7). It is evident from the representations (3.6) that the cost functions and the optimal cost function for the $\{\xi_n^h, n < \infty\}$ and $\psi(\cdot)$ processes are the same, modulo slight differences in the discount factor. But, in order to complete the demonstration of the equivalence of the discrete and continuous parameter problem, let us formally derive the dynamic programming equation for the optimal value function $V^h(x)$ for the continuous parameter Markov chain model $\psi^h(\cdot)$ and cost function (3.6) directly. The formality of the development is mainly in that we ignore the possibility of multiple events on small intervals $[0, \delta]$, but it can be shown that the contribution of these possibilities goes to zero as $\delta \rightarrow 0$. By the principle of optimality, for $x \in G_h^0$ and small $\delta > 0$, we can write

$$\begin{aligned} V^h(x) &= \min_{\alpha \in \mathcal{U}} \left[e^{-\beta \delta} E_{x,0}^{h,\alpha} V^h(\psi^h(\delta)) + k(x, \alpha) \delta \right] \\ &= \min_{\alpha \in \mathcal{U}} \left[V^h(x) \left(1 - \frac{\delta}{\Delta t^h(x, \alpha)} \right) e^{-\beta \delta} + \frac{\delta}{\Delta t^h(x, \alpha)} \sum_y e^{-\beta \delta} V^h(y) p^h(x, y | \alpha) + k(x, \alpha) \delta + o(\delta) \right]. \end{aligned}$$

Collecting the coefficients of $V^h(x)$, dividing all terms by δ , multiplying all terms by $\Delta t^h(x, \alpha)$, and letting $\delta \rightarrow 0$ yields (3.7). Thus, the minimal cost for the continuous parameter Markov chain interpolation is just that for the discrete parameter Markov chain model with a particular form of the discount factor. Let us note that if there is no discounting ($\beta = 0$), then

the dynamic programming equation is the same for the discrete and the continuous time problems.

A Useful Representation of $\psi^h(\cdot)$. We next give an important representation for $\psi^h(\cdot)$ which will be useful in the proofs of Chapter 10, and which also gives us a better intuitive feeling for the relationship between the processes $\psi^h(\cdot)$ and $x(\cdot)$. Let us define the following “limit of conditional expectations.” By (3.4) and the definition of $b_h(\cdot)$ in (1.3),

$$\begin{aligned} b_h(x, \alpha) &= \lim_{\delta \rightarrow 0} E_{t,x}^{h,\alpha} \frac{\Delta \psi^h(t)}{\delta} \\ &= b(x, \alpha) + \frac{o(\Delta t^h(x, \alpha))}{\Delta t^h(x, \alpha)}. \end{aligned} \quad (3.8)$$

Now, factoring out this conditional mean rate of change of $\psi^h(\cdot)$ or “compensator,” as it is commonly called, and letting $x(0) = \psi^h(0) = x$, we can write [the expression defines $M^h(\cdot)$]

$$\psi^h(t) = x + \int_0^t b_h(\psi^h(s), u^h(s)) ds + M^h(t). \quad (3.9)$$

The jumps of $M^h(\cdot)$ are those of $\psi^h(\cdot)$ and, hence go to zero as $h \rightarrow 0$. Between the jumps, the process is linear in t . The process $M^h(\cdot)$ is a martingale whose quadratic variation is $\int_0^t a_h(\psi^h(s), u^h(s)) ds$, where $a_h(\cdot)$ is defined by (1.3) and also equals

$$\begin{aligned} a_h(x, \alpha) &= \lim_{\delta \rightarrow 0} E_{x,t}^{h,\alpha} \frac{[\Delta \psi^h(t) - E_{x,t}^{h,\alpha} \Delta \psi^h(t)][\Delta \psi^h(t) - E_{x,t}^{h,\alpha} \Delta \psi^h(t)]'}{\delta} \\ &= \lim_{\delta \rightarrow 0} E_{x,t}^{h,\alpha} \frac{[M^h(t + \delta) - M^h(t)][M^h(t + \delta) - M^h(t)]'}{\delta} \\ &= a(x) + \frac{o(\Delta t^h(x, \alpha))}{\Delta t^h(x, \alpha)}. \end{aligned} \quad (3.10)$$

The resemblance of (3.9) to the diffusion (1.1) is more than accidental. It will be shown in Chapter 10 that, under quite reasonable conditions, any sequence of $\psi^h(\cdot)$ processes has a subsequence which converges in a particular sense and the limit satisfies (1.1), with an appropriate admissible control, and where the stochastic integral is the limit of the $M^h(\cdot)$.

4.4 A Random Walk Approximation to the Wiener Process

We now give a simple application of the concepts of the previous section in order to fix the ideas. Perhaps the most classical use of Markov chain

approximations to diffusion type processes (1.1) is the use of the symmetric random walk to approximate the Wiener process. First consider the case of no control. Let $x(t) = x + w(t)$, where $w(\cdot)$ is a real valued Wiener process, and, for $h > 0$, define the set of points on the line $S_h = \{0, \pm h, \pm 2h, \dots\}$ Let $\{\xi_n^h, n < \infty\}$ be a symmetric random walk on the set S_h , and define the interpolation intervals $\Delta t^h(x) = \Delta t^h = h^2$. Define the continuous time interpolation $\xi^h(\cdot)$ as in Section 4.2. Then $\xi^h(\cdot)$ can be represented in the form

$$\xi^h(t) = x + \sum_{i=0}^{t/h^2-1} h\rho_i,$$

where the ρ_i are mutually independent and take values ± 1 , each with probability $1/2$. Then it is well known [12, 42, 79] that a broad class of functionals of $x(\cdot)$ can be approximated by the same functionals of $\xi^h(\cdot)$ and that $\xi^h(\cdot)$ converges to $x(\cdot)$ in a weak or distributional sense. A similar result holds for the continuous parameter Markov chain interpolation $\psi^h(\cdot)$.

Now let us add a control, and write the system as

$$x(t) = x + w(t) + \int_0^t u(s)ds, \quad (4.1)$$

where we require that $|u(t)| \leq 1$. Let G be the interval $[0, B]$, where $B > 0$ is supposed to be an integer multiple of h . Let $h \leq 1$, and define the transition probabilities for the approximating Markov chain as:

$$p^h(x, x \pm h | \alpha) = \frac{1 \pm h\alpha}{2}.$$

It is easily verified that the chain is locally consistent with $x(\cdot)$ in the sense of (1.3). The interpolation interval is just h^2 . The dynamic programming equation for the cost function $W_1^h(x, u)$ of (2.2) is given by (2.4); namely,

$$V_1^h(x) = \min_{|\alpha| \leq 1} e^{-\beta h^2} \left[\frac{1 + h\alpha}{2} V_1^h(x + h) + \frac{1 - h\alpha}{2} V_1^h(x - h) + k(x, \alpha)h^2 \right] \quad (4.2)$$

for $x \in (0, B)$, with $V_1^h(x) = g(x)$ otherwise.

The dynamic programming equation for the cost (1.2) and system (4.1) with $G = [0, B]$ is given by (3.4.2), which in the present case takes the form:

$$\begin{aligned} 0 &= \min_{\alpha \in \mathcal{U}} \left[\frac{V_{xx}(x)}{2} + V_x(x)\alpha + k(x, \alpha) - \beta V(x) \right], & x \in (0, B), \\ V(x) &= g(x), & x = 0, B. \end{aligned} \quad (4.3)$$

The expression (4.2) can be related to a finite difference approximation to (4.3) as follows: Let us take a finite difference approximation to (4.3) and

use the following approximations, where h is the difference interval:

$$V_{xx}(x) \rightarrow \frac{V(x+h) + V(x-h) - 2V(x)}{h^2},$$

$$V_x(x) \rightarrow \frac{V(x+h) - V(x-h)}{2h}.$$

Then, using $V^h(x)$ to denote the finite difference approximation to (4.3), for $x \in (0, B)$ we have

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \frac{1}{1 + \beta h^2} \left[\frac{1 + \alpha h}{2} V^h(x+h) + \frac{1 - \alpha h}{2} V^h(x-h) + k(x, \alpha) h^2 \right]. \quad (4.4)$$

Note that (4.4) is the dynamic programming equation for a discounted cost problem for a controlled random walk with $\Delta t^h(x, \alpha) = h^2$ and discount factor $1/(1 + \beta \Delta t^h(x, \alpha))$. It is just (3.7) and is an $O(h^4)$ approximation to (4.2). The consistency of the equation (4.4), obtained by using the finite difference method with the dynamic programming equation obtained for the process $\psi^h(\cdot)$ in the last section or with (4.2), suggests that the control problem for the Markov chain might actually be a good approximation to that for the original problem. Appropriate finite difference approximations can often be used to obtain approximating Markov chains, as will be seen in Chapter 5. But we emphasize that from the point of view of the convergence proofs it is the controlled process which is approximated and not the formal PDE (4.3).

4.5 A Deterministic Discounted Problem

Many of the concepts which are involved in the use of the Markov chain approximation method and the convergence proofs can be illustrated by a discounted cost and a purely *deterministic* problem. In fact, the example that follows illustrates the practical use of the approximation and numerical methods for purely deterministic systems when feedback controls or optimal value functions are wanted. We will work with the following system:

$$\dot{x} = b(x, u(t)), \quad (5.1a)$$

$$W(x, u) = \int_0^\infty e^{-\beta t} k(x(t), u(t)) dt, \quad \beta > 0. \quad (5.1b)$$

The functions $b(\cdot)$ and $k(\cdot)$ are assumed to be bounded and continuous. The admissible controls are just the \mathcal{U} -valued measurable functions on the interval $[0, \infty)$, where \mathcal{U} is a compact set. In order to assure that the solution is well defined and to avoid needless complications in the rest of this chapter, we suppose that $b(\cdot, \alpha)$ is Lipschitz continuous, uniformly in α . Let

there be continuous functions $\Delta t^h(x, \alpha)$ which satisfy $k_2 h \geq \Delta t^h(x, \alpha) \geq k_1 h$ for some constants $k_i > 0$. Because the example is used for illustrative purposes only and we wish to keep it simple, we do not include a stopping boundary or target set. Thus, the control is over the infinite time interval and infinite space. In this sense, we are not necessarily working with a computationally feasible algorithm. But the simplifications enable the main points to be made without introducing excessive notation or concepts.

Remark on Stopping Boundaries. If a stopping boundary or target set is introduced, then we need to add a condition which, loosely speaking, reads as follows: For each $\epsilon > 0$, there is an ϵ -optimal control such that the corresponding trajectory of (5.1a) is not “tangent” to the boundary at the first contact time. This will be dealt with in detail in the context of the stochastic problem in Sections 10.4 and 10.5, and for the deterministic problem in Section 13.2.

An Approximation to (5.1a). We start by following a time honored procedure and use a discrete time approximation to (5.1a), but where the time intervals between successive updates might depend on the current state and control. It will be seen that a natural approach to the computational problem can be interpreted as a computational problem for a controlled “approximating” Markov chain. Let h denote an approximation parameter. Use the discrete time approximation to (5.1a) given by

$$\tilde{\xi}_{n+1}^h = \tilde{\xi}_n^h + b(\tilde{\xi}_n^h, u_n^h) \Delta t^h(\tilde{\xi}_n^h, u_n^h),$$

where u_n^h is the actual control which is used at the n^{th} update. Define the sequence $u^h = \{u_n^h, n < \infty\}$ and the interpolation interval $\Delta \tilde{t}_n^h = \Delta t^h(\tilde{\xi}_n^h, u_n^h)$. Define $t_n^h = \sum_{i=0}^{n-1} \Delta \tilde{t}_i^h$. Define the continuous parameter process $\tilde{\xi}^h(\cdot)$ by $\tilde{\xi}^h(t) = \tilde{\xi}_n^h, t \in [t_n^h, t_{n+1}^h]$. It is the $\tilde{\xi}^h(\cdot)$ which approximates the solution of (5.1a). A reasonable approximation to the cost function (5.1b) is

$$\tilde{W}^h(x, u^h) = \sum_{n=0}^{\infty} e^{-\beta t_n^h} k(\tilde{\xi}_n^h, u_n^h) \Delta \tilde{t}_n^h.$$

The dynamic programming equation for the optimal cost is

$$\tilde{V}^h(x) = \min_{\alpha \in \mathcal{U}} \left[e^{-\beta \Delta t^h(x, \alpha)} \tilde{V}^h(x + b(x, \alpha) \Delta t^h(x, \alpha)) + k(x, \alpha) \Delta t^h(x, \alpha) \right]. \quad (5.2)$$

Interpreting (5.2) in Terms of an Approximating Markov Chain. We will approximate (5.2) by a type of finite element method. The problem is illustrated in Figure 4.2.

In the figure, the sides of the triangles are $O(h)$. We take an approximation to $\tilde{V}^h(\cdot)$ of the following form: Approximate the optimal cost function

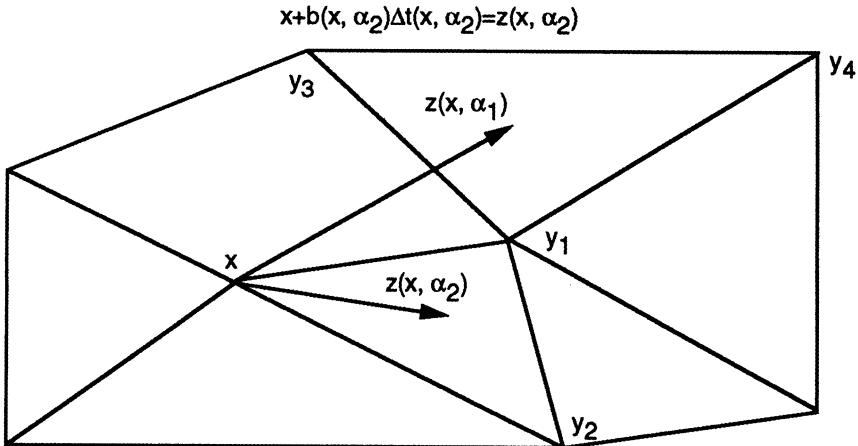


Figure 4.2. A piecewise linear approximation.

by a continuous function which is linear in each triangle of the figure. We will show that this leads directly to a Markov chain interpretation of (5.2), and that a simple probabilistic method can be used to prove the convergence of $\tilde{V}^h(x)$ to $V(x)$, as well as to get an approximate solution to (5.2). Let $z(x, \alpha)$ denote the point $x + b(x, \alpha)\Delta t^h(x, \alpha)$ which is reachable from x in time $\Delta t^h(x, \alpha)$ under control α . Refer to the figure, where the values are plotted for two values of α . Let $Y^h(x, \alpha)$ denote the corners of the triangle in which $z(x, \alpha)$ falls. For example, in the figure $Y^h(x, \alpha_2) = \{x, y_1, y_2\}$. We can represent the point $z(x, \alpha)$ as a convex combination of the points in $Y^h(x, \alpha)$. Let $p^h(x, y|\alpha)$ denote the weights used for the convexification. These weights are non-negative and sum to unity. Hence, they can be considered to be transition probabilities for a controlled Markov chain whose state space is just the set of all corner points in the figure.

Now let $V^h(\cdot)$ denote the *continuous piecewise linear approximation* to $\tilde{V}^h(\cdot)$ and rewrite (5.2) in terms of the transition probabilities as follows:

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \left[e^{-\beta \Delta t^h(x, \alpha)} \sum_{y \in Y^h(x, \alpha)} p^h(x, y|\alpha) V^h(y) + k(x, \alpha) \Delta t^h(x, \alpha) \right]. \quad (5.3)$$

Comparing (5.3) to (2.4), it is seen that the finite element approximation yields a solution which is just an optimal cost function for a controlled Markov chain. We will next show, via a straightforward probabilistic method, that $V^h(x)$ converges to $V(x)$ as $h \rightarrow 0$.

Let $\{\xi_n^h, n < \infty\}$ denote the approximating controlled chain and $\{\bar{u}_n^h, n < \infty\}$ the actual random variables which are the optimal control actions. We need not know what they are, but they can be obtained from a solution of (5.3). In particular, let $\bar{u}^h(x)$ be a minimizing value of α in (5.3). Then

$\bar{u}_n^h = \bar{u}^h(\xi_n^h)$. Define $\Delta t_n^h = \Delta t^h(\xi_n^h, \bar{u}_n^h)$ and redefine $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$. Then we can write

$$E[\xi_{n+1}^h - \xi_n^h | \xi_n^h = x, \bar{u}_n^h = \alpha] = b(x, \alpha) \Delta t^h(x, \alpha) = O(h), \quad (5.4a)$$

$$\text{cov}[\xi_{n+1}^h - \xi_n^h | \xi_n^h = x, \bar{u}_n^h = \alpha] = O(h^2). \quad (5.4b)$$

Convergence of the Approximations $\xi^h(\cdot)$ and $V^h(x)$: Part (a). Let E_n^h denote the expectation, given the state and control actions up to and including time n . In order to see the resemblance of the stochastic problem to the original deterministic problem, rewrite $\xi^h(\cdot)$ in the more convenient form

$$\xi^h(t) = x + \sum_{n: t_{n+1}^h \leq t} \Delta \xi_n^h.$$

Write $\Delta \xi_n^h = \xi_{n+1}^h - \xi_n^h$ in the form $\Delta \xi_n^h = E_n^h \Delta \xi_n^h + (\Delta \xi_n^h - E_n^h \Delta \xi_n^h)$. Then

$$\xi^h(t) = x + \sum_{n: t_{n+1}^h \leq t} E_n^h \Delta \xi_n^h + \sum_{n: t_{n+1}^h \leq t} (\Delta \xi_n^h - E_n^h \Delta \xi_n^h). \quad (5.5)$$

We will next show that only the “mean increments” (the first sum on the right) of $\xi^h(\cdot)$ are important and that there is some control such that $\xi^h(\cdot)$ is a good approximation to a solution to (5.1a) under that control.

The right hand sum in (5.5) is a continuous time interpolation of a martingale. By (5.4b), its variance is $E \sum_{n: t_{n+1}^h \leq t} \Delta t_n^h O(h) = O(h)t$. By (1.1.3), this implies that for any $t < \infty$

$$E \sup_{t \leq T} \left| \sum_{n: t_{n+1}^h \leq t} (\Delta \xi_n^h - E_n^h \Delta \xi_n^h) \right|^2 = O(h).$$

Thus the effects of that right hand term disappear in the limit. The basic reason for this is that the spatial *and* the temporal scales are both of the order of h .

We write the right hand sum in (5.5) simply as $O(h)$, and this will be the order of that term for approximations of deterministic problems in general. Define the continuous parameter interpolation $\bar{u}^h(\cdot)$ of $\{\bar{u}_n^h, n < \infty\}$ by $\bar{u}^h(t) = \bar{u}_n^h$ on the interval $[t_{n+1}^h, t_n^h]$. Now, using (5.4a), we have

$$\begin{aligned} \xi^h(t) &= x + \sum_{n: t_{n+1}^h \leq t} E_n^h \Delta \xi_n^h + O(h) \\ &= x + \int_0^t b(\xi^h(s), \bar{u}^h(s)) ds + O(h). \end{aligned} \quad (5.6)$$

We now proceed to show that there is some admissible control such that the paths of $\xi^h(\cdot)$ are actually good approximations to a solution of (5.1a)

under that control. Because $\Delta\xi_n^h = O(h)$ and $\Delta t_n^h \geq k_1 h$, the (piecewise linear interpolations of the) paths of the process $\xi^h(\cdot)$ are equicontinuous (in ω and h). Thus, for each fixed value of the probability space variable ω , each subsequence of $\{\xi^h(\cdot)\}$ has a further subsequence which converges to some limit uniformly on each bounded time interval. Suppose that the same were true of the sequence of interpolated control paths $\bar{u}^h(\cdot)$. We note now, for future reference, that this is hard to guarantee. This will be the reason for the introduction of an expanded class of admissible controls in Section 4.6 below. However, until further notice, we do proceed under the assumption that the paths of the control processes are equicontinuous in h and ω . Even if this assumption is not true, the convergence result for $V^h(x)$ will remain true. The details for the general case will be given in Section 4.6. Continuing, fix the sample space variable ω and let $h_n(\omega)$ index a convergent subsequence of (the piecewise linear interpolations of) $\{\xi^h(\cdot), \bar{u}^h(\cdot)\}$, with limit denoted by $x(\cdot, \omega), u(\cdot, \omega)$. By the convergence and the compactness of \mathcal{U} , we have $u(t, \omega) \in \mathcal{U}$. The uniform convergence implies that

$$x(t, \omega) = x + \int_0^t b(x(s, \omega), u(s, \omega)) ds. \quad (5.7)$$

Thus, the limit path satisfies the original ODE (5.1a) with an admissible control. Also, it is easily seen that $V^{h_n(\omega)}(x)$ converges to $W(x, u(\omega))$. Due to the minimality of $V(x)$, we have

$$W(x, u(\omega)) \geq V(x). \quad (5.8)$$

Because this holds for each ω ,

$$\liminf_h V^h(x) \geq V(x). \quad (5.9)$$

Convergence of $V^h(x)$ to $V(x)$: Part (b). In order to get the desired convergence result, we need to get the reverse inequality to (5.9). In order to do this, we will jump ahead a little and assume a result which will be discussed in Section 4.6: Namely, that there exists an optimal admissible control for the problem (5.1) and that the control can be arbitrarily well approximated by a control which is piecewise constant. That is, given any $\epsilon > 0$, there is an ϵ -optimal control $u^\epsilon(\cdot)$ of the following form: There is $\delta > 0$ and a finite number of points \mathcal{U}_ϵ in \mathcal{U} such that $u^\epsilon(\cdot)$ is \mathcal{U}_ϵ -valued and is constant on the intervals $[i\delta, i\delta + \delta]$. This fact remains true when working with the expanded class of controls which will be introduced later. We will next apply this ϵ -optimal control to the Markov chain and use the minimality of $V^h(x)$ for the controlled Markov chain problem to get the reverse inequality to (5.9).

The procedure is as follows. Fix $\epsilon > 0$. Define a sequence $u^{h,\epsilon}(\cdot)$ of controls for the chain by adapting the above ϵ -optimal control $u^\epsilon(\cdot)$ in the following natural way: Let $\{\xi_n^{h,\epsilon}, n < \infty\}$ denote the chain associated

with this new (to be defined) control. Let h be small enough such that $\delta > \inf_{x,\alpha} \Delta t^h(x, \alpha)$. Define the sequences $u_n^{h,\epsilon}$, $\Delta t_n^h = \Delta t^h(\xi_n^{h,\epsilon}, u_n^{h,\epsilon})$, and $t_n^h = \sum_0^{n-1} \Delta t_i^h$, recursively by $u_n^{h,\epsilon} = u^\epsilon(i\delta)$ for all n such that $t_n^h \in [i\delta, i\delta + \delta]$, for each i . The constructed control is an admissible control for the Markov chain. Let $\xi^{h,\epsilon}(\cdot)$ and $u^{h,\epsilon}(\cdot)$ denote the continuous parameter interpolation [interpolation intervals Δt_n^h]. Note that $u^{h,\epsilon}(\cdot)$ converges to $u^\epsilon(\cdot)$, except possibly at the points $i\delta$. Fix ω and choose a convergent subsequence of $\xi^{h,\epsilon}(\cdot)$ (the sequence need not be the same for each ω). Let $x^\epsilon(\cdot, \omega)$ denote the limit. Then, following the analysis which led to (5.7), we get that

$$x^\epsilon(t, \omega) = x + \int_0^t b(x^\epsilon(s, \omega), u^\epsilon(s)) ds.$$

The limit paths $x^\epsilon(\cdot, \omega)$ are all the same, irrespective of the chosen subsequence or of ω , because the solution to (5.1a) is unique under the chosen control $u^\epsilon(\cdot)$. This implies that the sequence $W^h(x, u^{h,\epsilon})$ converges to the ϵ -optimal cost $W(x, u^\epsilon)$. Now, using the optimality of $V^h(x)$, we have

$$\limsup_h V^h(x) \leq \lim_h W^h(x, u^{h,\epsilon}) = W(x, u^\epsilon) \leq V(x) + \epsilon. \quad (5.10)$$

Inequalities (5.9) and (5.10) imply that $V^h(x) \rightarrow V(x)$.

Remark on the Analogous Argument for a Stochastic Problem. For the general stochastic problem of Section 4.1, we will not be able to duplicate the above proof by choosing a convergent subsequence for each ω . Even if there is equicontinuity, there will be no guarantee that the limit is actually a sample value of a random variable. But somewhat more general compactness methods can be used there. These will be developed in Chapter 9 and used in Chapter 10.

We now turn to the general case for the deterministic problem where the control sequences $u^h(\cdot)$ are not “nice enough” for there to exist the convergent subsequences, and show how this case can be handled. This requires a diversion, but the concepts will also be of use for the general stochastic problem.

4.6 Deterministic Relaxed Controls

There need not exist an optimal control for problems such as (5.1) in the sense that there is a \mathcal{U} -valued measurable function $\bar{u}(\cdot)$ such that

$$\inf_u W(x, u) = V(x) = W(x, \bar{u}),$$

where the infimum is over all the \mathcal{U} -valued measurable functions. See the example below. Our primary concern is with getting a good approximation

to $V(\cdot)$ and with feedback controls which yield costs which are close to the infimum. The numerical methods will give feedback controls, but in order to be able to prove that the values of the costs which are given by the numerical algorithms converge to the infimum of the costs as the approximation parameter h converges to zero, we need to know that there is an optimal control in some appropriate sense. In particular, as seen in the argument at the end of the last section, we will need to know that there is a reasonable class of admissible controls, and an optimal control in that class which can be approximated by a “nice” piecewise constant control, with arbitrarily small penalty. The class of *relaxed controls* to be defined below was introduced for just such a purpose [9, 122]. In particular,

$$\inf_{\text{relaxed controls}} W(x, u) = \inf_{\text{ordinary controls}} W(x, u).$$

The class of relaxed controls will be of importance in the proofs only. They do not enter into the numerical algorithms or into actual applications.

Example of Nonexistence of an Optimal Control. For motivation, consider the following artificial example: $\mathcal{U} = [-1, 1]$, $\beta > 0$,

$$\dot{x}(t) = b(x(t), u(t)) = u(t),$$

$$W(x, u) = \int_0^\infty e^{-\beta t} [x^2(t) + (u^2(t) - 1)^2] dt. \quad (6.1)$$

Note that $V(0) = 0$. To see this, define the sequence of controls $u^n(\cdot)$ by

$$u^n(t) = (-1)^k \text{ on } [k/n, (k+1)/n], \quad k = 0, 1, \dots$$

It is not hard to see that $W(0, u^n) \rightarrow 0$ as $n \rightarrow \infty$. In a sense, when $x(0) = 0$, the “optimal control wants to take values ± 1 simultaneously.” But there is no optimal control in the usual sense.

The example is admittedly artificial, but more realistic examples arise when the control value space is not “rich enough”; more particularly, when the set of values $(b(x, \mathcal{U}), k(x, \mathcal{U}))$ is not convex for each x . As will be seen, by definition the relaxed control is a measure on the control value space, and allows all values of \mathcal{U} to be used “simultaneously” with appropriate weights. In preparation for the definition, we next show how to represent an ordinary control as a measure.

The Representation of a Control as a Measure. Let $u(\cdot)$ be an admissible control and let $\mathcal{B}(\mathcal{U})$ and $\mathcal{B}(\mathcal{U} \times [0, \infty))$ denote the σ -algebras over the Borel sets in \mathcal{U} and $\mathcal{U} \times [0, \infty)$, resp. Define the measures $m_t(\cdot)$ on $\mathcal{B}(\mathcal{U})$ and $m(\cdot)$ on $\mathcal{B}(\mathcal{U} \times [0, \infty))$ by

$$\begin{aligned} m_t(A) &= I_A(u(t)), \\ m(A \times [0, t]) &= \int_0^t m_s(A) ds. \end{aligned} \quad (6.2)$$

We can now write (5.1) as

$$\begin{aligned}\dot{x}(t) &= \int_{\mathcal{U}} b(x(t), \alpha) m_t(d\alpha), \\ x(t) &= x + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m_s(d\alpha) ds = x + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m(d\alpha ds), \\ W(x, m) &= \int_0^\infty \int_{\mathcal{U}} e^{-\beta s} k(x(s), \alpha) m(d\alpha ds).\end{aligned}\tag{6.3}$$

$m(A \times [0, t])$ is just the total integrated time over the interval $[0, t]$ that the control $u(\cdot)$ takes values in the set $A \subset \mathcal{U}$. A relaxed control is just a generalization of such $m(\cdot)$ and we now give the general definition.

Definition: Relaxed Control. An *admissible relaxed control* or simply a *relaxed control* $m(\cdot)$ is a measure on $\mathcal{B}(\mathcal{U} \times [0, \infty))$ such that $m(\mathcal{U} \times [0, t]) = t$ for all t . Given a relaxed control $m(\cdot)$, there is a derivative $m_t(\cdot)$ such that $m(d\alpha dt) = m_t(d\alpha)dt$. In fact, we can define the derivative by

$$m_t(A) = \lim_{\delta \rightarrow 0} \frac{m(A \times [t - \delta, t])}{\delta}.$$

Example. Let $\alpha_i \in \mathcal{U}, i = 1, 2$, and let $m_t(\cdot)$ be the measure which takes the value $1/2$ at each of the points α_i . Then the ODE in (6.3) can be written as:

$$\dot{x} = \frac{1}{2}[b(x, \alpha_1) + b(x, \alpha_2)] = \int_{\mathcal{U}} b(x, \alpha) m_t(d\alpha).\tag{6.4}$$

With the use of a relaxed control, the set of possible velocities and cost rates $(b(x, \mathcal{U}), k(x, \mathcal{U}))$ is replaced by its convex hull. The relaxed control $m(\cdot)$ for which $W(0, m) = 0$ and $x(t) = 0$ in (6.1) is the one for which $m_t(\cdot)$ is concentrated on the points ± 1 each with mass $1/2$ for all t . ■

Remarks on Relaxed Controls. To see why relaxed controls are needed when working with a sequence of controls and corresponding solutions to (5.1a), consider the following example. Let $u^n(t) = (-1)^{[nt]}$, where $[nt]$ denotes the integer part of nt , and let $m^n(\cdot)$ denote its relaxed control representation, as in (6.2). That is, the measure $m_t^n(\cdot)$ is concentrated at the point $u^n(t)$. The sequence of controls $\{u^n(\cdot), n < \infty\}$ does not converge in any of the usual senses. The sequence $\{m^n(\cdot), n < \infty\}$ converges in the “weak” sense; i.e., for any continuous function $\phi(\cdot)$ and $t < \infty$,

$$\int_0^t \int_{\mathcal{U}} \phi(\alpha, s) m^n(d\alpha ds) \rightarrow \int_0^t \int_{\mathcal{U}} \phi(\alpha, s) m(d\alpha ds),\tag{6.5}$$

where $m(\cdot)$ is the relaxed control with derivative $m_t(-1) = m_t(1) = 1/2$. If $x^n(\cdot)$ is the solution to (5.1a) under control $u^n(\cdot)$, then $x^n(\cdot)$ converges to $x(\cdot)$ which satisfies (6.4) with $\alpha_1 = 1$ and $\alpha_2 = -1$. In fact, for this example, $m^n([0, t] \times \{+1\})$ is just the total part of the time interval $[0, t]$ on which $u^n(s) = 1$, and it equals $t/2 + O(1/n)$.

For any sequence of relaxed admissible controls, there is always a subsequence which converges in the sense of (6.5). The introduction of relaxed controls has the effect of making the control appear essentially linearly in the dynamics and cost function.

We saw in an example above that there might not be an optimal control in the class of ordinary admissible controls. But there always is one in the class of relaxed controls. More detail is in Chapters 9 and 10. The following approximation result is important in applications because it says that any admissible relaxed control can be well approximated by a “nice” ordinary admissible control.

Approximation Theorem: The “Chattering” Theorem. *Let the ODE in (6.3) have a unique solution under some given relaxed control $m(\cdot)$, and let $k(\cdot)$ and $b(\cdot)$ be bounded and continuous. For any $\epsilon > 0$ and $T < \infty$, there is a $\delta > 0$, a finite set $\mathcal{U}_\epsilon \subset \mathcal{U}$, and a \mathcal{U}_ϵ -valued ordinary control $u^\epsilon(\cdot)$ which is constant on the intervals $[i\delta, i\delta + \delta)$ and is such that*

$$\begin{aligned} \sup_{t \leq T} |x(t, m) - x(t, u^\epsilon)| &\leq \epsilon, \\ |W(x, m) - W(x, u^\epsilon)| &\leq \epsilon. \end{aligned} \tag{6.6}$$

Here $x(t, m)$ is the solution under the denoted control $m(\cdot)$.

Completion of the Argument of Section 4.5. Even if the sequence of controls $\bar{u}^h(\cdot)$ used in Section 4.4 does not have a convergent subsequence for each ω , the sequence of its relaxed control representations $m^h(\cdot)$ always will. Furthermore any relaxed control (optimal or ϵ -optimal) can be arbitrarily well approximated by a piecewise constant ordinary control. In addition, the infimum of the cost over the classes of ordinary and relaxed controls are the same. With these points in mind, we can complete the proof of convergence in Section 4.5 in general.

5

Construction of the Approximating Markov Chain

In this chapter we develop some canonical methods for obtaining approximating Markov chains which are locally consistent with the controlled diffusion in the sense used in (4.1.3). We also deal with the controlled jump diffusion and reflected processes in Sections 5.6 and 5.7. The chapter outlines two classes of basic methods and a number of variations of each of them. One purpose is to describe methods which are readily programmable. But we also wish to show the versatility and intuitive nature of the general approach. There are many variations of the methods discussed. Once the general procedures are clear, the reader can adapt them to particular problems which might not be directly covered by the discussion. The development does not directly cover processes on manifolds such as the surface of a sphere or torus, but the various possibilities should be apparent.

The first method to be discussed is called a “finite difference” method, although its validity does not depend on the validity of any finite difference approach to the solution of partial differential equations. The finite difference approximations are used as guides to the construction of locally consistent approximating Markov chains only. It turns out that when a carefully chosen finite difference approximation is applied to the differential operator of the controlled process, the coefficients of the resulting discrete equation can serve as the desired transition probabilities and interpolation interval. Once these are available, we use purely probabilistic methods for dealing with them, because the partial differential equations which are approximated might not have the smoothness which is required for validity of the classical finite difference approach. Also, once the transition probabilities are obtained with this method, they can be altered in many ways, according to numerical convenience, while keeping consistency. In order to motivate the method, we start by discussing several special examples in Section 5.1. Various numerical implications and variations of the choices are discussed at the end of the section and in Section 5.2. The general “finite difference” method is described in Section 5.3. The method, as developed in its simplest form, sometimes fails if the off diagonal terms of the noise covariance matrix are too large. Several methods for overcoming this problem are discussed in Section 5.3. The notion of local consistency of an approximating chain is analogous to the notion of consistency for a

finite difference approximation, except that here it is the controlled process which is being approximated, and not a partial differential equation.

Section 5.4 describes a very versatile “direct” method for getting good approximating chains, essentially by decomposing the “local” effects of the “drift” and “noise” parts of the original process. It is a generalization of the approach of Sections 5.1 to 5.3. Several illustrations show how the state space and transition functions of the approximating chains can be tailored to the particular problem.

Section 5.5 contains some introductory comments concerning variable grids. The spacing of the points in the approximating chain might vary from one region of the state space to another because we need greater accuracy in certain regions than in others. Because of this, we might lose local consistency on the boundaries of the regions. A particular example of a “variable grid” and the lack of local consistency is discussed. The probabilistic approach to the convergence proofs works when the approximated process spends little time in a neighborhood of the points where we lack consistency. This covers many cases of interest. The example illustrates the versatility of the method. The approximating chain for the jump diffusion process is discussed in Section 5.6. The approximating chains are derived by simply piecing together in an obvious way the approximation for the diffusion alone with an approximation to the jumps. Approximations for reflecting boundaries are dealt with in Section 5.7. The Skorokhod Problem model for a reflecting jump diffusion is used. The essential requirement for the approximating chain on the reflecting boundary is that the mean direction of motion be an approximation to a reflection direction for the original problem. In Section 5.8, the dynamic programming equations for the problems of Chapter 2 are written in the notation of the approximating Markov chains for use in subsequent chapters.

5.1 Finite Difference Type Approximations: One Dimensional Examples

In this and in the next three sections, several convenient methods for getting locally consistent approximating chains will be demonstrated. As with any method of numerical approximation, once the general concepts are understood one can use ingenuity to exploit the unique features of each special case. It is useful to start with a discussion of the relation between the classical finite difference approximation method for elliptic PDE’s and the Markov chain approximation method, so that we can see that we are actually on familiar ground. It will be seen that an essentially standard use of a finite difference approximation of the differential operator of the controlled process will yield the transition probabilities and interpolation interval as the coefficients in the finite difference representation. The general

idea of the approach via a finite difference method can be best introduced by means of a few simple one dimensional examples. These will illustrate the essentially automatic nature of the generation of the transition probabilities $p^h(x, y|\alpha)$ of the chain and the interpolation intervals $\Delta t^h(x, \alpha)$. In these motivating examples, all the stochastic and partial differential equations and the uses of Itô's formula will be dealt with in a purely formal way. We are only concerned with getting a locally consistent chain. Once the formalities of the "finite difference" derivation are over, it can be readily verified that the derived chain satisfies the required consistency properties.

Recall that the controlled chain $\{\xi_n^h, n < \infty\}$ is said to be *locally consistent* with the controlled diffusion process

$$dx = b(x, u)dt + \sigma(x)dw \quad (1.1)$$

if (4.1.3) holds for interpolation intervals satisfying $\sup_{x, \alpha} \Delta t^h(x, \alpha) \rightarrow 0$, as $h \rightarrow 0$. For these one dimensional examples, the process $x(\cdot)$ will be of interest on the interval $G = [0, B]$, $B > 0$. It is not really necessary to specify the interval, but it might help to fix ideas for the actual control problem. We will use $h > 0$ as the approximation parameter, and suppose that B is an integral multiple of h for all values of h of interest. Define the sets $S_h = \mathbb{R}_h = \{0, \pm h, \pm 2h, \dots\}$, and $G_h^0 = G^0 \cap S_h$, where G^0 is the interior of G .

Example 1. An Uncontrolled Wiener Process. Define the uncontrolled process $x(t) = x + \sigma w(t)$, where $w(\cdot)$ is a standard real valued Wiener process, x is the initial condition and σ is a constant. Define the first exit time $\tau = \min\{t : x(t) \notin (0, B)\}$ and the cost functional

$$W(x) = E_x \int_0^\tau k(x(s))ds,$$

where $k(\cdot)$ is a bounded and continuous function. This function, as well as the interval G , play only an auxiliary and formal role in the development. Their values are not important. By Itô's formula, if $W(\cdot)$ is smooth enough, then it satisfies the differential equation (Section 3.3)

$$\mathcal{L}W(x) + k(x) = 0, \quad x \in (0, B), \quad (1.2)$$

$$W(0) = W(B) = 0,$$

where $\mathcal{L} = (\sigma^2/2)(d^2/dx^2)$ is the differential operator of the process $x(\cdot)$.

We will obtain the desired transition probabilities and interpolation interval simply by trying to solve the differential equation (1.2) by finite differences. The standard approximation

$$f_{xx}(x) \rightarrow \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} \quad (1.3)$$

for the second derivative will be used. Now use (1.3) in (1.2), denote the result by $W^h(x)$, and get (for $x \in G_h^0$)

$$\frac{\sigma^2}{2} \frac{[W^h(x+h) + W^h(x-h) - 2W^h(x)]}{h^2} + k(x) = 0,$$

which can be rewritten as

$$W^h(x) = \frac{1}{2}W^h(x+h) + \frac{1}{2}W^h(x-h) + \frac{h^2}{\sigma^2}k(x), \quad x \in G_h^0. \quad (1.4a)$$

The boundary conditions are

$$W^h(0) = W^h(B) = 0. \quad (1.4b)$$

Equation (1.4) has a simple interpretation in terms of a Markov chain (recall a similar problem in Section 4.4). Let $\{\xi_n^h, n < \infty\}$ be the symmetric random walk on the state space S_h and define $\Delta t^h = h^2/\sigma^2$. The transition probabilities of the random walk are $p^h(x, x \pm h) = 1/2$. In terms of these transition probabilities, we can rewrite (1.4a) as

$$W^h(x) = p^h(x, x+h)W^h(x+h) + p^h(x, x-h)W^h(x-h) + k(x)\Delta t^h, \quad (1.5)$$

$x \in G_h^0$. Using N_h for the first escape time of the chain from the set G_h^0 , the solution of (1.5) can be written as (see Chapter 2) a functional of the path of the chain in the following way:

$$W^h(x) = E_x \sum_{n=0}^{N_h-1} k(\xi_n^h)\Delta t^h. \quad (1.6)$$

Note that $E_n^h \Delta \xi_n^h = 0$ and the conditional covariance is $E_n^h (\Delta \xi_n^h)^2 = h^2 = \sigma^2 \Delta t^h$. Thus, the chain is locally consistent with the Wiener process, in the sense of (4.1.3). The use of a random walk to approximate a Wiener process is well known. The continuous time interpolations $\xi^h(\cdot)$ (Section 4.2) and $\psi^h(\cdot)$ (Section 4.3) both converge in distribution (Chapters 9 and 10) to a Wiener process. In fact, Theorem 10.5.1 implies that $W^h(x) \rightarrow W(x)$. The convergence result is also a consequence of the Donsker invariance principle [42]. The finite difference approximation of the differential operator \mathcal{L} of the Wiener process gives the transition probabilities and the interpolation interval immediately, whether or not $W(x)$ actually satisfies (1.2). [In this particular case, (1.2) does hold.] The partial differential equation (1.2) and the function $k(\cdot)$ only served the formal and auxiliary purpose of being a vehicle for the calculation of the transition probabilities and the interpolation interval. The transition probabilities were found by choosing an appropriate finite difference approximation to (1.2), collecting terms, and dividing all terms in the resulting expression by the coefficient of $W^h(x)$.

Example 2. An Uncontrolled Deterministic Case. This example uses an uncontrolled ordinary differential equation model. Consider the system $\dot{x} = b(x)$, where x is real valued and $b(\cdot)$ is bounded and continuous. First, suppose that $\inf_x |b(x)| \neq 0$. As in Example 1, define the functional $W(x) = \int_0^x k(x(s))ds$. Formally, $W(\cdot)$ solves the ODE (1.2), where now $\mathcal{L} = b(x)d/dx$ is just a first order derivative operator. In particular, if $W(\cdot)$ is smooth enough we have

$$\begin{aligned} W_x(x)b(x) + k(x) &= 0, \quad x \in (0, B), \\ W(0) = W(B) &= 0. \end{aligned} \tag{1.7}$$

Owing to the unique direction of flow for each initial condition, only one of the two boundary conditions is relevant for all $x \in (0, B)$. As was done with Example 1, the Markov chain approximation to $x(\cdot)$ will be obtained by use of a finite difference approximation to the derivative in (1.7). But we will need to choose the difference approximation to the derivative $W_x(x)$ carefully, if the finite difference equation is to have an interpretation in terms of a Markov chain. Define the one sided difference approximations:

$$\begin{aligned} f_x(x) &\rightarrow \frac{f(x+h) - f(x)}{h} && \text{if } b(x) \geq 0, \\ f_x(x) &\rightarrow \frac{f(x) - f(x-h)}{h} && \text{if } b(x) < 0. \end{aligned} \tag{1.8}$$

That is, if the velocity at a point is non-negative, then use the forward difference, and if the velocity at a point is negative, then use the backward difference.

Such schemes are known as the “upwind” approximation method in numerical analysis. Define the *positive and negative parts* of a real number by: $a^+ = \max[a, 0]$, $a^- = \max[-a, 0]$. Using (1.8) in (1.7) yields

$$\frac{W^h(x+h) - W^h(x)}{h} b^+(x) - \frac{W^h(x) - W^h(x-h)}{h} b^-(x) + k(x) = 0, \tag{1.9}$$

for $x \in (0, B)$ and with boundary conditions $W^h(0) = W^h(B) = 0$. Define the functions

$$p^h(x, x+h) = I_{\{b(x) \geq 0\}}, \quad p^h(x, x-h) = I_{\{b(x) < 0\}}$$

and the interpolation interval $\Delta t^h(x) = h/|b(x)|$. The $p^h(x, x \pm h)$ are transition probabilities for a Markov chain on the state space S_h . Now, collecting terms in (1.9), noting that $b^+(x) + b^-(x) = |b(x)|$, and dividing by the coefficient of $W^h(x)$ yields, for $x \in G_h^0$,

$$\begin{aligned} W^h(x) &= W^h(x+h) \frac{b^+(x)}{|b(x)|} + W^h(x-h) \frac{b^-(x)}{|b(x)|} + k(x) \frac{h}{|b(x)|} \\ &= W^h(x+h)p^h(x, x+h) + W^h(x-h)p^h(x, x-h) \\ &\quad + k(x)\Delta t^h(x). \end{aligned} \tag{1.10}$$

The $p^h(x, y)$ define a Markov chain $\{\xi_n^h, n < \infty\}$ on S_h . If $\inf_x |b(x)| \neq 0$, then the chain together with the interpolation interval $h/|b(x)|$ is locally consistent with the “process” defined by the solution to $\dot{x} = b(x)$ in that (4.1.3) holds: In particular, $E_{x,n}^h \Delta \xi_n^h = b(x) \Delta t^h(x)$ and $\text{cov}_{x,n}^h \Delta \xi_n^h = O(h^2) = o(\Delta t^h(x))$.

A Modification if $\inf_x |b(x)| = 0$. If $\inf_x |b(x)| = 0$, then the above calculation breaks down in that $\Delta t^h(x) = \infty$ at some point x . Because that state is absorbing anyway, the infinite value is not surprising. If we wish to use interpolation intervals which go to zero as $h \rightarrow 0$, the degeneracy can be circumvented by simply allowing transitions of the states of the chain to themselves. To see how this can be done, let $0 \leq p^h(x, x) \leq 1$ and rewrite (1.10) as

$$\begin{aligned} W^h(x) &= p^h(x, x)W^h(x) + (1 - p^h(x, x))\frac{b^+(x)}{|b(x)|}W^h(x + h) \\ &\quad + (1 - p^h(x, x))\frac{b^-(x)}{|b(x)|}W^h(x - h) + (1 - p^h(x, x))\Delta t^h(x)k(x). \end{aligned} \tag{1.10'}$$

Define the new transition probabilities

$$p^h(x, x \pm h) = (1 - p^h(x, x))(b^\pm(x)/|b(x)|),$$

and the new interpolation interval $\Delta t^h(x) = (1 - p^h(x, x))h/|b(x)|$. One can now readily choose the $p^h(x, x)$ to get a locally consistent transition probability and interpolation interval. For example, if $b(x) = 0$, then set $p^h(x, x) = 1$ and $\Delta t^h(x) = h$. Note that (1.10) and (1.10') have the same solution if $p^h(x, x) < 1$ for all x .

The Solution to (1.10'). In order to facilitate writing the solution to (1.10'), suppose that there is a point x_0 such that for $x > x_0$, we have $b(x) > 0$, and for $x < x_0$ we have $b(x) < 0$, and $b(x_0) = 0$. Then, in terms of $\{\xi_n^h, n < \infty\}$ the solution to (1.10') can be written as

$$W^h(x) = E_x \sum_{n=0}^{N_h-1} k(\xi_n^h) \Delta t^h(\xi_n^h),$$

where N_h is again the first escape time of the chain from G_h^0 . If $x_0 \in (0, B)$, then the process might get stuck at x_0 with value $W^h(x_0) = 0$ if $k(x_0) = 0$, and $W^h(x_0) = \pm\infty$ otherwise, according to the sign of $k(x_0)$. It turns out that $W^h(x) \rightarrow W(x)$, and the interpolated processes $\xi^h(\cdot)$ and $\psi^h(\cdot)$ both converge to the solution to $\dot{x} = b(x)$. Thus, the finite difference method automatically gives a chain which satisfies the consistency conditions and can be used to approximate the original process as well as functionals of it.

Remark on the Choice of Finite Difference Approximation (1.8) and on $\Delta t^h(x)$. Note that the choice of the finite difference approximation

depends on the sign of $b(x)$. I.e., the “direction” of the approximation depends on the sign of $b(x)$ because we want the direction of movement of the chain at a point x to reflect the sign of the velocity at x . Such a choice is not unusual in numerical analysis, and will be commented on in Examples 3 and 4 also. Note also that the $\Delta t^h(x)$ is just the inverse of the absolute value of the velocity at x times the spatial difference interval. Thus, in time $\Delta t^h(x)$ the continuous time interpolations $\xi^h(\cdot)$ and $\psi^h(\cdot)$ move an average of h units.

Example 3. A Diffusion With Drift: No Control. Now we combine the two cases of Examples 1 and 2. Let $x(\cdot)$ be the process which satisfies

$$dx = b(x)dt + \sigma(x)dw, \quad x(0) = x, \quad (1.11)$$

where we assume that $b(\cdot)$ and $\sigma(\cdot)$ are bounded and continuous and

$$\inf_x (\sigma^2(x) + |b(x)|) > 0.$$

If this last restriction does not hold, then we can continue by adding transitions from appropriate states x to themselves as discussed in Example 2. Let $W(x)$ be defined as in Example 1. Then if $W(\cdot)$ is smooth enough, Itô’s formula implies (1.2), where \mathcal{L} is the differential operator of the process (1.11). In particular,

$$W_x(x)b(x) + W_{xx}(x)\sigma^2(x)/2 + k(x) = 0, \quad x \in (0, B), \quad (1.12)$$

$$W(0) = W(B) = 0.$$

Use the finite difference approximations (1.3) and (1.8), and again let $W^h(x)$ denote the finite difference approximation. [A possible alternative to (1.8) is given in (1.18) below.] Substituting these approximations into (1.12), collecting terms, multiplying by h^2 , and dividing all terms by the coefficient of $W^h(x)$ yields the approximating equation

$$\begin{aligned} W^h(x) &= \frac{\sigma^2(x)/2 + hb^+(x)}{\sigma^2(x) + h|b(x)|} W^h(x+h) + \frac{\sigma^2(x)/2 + hb^-(x)}{\sigma^2(x) + h|b(x)|} W^h(x-h) \\ &\quad + k(x) \frac{h^2}{\sigma^2(x) + h|b(x)|}. \end{aligned} \quad (1.13)$$

Let us rewrite this as

$$W^h(x) = p^h(x, x+h)W^h(x+h) + p^h(x, x-h)W^h(x-h) + k(x)\Delta t^h(x), \quad (1.14)$$

where the p^h and Δt^h are defined in the obvious manner. Let $p^h(x, y) = 0$ for $y \neq x \pm h$. The $p^h(x, y)$ are non-negative and sum (over y) to unity for each x . Thus, they can be considered to be transition probabilities for

a Markov chain on the state space S_h . Let $\{\xi_n^h, n < \infty\}$ denote the chain, and let us check for local consistency. We have

$$E_{x,n}^h \Delta \xi_n^h = h \frac{\sigma^2(x)/2 + hb^+(x)}{\sigma^2(x) + h|b(x)|} - h \frac{\sigma^2(x)/2 + hb^-(x)}{\sigma^2(x) + h|b(x)|} = b(x)\Delta t^h(x), \quad (1.15)$$

$$\begin{aligned} E_{x,n}^h (\Delta \xi_n^h)^2 &= h^2 \left(\frac{\sigma^2(x)/2 + hb^+(x)}{\sigma^2(x) + h|b(x)|} + \frac{\sigma^2(x)/2 + hb^-(x)}{\sigma^2(x) + h|b(x)|} \right) \\ &= \sigma^2(x)\Delta t^h(x) + o(\Delta t^h(x)). \end{aligned} \quad (1.16)$$

Also, $E_{x,n}^h [\Delta \xi_n^h - E_{x,n}^h \Delta \xi_n^h]^2$ has the representation (1.16) where $o(\Delta t^h(x)) = \Delta t^h(x)O(h)$. Thus the chain (with the given interpolation interval) is locally consistent with the process defined by (1.11). Again the finite difference approximations were used only in a mechanical way to get the transition probabilities and interpolation interval for a locally consistent approximating chain.

Note that the interpolation interval $\Delta t^h(x)$ depends on both the drift and the diffusion coefficients. As the magnitudes of these coefficients increase, the interval decreases in size. The relative effects of the two coefficients on the interpolation interval are affected by the two natural time scales: The time scale for the pure drift case (Example 2) is $O(h)$, and that for the pure diffusion case (Example 1) is $O(h^2)$.

A Central Difference Approximation. If the diffusion term always “dominates” the drift term for the values h of interest, then the one sided difference approximations (1.8) can be replaced by a two sided or symmetric finite difference approximation. This seems to be preferable in certain circumstances [101]. Suppose that

$$\inf_x [\sigma^2(x) - h|b(x)|] \geq 0 \quad (1.17)$$

for all h of interest, and use the symmetric difference approximation

$$f_x(x) \rightarrow \frac{f(x+h) - f(x-h)}{2h} \quad (1.18)$$

for the first derivative in lieu of (1.8). Then repeating the procedure which led to (1.14) yields the following finite difference equation and *new* transition probabilities and interpolation interval.

$$\begin{aligned} W^h(x) &= \frac{\sigma^2(x) + hb(x)}{2\sigma^2(x)} W^h(x+h) \\ &\quad + \frac{\sigma^2(x) - hb(x)}{2\sigma^2(x)} W^h(x-h) + k(x) \frac{h^2}{\sigma^2(x)} \\ &= p^h(x, x+h)W^h(x+h) + p^h(x, x-h)W^h(x-h) \\ &\quad + k(x)\Delta t^h(x). \end{aligned}$$

Local consistency can be shown as in (1.15) and (1.16).

If (1.17) does not hold, then the $p^h(x, x \pm h)$ in (1.18) are not all non-negative. Thus, they cannot serve as transition probabilities and, in fact, serious instability problems might arise.

A Modification to Allow a Symmetric Finite Difference Interval. The algorithm can be modified to allow use of (1.18) in lieu of (1.8) even if (1.17) fails. But this modification is deceptive because the result is (1.13). Replace (1.1) by

$$dx = b(x) + \sigma(x)dw + \sqrt{h|b(x)|}dw', \quad (1.19)$$

where $w'(\cdot)$ is a Wiener process which is independent of $w(\cdot)$. Now, we can use (1.18). The result is just the equation below (1.18) with $\sigma^2(x)$ replaced by $\sigma^2(x) + h|b(x)|$, which is just (1.13).

Example 4. A One Dimensional Example with Control. We now turn to a discussion of a controlled one dimensional diffusion. Let the control $u(\cdot)$ be of the feedback type, with values $u(x)$ in \mathcal{U} , a compact control set. Let $x(\cdot)$ be defined by

$$dx = b(x, u(x))dt + \sigma(x)dw. \quad (1.20)$$

Again, let $\tau = \min\{t : x(t) \notin (0, B)\}$ and define the cost function

$$W(x, u) = E_x^u \left[\int_0^\tau k(x(s), u(x(s)))ds + g(x(\tau)) \right],$$

$$W(x, u) = g(x), \text{ for } x = 0, B.$$

As in the previous examples, the functions $k(\cdot)$ and $g(\cdot)$ and the interval $[0, B]$ play a purely auxiliary role. Formally applying Itô's formula to the function $W(x, u)$ yields the equation (see Section 3.2)

$$\mathcal{L}^{u(x)} W(x, u) + k(x, u(x)) = 0, \quad x \in (0, B), \quad (1.21)$$

with boundary conditions $W(0, u) = g(0)$, $W(B, u) = g(B)$, where \mathcal{L}^α is the differential operator of $x(\cdot)$ when the control is fixed at α .

To get a locally consistent Markov chain and interpolation interval, we simply follow the procedure used in Example 3; namely, use the finite difference approximations (1.3) and (1.8) for the $W_{xx}(x, \alpha)$ and $W_x(x, \alpha)$ in (1.21). Define

$$\begin{aligned} p^h(x, x + h|\alpha) &= \frac{\sigma^2(x)/2 + hb^+(x, \alpha)}{\sigma^2(x) + h|b(x, \alpha)|}, \\ p^h(x, x - h|\alpha) &= \frac{\sigma^2(x)/2 + hb^-(x, \alpha)}{\sigma^2(x) + h|b(x, \alpha)|}, \end{aligned} \quad (1.22)$$

$$\Delta t^h(x, \alpha) = \frac{h^2}{\sigma^2(x) + h|b(x, \alpha)|}.$$

For $y \neq x \pm h$, set $p^h(x, y|\alpha) = 0$. Then the constructed p^h are transition probabilities for a controlled Markov chain. Local consistency of this chain and interpolation interval can be shown exactly as for Example 3. Also, by following the procedure which led to (1.13), we see that the formal finite difference form of (1.21) is just

$$W^h(x, u) = \sum_y p^h(x, y|u(x)) W^h(y, u) + k(x, u(x)) \Delta t^h(x, u(x)) \quad (1.23)$$

for $x \in G_h^0$ and with the same boundary conditions as $W(x, u)$ satisfies. If (1.23) has a unique solution, then it is the cost associated with the controlled chain, namely,

$$W^h(x, u) = E_x^u \sum_0^{N_h-1} k(\xi_n^h) \Delta t^h(\xi_n^h, u(\xi_n^h)).$$

The dynamic programming equation for the optimal value function is

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_y p^h(x, y|\alpha) V^h(y) + k(x, \alpha) \Delta t^h(x, \alpha) \right] \quad (1.24)$$

with the same boundary conditions as for $W(x, u)$.

We emphasize that no claim is made that the convergence of the finite difference approximations can be proved via the classical methods of numerical analysis. The finite difference approximation is used only to get the transition probabilities of a Markov chain which is locally consistent with (1.20).

5.2 Numerical Simplifications and Alternatives for Example 4

5.2.1 ELIMINATING THE CONTROL DEPENDENCE IN THE DENOMINATORS OF $p^h(x, y|\alpha)$ AND $\Delta t^h(x, \alpha)$

The possible presence of the control parameter α in the denominators of the expressions for the transition probabilities and interpolation interval in (1.22) will complicate getting the solution of (1.24) if it complicates the procedure needed to evaluate the minimum on the right hand side. All of the current methods used to solve (1.24) require the computation of minima analogous to the right hand side of (1.24). The *iteration in policy space method*, described in Chapter 6 and also briefly below, is one generally

popular approach. This method requires the solution of the sequence of equations for the costs associated with a “minimizing” sequence of controls, and each control is determined via an operation similar to the minimization on the right side of (1.24).

The ease of solution of the approximating equations, which are (1.23) and (1.24) in our case, is a key concern in deciding on the approximating chain. For any particular problem, there are many alterative locally consistent approximating chains which can be used, some being more convenient than others. One has to weigh the work needed to obtain the approximating chain against the amount of work needed to actually solve (1.23) or (1.24). Given the transition probabilities for any locally consistent chain, one can often simplify them from the computational point of view, while maintaining local consistency. Getting a useful chain is often just a matter of common sense. The intuition of numerical analysis is combined with the intuition obtained from the physical interpretation of the Markov chain approximations. We will illustrate some useful methods for simplifying the dependence of the transition probabilities and interpolation interval on the control. The remarks in this section deal with the Example 4 of the last section but are of general applicability.

One simple way of eliminating the dependence of the denominators in the expressions for $p^h(x, y|\alpha)$ and $\Delta t^h(x, \alpha)$ on the control parameter α starts with (1.22) and proceeds as follows. Define the functions $B(x) = \max_{\alpha \in \mathcal{U}} |b(x, \alpha)|$ and define

$$\bar{p}^h(x, x \pm h|\alpha) = \frac{\sigma^2(x)/2 + hb^\pm(x, \alpha)}{\sigma^2(x) + hB(x)}, \quad (2.1)$$

$$\Delta \bar{t}^h(x) = \frac{h^2}{\sigma^2(x) + hB(x)}.$$

The $\bar{p}^h(x, y|\alpha)$ might sum (over y) to less than unity for some x and α . To handle this, we define the “residual”

$$\bar{p}^h(x, x|\alpha) = 1 - \sum_{y \neq x} \bar{p}^h(x, y|\alpha) = \frac{h(B(x) - |b(x, \alpha)|)}{\sigma^2(x) + hB(x)}. \quad (2.2)$$

It can be readily shown that the chain associated with the transition probabilities $\bar{p}^h(x, y|\alpha)$ and interpolation interval $\Delta \bar{t}^h(x)$ is locally consistent with (1.20). The difference between the barred and the unbarred values is $O(h)$. The symmetric finite difference approximation (1.18) to the first derivative can also be used if

$$\min_x [\sigma^2(x) - hB(x)] \geq 0.$$

It will be seen below that the use of (2.1) and (2.2) is equivalent to the use of (1.22). They yield the same cost functions and minimum cost.

Controlled Variance. So far, we have not considered cases where the variance $\sigma(\cdot)$ might depend on the control parameter. The mathematics required for the convergence proofs is somewhat more complicated when there is such dependence. See [80] for more details. But the numerical algorithms are obtained in the same way as when only the drift $b(\cdot)$ is controlled. When the variance depends on the control, it is often necessary to carefully examine the computational complexity of taking minima in (1.24) or in (2.7) below. The method described above can also be used in this case, as follows by simply defining the new transition probabilities and transition intervals:

$$\begin{aligned} D^h(x) &= \max_{\alpha \in \mathcal{U}} [\sigma^2(x, \alpha) + h|b(x, \alpha)|], \\ \bar{p}^h(x, x \pm h|\alpha) &= \frac{\sigma^2(x, \alpha)/2 + hb^\pm(x, \alpha)}{D^h(x)}, \\ \bar{p}^h(x, x|\alpha) &= 1 - \sum_{y \neq x} \bar{p}^h(x, y|\alpha) = \frac{D^h(x) - [\sigma^2(x, \alpha) + h|b(x, \alpha)|]}{D^h(x)}, \\ \Delta \bar{t}^h(x) &= \frac{h^2}{D^h(x)}. \end{aligned}$$

5.2.2 A USEFUL NORMALIZATION IF $p^h(x, x|\alpha) \neq 0$

The numerical methods for (1.23) and (1.24) generally converge faster if $p^h(x, x|\alpha) = 0$. From an intuitive point of view, this is due to the observation that the positivity of the $p^h(x, x|\alpha)$ implies that the Markov chain “mixes” more slowly. Loosely speaking, the faster the mixing, the faster the convergence. Suppose that the $\bar{p}^h(x, y|\alpha)$ defined in (2.1) and (2.2) are used for the $p^h(x, y|\alpha)$ in (1.23) or (1.24). We have $\bar{p}^h(x, x|\alpha) = O(h) \neq 0$. The difference from zero is not large, but numerical studies suggest that even in this case there is a slight advantage to eliminating the transition of a state to itself. Let us see how this can be done.

The Equivalence of ((2.1), (2.2)) and (1.22). Let $\bar{W}^h(x, u)$ denote the cost under (2.1) and (2.2). Then

$$\begin{aligned} \bar{W}^h(x, u) &= \sum_{y \neq x} \bar{p}^h(x, y|u(x)) \bar{W}^h(y, u) + \bar{p}^h(x, x|u(x)) \bar{W}^h(x, u) \\ &\quad + k(x, u(x)) \Delta \bar{t}^h(x, u(x)), \end{aligned} \tag{2.3}$$

which can be written as

$$\bar{W}^h(x, u) = \frac{\sum_{y \neq x} \bar{p}^h(x, y|u(x)) \bar{W}^h(y, u) + k(x, u(x)) \Delta \bar{t}^h(x, u(x))}{1 - \bar{p}^h(x, x|u(x))}.$$

This equation equals

$$\bar{W}^h(x, u) = \sum_{y \neq x} p^h(x, y|u(x)) \bar{W}^h(y, u) + k(x, u(x)) \Delta t^h(x, u(x)). \quad (2.4)$$

Comparing this equation to (1.23), we see that $W^h(x, u) = \bar{W}^h(x, u)$ for all feedback controls $u(\cdot)$ for which (1.23) or (2.4) has a unique solution. Thus one can use either equation with the same results. This procedure for eliminating the transitions from any state to itself is called *normalization*.

These observations apply to any situation where we start with a transition probability under which some states communicate to themselves. Just to complete the picture, suppose that we are given a locally consistent chain with transition probabilities and interpolation interval $\tilde{p}^h(x, y|\alpha)$, $\Delta \tilde{t}^h(x, \alpha)$, respectively, where $\tilde{p}^h(x, x|\alpha)$ might be positive for some x and α . Let $\tilde{W}^h(x, u)$ denote the associated value of the cost. Then

$$\begin{aligned} \tilde{W}^h(x, u) &= \sum_{y \neq x} \tilde{p}^h(x, y|u(x)) \tilde{W}^h(y, u) + \tilde{p}^h(x, x|u(x)) \tilde{W}^h(x, u) \\ &\quad + k(x, u(x)) \Delta \tilde{t}^h(x, u(x)). \end{aligned} \quad (2.3')$$

For $x \neq y$, define

$$\begin{aligned} \hat{p}^h(x, y|\alpha) &= \frac{\tilde{p}^h(x, y|\alpha)}{1 - \tilde{p}^h(x, x|\alpha)}, \\ \hat{\Delta t}^h(x, \alpha) &= \frac{\Delta \tilde{t}^h(x, \alpha)}{1 - \tilde{p}^h(x, x|\alpha)}, \end{aligned} \quad (2.5)$$

and set $\hat{p}^h(x, x|\alpha) = 0$. The chain with this new transition probability and interpolation interval obviously yields the same cost as does the original chain [under the same control $u(\cdot)$] if the solution to (2.3') is unique, because by construction of the $\hat{p}^h(x, y|\alpha)$ and $\hat{\Delta t}^h(x)$, the equations for the cost are the same in both cases. It is readily verified that (2.5) yields a locally consistent chain.

A Comment on Normalization for the Discounted Cost Problem. If there is a discount rate $\beta > 0$, then we need to be more careful with the use of the type of transformations just discussed, because the discount factor also depends on the interpolation interval. The transformed transition probabilities (as in (2.5)) will still give a locally consistent Markov chain, of course, because consistency does not depend on whether or not there is discounting. But we should use the discount factor $e^{-\beta \hat{\Delta t}^h(x, \alpha)}$ (or an appropriate approximation) which corresponds to the new interpolation interval $\hat{\Delta t}^h(x, \alpha)$.

The Approximation in Policy Space Method. Introduction. Many computational methods for solving (1.24) are versions of the so-called *ap-*

proximation in policy space technique. (See Chapter 6 for more detail.) With this method, one sequentially computes a “minimizing” sequence of feedback control policies $\{u_n(\cdot)\}$ for the chain, and the control $u_{n+1}(\cdot)$ is obtained from an approximation to the solution to the equation for the cost under $u_n(\cdot)$. Suppose that a “tentative” optimal control for the controlled chain and given cost is $u_n(\cdot)$, and we wish to compute $u_{n+1}(\cdot)$. To do this we can work with either the original $p^h(x, y|\alpha)$ or the form (2.1), (2.2) (to eliminate the control dependence of the denominator). Suppose that we are given locally consistent transition probabilities $p^h(x, y|\alpha)$ and interpolation interval $\Delta t^h(x, \alpha)$ where $p^h(x, x|\alpha) \equiv 0$, and compute $\bar{p}^h(x, y|\alpha)$ and $\Delta \bar{t}^h(x, \alpha)$ from them as in (2.1) and (2.2). The cost function $W^h(x, u)$ satisfies (1.23) for the original transition probabilities and (2.6) for the “bar” system:

$$\begin{aligned} W^h(x, u_n) &= \sum_y \bar{p}^h(x, y|u_n(x)) W^h(y, u_n) + k(x, u_n(x)) \Delta \bar{t}^h(x, u_n(x)) \\ &= \bar{p}^h(x, x|u_n(x)) W^h(x, u_n) + \sum_{y \neq x} \bar{p}^h(x, y|u_n(x)) W^h(y, u_n) \\ &\quad + k(x, u_n(x)) \Delta \bar{t}^h(x, u_n(x)) \end{aligned} \tag{2.6}$$

for $x \in G_h^0$ with the boundary conditions $W^h(x, u_n) = g(x)$, for $x \in \{0, B\}$. In fact, equation (2.6) is equivalent to (1.23) for $u(\cdot) = u_n(\cdot)$, where the original $p^h(x, y|\alpha)$ and $\Delta t^h(x, \alpha)$ are used, in that the solutions are the same. Thus, we could use $p^h(x, y|\alpha)$ and $\Delta t^h(x, \alpha)$ in (2.6) and get the same solution. In order to calculate the next control, one gets an approximate value for $W^h(\cdot, u_n)$ and then calculates $u_{n+1}(\cdot)$ from either

$$u_{n+1}(x) = \arg \min_{\alpha \in \mathcal{U}} \left[\sum_y p^h(x, y|\alpha) W^h(y, u_n) + k(x, \alpha) \Delta t^h(x, \alpha) \right] \tag{2.7a}$$

or

$$u_{n+1}(x) = \arg \min_{\alpha \in \mathcal{U}} \left[\sum_y \bar{p}^h(x, y|\alpha) W^h(y, u_n) + k(x, \alpha) \Delta \bar{t}^h(x, \alpha) \right] \tag{2.7b}$$

depending on computational convenience. It is sometimes numerically simpler to use the $(\bar{p}^h(x, y|\alpha), \Delta \bar{t}^h(x, \alpha))$ because the denominators of these expressions do not depend on α . The sequence of minimizing controls might not be the same for the two cases. But owing to the local consistency of both approximations, the asymptotic results ($h \rightarrow 0$) will be the same.

One needs to be more careful if there is a discount factor, as noted above.

5.2.3 ALTERNATIVE MARKOV CHAIN APPROXIMATIONS FOR EXAMPLE 4 OF SECTION 5.1: SPLITTING THE OPERATOR

The general procedure outlined in Example 4 has many variations, some being more convenient for particular cases. Some of the issues involved in the choice of the chain can be seen via a concrete example. We will go through some “typical” calculations in order to illustrate some of the possibilities for exploiting the structure of particular problems. We will work with the particular one dimensional form

$$\begin{aligned} dx &= (\bar{b}(x) + u(x))dt + \sigma(x)dw, \quad |u(x)| \leq 1, \\ k(x, \alpha) &= \bar{k}(x) + k_0|\alpha|, \quad k_0 > 0, \end{aligned} \quad (2.8)$$

where the functions are appropriately bounded and smooth. Similar considerations work for the general r -dimensional problem. For (2.8), (1.22) yields

$$p^h(x, x \pm h|\alpha) = \frac{\sigma^2(x)/2 + h(\bar{b}(x) + \alpha)^{\pm}}{\sigma^2(x) + h|\bar{b}(x) + \alpha|},$$

which is an awkward expression if one wishes to use it in a minimization such as in (2.7).

We will show how to simplify it via a simple adaptation of the finite difference procedure of Example 4 of Section 5.1. First rewrite $\mathcal{L}^\alpha f(x)$ in the “split up” form, where the two terms in which the first derivative appears are separated:

$$\mathcal{L}^\alpha f(x) = f_x(x)\bar{b}(x) + f_x(x)\alpha + \frac{\sigma^2(x)}{2}f_{xx}(x). \quad (2.9)$$

Each of the two terms involving the first derivative will be approximated separately. We will use the difference approximation

$$f_x(x)\bar{b}(x) \longrightarrow \begin{cases} \bar{b}(x)\frac{f(x+h) - f(x)}{h} & \text{for } \bar{b}(x) \geq 0 \\ \bar{b}(x)\frac{f(x) - f(x-h)}{h} & \text{for } \bar{b}(x) < 0, \end{cases} \quad (2.10a)$$

$$f_x(x)\alpha \longrightarrow \begin{cases} \alpha\frac{f(x+h) - W(x)}{h} & \text{for } \alpha \geq 0 \\ \alpha\frac{f(x) - f(x-h)}{h} & \text{for } \alpha < 0. \end{cases} \quad (2.10b)$$

Following the procedure of Example 4, but using (2.10) for the first derivative terms, yields the expression (1.23) but with the following new definitions of the transition probabilities and interpolation interval: Define the

normalization $Q^h(x, \alpha) = \sigma^2(x) + h|\alpha| + h|\bar{b}(x)|$ and

$$\begin{aligned} p^h(x, x \pm h|\alpha) &= \frac{\sigma^2(x)/2 + h\bar{b}^\pm(x) + h\alpha^\pm}{Q^h(x, \alpha)}, \\ \Delta t^h(x, \alpha) &= \frac{h^2}{Q^h(x, \alpha)}. \end{aligned} \quad (2.11)$$

The chain defined by (2.11) is locally consistent with the controlled process defined by (2.8).

The transition probabilities and interpolation interval in (2.11) differ from (1.22) by having the absolute values of α and $\bar{b}(x)$ separated in the denominators and $(\alpha + \bar{b}(x))^\pm$ being replaced by $\alpha^\pm + \bar{b}(x)^\pm$ in the numerators, an advantage from the numerical point of view. The denominators still have an α -dependence. This can be eliminated via the method which led to (2.1), (2.2). To do this, define the maximum of $Q^h(x, \alpha)$ over $\alpha \in \mathcal{U} = [0, 1]$, namely, $\bar{Q}^h(x) = \sigma^2(x) + h + h|\bar{b}(x)|$. Then use

$$\begin{aligned} \bar{p}^h(x, x \pm h|\alpha) &= (\sigma^2(x)/2 + h\bar{b}^\pm(x) + h\alpha^\pm)/\bar{Q}^h(x), \\ \bar{p}^h(x, x|\alpha) &= h(1 - |\alpha|)/\bar{Q}^h(x), \\ \Delta \bar{t}^h(x) &= h^2/\bar{Q}^h(x). \end{aligned} \quad (2.12)$$

The chain defined by (2.12) is locally consistent with the controlled diffusion (2.8).

Solving for the Minimum in (2.7). The form of the transition probabilities given in (2.12) is quite convenient for getting the minimum in expressions such as (2.7b), and we will go through the calculation for our special case. The expression (2.7b) is equivalent to:

$$\begin{aligned} u_{n+1}(x) &= \arg \min_{|\alpha| \leq 1} [hW^h(x + h, u_n)\alpha^+ + hW^h(x - h, u_n)\alpha^- \\ &\quad + k_0|\alpha|h^2 - hW^h(x, u_n)|\alpha|]. \end{aligned}$$

It is easily seen that the minimizing α will take one of the values $\{-1, 0, 1\}$. Suppose that $W^h(x + h, u_n) \leq W^h(x - h, u_n)$. Then the minimizing α will be non-negative and the $W^h(x - h, u_n)$ can be dropped from the expression. Otherwise, it will be nonpositive and the $W^h(x + h, u_n)$ can be dropped. In fact, for the first case, if

$$hW^h(x + h, u_n) - hW^h(x, u_n) + k_0h^2 \leq 0,$$

then we can use $u_{n+1}(x) = 1$ for the minimizing control value. In general,

we can conclude that

$$\text{if } [W^h(x + h, u_n) - W^h(x, u_n)]/h \leq -k_0$$

and $W^h(x + h, u_n) \leq W^h(x - h, u_n)$, then $u_{n+1}(x) = 1$,

$$\text{if } [W^h(x, u_n) - W^h(x - h, u_n)]/h \geq k_0,$$

and $W^h(x + h, u_n) > W^h(x - h, u_n)$, then $u_{n+1}(x) = -1$,

otherwise, $u_{n+1}(x) = 0$.

Numerical minimization can be quite time consuming, and should be avoided if possible. Thus, it is important to keep in mind that the minimum in (2.7) should be solved for explicitly if at all possible. Also, in an actual program one writes the simplest decision tree for the choices above, and division by h is not required.

The Symmetric Difference Approximation for the First Derivative. If

$$\max_{x \in G_h} [\sigma^2(x) - h - h|\bar{b}(x)|] \geq 0,$$

then the symmetric finite difference approximation (1.18) can be used for the first derivatives in (2.9), and the calculations are simpler, because the denominators of the $p^h(x, y|\alpha)$ and $\Delta t^h(x, \alpha)$ will not then depend on α . If the above maximum is negative, then we still might be able to use the symmetric finite difference for one of the first derivative terms, as follows. Suppose that $\min_{x \in G_h} [\sigma^2(x) - h] \geq 0$. Then use the one sided finite difference for the representation of $W_x(x, u)\bar{b}(x)$ and the symmetric finite difference for the approximation of $W_x(x, u)\alpha$. With the use of this approximation, we still get a Markov chain which is locally consistent with (2.8), and the calculation of the infima in (2.7) is still relatively simple. The transition probabilities are given by

$$p^h(x, x \pm h|\alpha) = [\sigma^2(x)/2 + h\bar{b}(x)^\pm \pm h\alpha/2]/[\sigma^2(x) + h|\bar{b}(x)|].$$

The other $p^h(x, y|\alpha)$ are zero, and

$$\Delta t^h(x) = h^2/[\sigma^2(x) + h|\bar{b}(x)|].$$

Sometimes the value of h used in the numerical calculation is large. For example, in the multigrid method discussed in Chapter 6, some of the calculations are done on a very coarse grid. Also, it is sometimes convenient to get a rough solution to the dynamic programming equation with a large value of h , and then use that as a starting solution for a finer grid. In such cases, we need to be careful that the non-negativity conditions hold.

5.3 The General Finite Difference Method

The method of finite differences for getting the transition probabilities and interpolation time for a locally consistent chain for the general vector case uses the same procedure as the scalar case examples in Sections 5.1 and 5.2. We will work with the controlled process $x(\cdot)$ satisfying:

$$dx = b(x, u(x)) + \sigma(x)dw, \quad x \in \mathbb{R}^r. \quad (3.1)$$

Define the covariance matrix $a(x) = \sigma(x)\sigma(x)' = \{a_{ij}(x)\}, i, j = 1, \dots, r$, and recall the definition of the differential operator \mathcal{L}^α of (3.1)

$$\mathcal{L}^\alpha = \sum_{i=1}^r b_i(x, \alpha) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^r a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}. \quad (3.2)$$

Let e_i denote the unit vector in the i^{th} coordinate direction and let \mathbb{R}_h^r denote the uniform h -grid on \mathbb{R}^r ; i.e., $\mathbb{R}_h^r = \{x : x = h \sum_i e_i n_i : n_i = 0, \pm 1, \pm 2, \dots\}$. Until further notice, we use $S_h = \mathbb{R}_h^r$ as the state space of the approximating Markov chain. It is not necessary to use such a uniform grid but it simplifies the explanation.

Recall the procedure used for the scalar cases in Sections 5.1 and 5.2. Our aim in this and in the following sections is just to get an approximating Markov chain which is locally consistent with (3.1). The procedure of Section 5.1 started with a partial differential equation of the type (3.3):

$$\mathcal{L}^{u(x)} W(x, u) + k(x, u(x)) = 0, \quad (3.3)$$

where the function $k(\cdot)$ played only an auxiliary role. Then suitable finite difference approximations for the derivatives at the point x were substituted into (3.3), terms collected, and all terms divided by the coefficient of $W^h(x, u)$. We use $W^h(x, u)$ to denote the finite difference approximation to (3.3). This procedure gave the transition probabilities and interpolation interval automatically, as the coefficients in the resulting finite difference equation. We follow the same procedure and show that it works, together with all of the variants discussed in Sections 5.1 and 5.2.

The finite difference approximations used in this section are simple. E.g., only derivatives along the coordinate directions are taken, and the grid is uniform. Other approximations might be useful in particular situations and some further comments appear below. The method is a special case of the approach in the next section.

The Diagonal Case. For ease of explanation, we first write the expressions for the case where $a_{ij}(x) = 0$ for $i \neq j$. Proceed as in Section 5.1, Example 4. As in that example, the boundary conditions and boundary cost play no role in getting the chain, and we ignore them. For the second derivative,

we use the standard approximation

$$f_{x_i x_i}(x) \longrightarrow \frac{f(x + e_i h) + f(x - e_i h) - 2f(x)}{h^2}. \quad (3.4)$$

For the approximations to the first derivatives, we can adapt any of the previously discussed schemes. Let us start with the use of one sided approximations analogous to (1.8), namely,

$$f_{x_i}(x) \longrightarrow \begin{cases} [f(x + e_i h) - f(x)]/h & \text{if } b_i(x, \alpha) \geq 0 \\ [f(x) - f(x - e_i h)]/h & \text{if } b_i(x, \alpha) < 0. \end{cases} \quad (3.5)$$

Note that the direction (forward or backward) of the finite difference approximation (3.5) is the direction of the velocity component $b_i(x, \alpha)$. Let $x \in S_h$.

Use (3.4) and (3.5) in (3.3), denote the solution by $W^h(x, u)$, collect terms, clear the fraction by multiplying all terms by h^2 , then divide all terms by the coefficient of $W^h(x, u)$ to get the finite difference equation

$$W^h(x, u) = \sum_y p^h(x, y|u(x)) W^h(y, u) + k(x, u(x)) \Delta t^h(x, u(x)). \quad (3.6)$$

We define

$$\begin{aligned} p^h(x, x \pm e_i h | \alpha) &= \frac{a_{ii}(x)/2 + h b_i^\pm(x, \alpha)}{Q^h(x, \alpha)}, \\ \Delta t^h(x, \alpha) &= \frac{h^2}{Q^h(x, \alpha)}, \\ Q^h(x, \alpha) &= \sum_i [a_{ii}(x) + h |b_i(x, \alpha)|], \end{aligned} \quad (3.7)$$

and suppose that $\Delta t^h(x, \alpha) = O(h)$. For $y \in S_h$ not of the form $x \pm h e_i$ for some i , set $p^h(x, y | \alpha) = 0$.

The expressions (3.7) reduce to (1.22) when $x(t)$ is real valued. The constructed $p^h(x, y | \alpha)$ are non-negative. For each x and α , they sum (over y) to unity. Thus, they can be considered to be the transition probabilities for a controlled Markov chain. The local consistency of (3.7) with $x(\cdot)$ is easy to check. In particular,

$$E_{n,x}^{h,\alpha} \Delta \xi_n^h = b(x, \alpha) \Delta t^h(x, \alpha), \quad (3.8)$$

$$\text{cov}_{n,x}^{h,\alpha} \Delta \xi_n^h = a(x) \Delta t^h(x, \alpha) + O(h) \Delta t^h(x, \alpha) + [O(\Delta t^h(x, \alpha))]^2. \quad (3.9)$$

Thus, for the case where the $a(\cdot)$ matrix is diagonal, the finite difference method easily yields a locally consistent set of transition probabilities. All the variations discussed in Section 5.2 can be used, and will be illustrated in a two dimensional example in Subsection 5.3.2 below and at the end of the next section.

Remark on (3.5): Other Transition Directions. The approximation (3.5) uses the directions $\pm e_i$ and the transitions due to the drift terms are to the points of the type $x \pm e_i$. Other forms are possible, provided only that local consistency holds. For example, in two dimensions, a drift vector pointing in a direction between, say, e_2 and $e_1 + e_2$, need only use a randomization involving the points $x + he_2$ and $x + he_1 + he_2$. This general idea is also included in the approach of Section 5.4.

5.3.1 THE GENERAL CASE

Now consider the case where the off-diagonal terms $a_{ij}(x)$, $i \neq j$, are not all zero. For this case, the method illustrated so far is still useful, but it breaks down if some of the $a_{ij}(x)$ are large with respect to the diagonal terms $a_{ii}(x)$. This can be remedied by a more careful choice of the approximation, as discussed below or by the approach in the next section. But we prefer to start with the naive procedure. Analogous to what was done before for the finite difference approximation of the first derivative terms, we need to let the approximation used for each off-diagonal term $W_{x_i x_j}(x, u)$, $i \neq j$, depend on the sign of its coefficient $a_{ij}(x)$ to guarantee that the coefficients of the $W^h(y, u)$ in the finite difference equation will be positive and sum to unity so that they have the interpretation of a transition function for a Markov chain. The procedure which follows might seem somewhat complicated because the expressions for the finite difference approximations are complicated, but similar methods are common in the numerical analysis of elliptic equations, as are the particular approximations (3.10) and (3.11) below. For illustrative purposes, for the mixed derivatives, where $i \neq j$, we use the following standard finite difference approximations: For $a_{ij}(x) \geq 0$, we will use the difference approximations

$$\begin{aligned} f_{x_i x_j}(x) \rightarrow & [2f(x) + f(x + e_i h + e_j h) + f(x - e_i h - e_j h)]/2h^2 \\ & - [f(x + e_i h) + f(x - e_i h) + f(x + e_j h) \\ & + f(x - e_j h)]/2h^2. \end{aligned} \tag{3.10}$$

If $a_{ij}(x) < 0$, we will use

$$\begin{aligned} f_{x_i x_j}(x) \rightarrow & -[2f(x) + f(x + e_i h - e_j h) + f(x - e_i h + e_j h)]/2h^2 \\ & + [f(x + e_i h) + f(x - e_i h) + f(x + e_j h) \\ & + f(x - e_j h)]/2h^2. \end{aligned} \tag{3.11}$$

Each of (3.10) and (3.11) are consistent and standard approximations to the mixed second derivative. The reasons for these choices will soon be clear. Assume that

$$a_{ii}(x) - \sum_{j:j \neq i} |a_{ij}(x)| \geq 0, \tag{3.12}$$

for all i, x . The condition (3.12) depends on the coordinate system which is used, and several ways of relaxing it will be discussed later in this section

and in the next section. Define the normalizing coefficient

$$Q^h(x, \alpha) = \sum_i a_{ii}(x) - \sum_{i,j:i \neq j} |a_{ij}(x)|/2 + h \sum_i |b_i(x, \alpha)|. \quad (3.13)$$

Define the interpolation interval

$$\Delta t^h(x, \alpha) = \frac{h^2}{Q^h(x, \alpha)}. \quad (3.14)$$

Suppose that $\Delta t^h(x, \alpha) = O(h)$. Define the transition probabilities

$$\begin{aligned} p^h(x, x \pm e_i h | \alpha) &= [a_{ii}(x)/2 - \sum_{j:j \neq i} |a_{ij}(x)|/2 + h b_i^\pm(x, \alpha)] / Q^h(x, \alpha), \\ p^h(x, x + e_i h + e_j h | \alpha) &= p^h(x, x - e_i h - e_j h | \alpha) = a_{ij}^+(x) / 2Q^h(x, \alpha), \\ p^h(x, x - e_i h + e_j h | \alpha) &= p^h(x, x + e_i h - e_j h | \alpha) = a_{ij}^-(x) / 2Q^h(x, \alpha). \end{aligned} \quad (3.15)$$

For y not taking one of the listed values, define $p^h(x, y | \alpha) = 0$.

Equation (3.15) is obtained in the usual way: Substitute (3.4), (3.5), (3.10) and (3.11) into (3.3), use the symmetry of the $a(x)$ matrix, collect terms, and divide by the coefficient of $W^h(x, u)$ to get the finite difference equation (3.6). The transition probabilities and interpolation interval are the coefficients in that equation.

The chain defined by (3.15) with the interpolation interval (3.14) is locally consistent with the diffusion (3.1). The consistency is straightforward to check. For example, letting $\xi_n^{h,i}$ denote the i^{th} component of the vector ξ_n^h , we have for $i \neq j$,

$$E_{x,n}^{h,\alpha} \Delta \xi_n^{h,i} \Delta \xi_n^{h,j} = 2h^2 \frac{a_{ij}(x)}{2Q^h(x, \alpha)} = a_{ij}(x) \Delta t^h(x, \alpha).$$

Similarly to get the “consistency relation”

$$E_{x,n}^{h,\alpha} (\Delta \xi_n^{h,i})^2 = a_{ii}(x) \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)),$$

note that the probability that the i^{th} coordinate changes by $+h$ or $-h$ is

$$[a_{ii}(x) + O(h)|b_i(x, \alpha)|] / Q^h(x, \alpha).$$

The different choices for the finite difference approximation to the mixed second order derivatives are made in order to guarantee that the coefficients of the off-diagonal terms $p^h(x, x + e_i h \pm e_j h | \alpha)$, $p^h(x, x - e_i h \pm e_j h | \alpha)$, $i \neq j$, are non-negative. Also these choices guarantee that the coefficients sum to unity, so that they can be considered to be transition probabilities for an appropriate approximating Markov chain.

On Condition (3.12). Sometimes the condition (3.12) is not satisfied. We now mention some simple ways to circumvent this. Other more sophisticated ways will be discussed in the next section. The various alternative

methods indicate the versatility of the general Markov chain approximation scheme.

If the condition (3.12) fails because one of the $a_{ij}(x)$ is large relative to the diagonal term $a_{ii}(x)$, then we can redo the finite difference calculations with the finite difference grid rotated so that the coordinate lines are more closely aligned with the principal directions of the matrix $a(x)$; i.e., use a linear transformation of the coordinates. Even curvilinear coordinates can be used, if the principal directions vary substantially with x , but the required programming might be too difficult. The condition can be relaxed by letting the finite difference interval depend on the coordinate direction. This involves somewhat more work in programming the algorithm than for the constant interval case, but it is a widely used procedure. Similarly, one could consider using the “corner” points as the primary ones. There are many variations to this approach, and one need only keep track of the local consistency. We now illustrate it for a particular two dimensional case. For this case, a simple linear transformation of the coordinates (which diagonalizes the matrix $a(\cdot)$) might actually be easier to use in the sense that once it is done the subsequent programming will be simpler, provided that the transformation does not complicate dealing with whatever boundary conditions there might be.

Suppose that $a_{11}(x) = 1, a_{12}(x) = a_{21}(x) = 2, a_{22}(x) = 5$, so that (3.12) fails. Let h_i denote the finite difference interval in the i^{th} coordinate direction. Now use the finite difference procedures (3.4), (3.5), (3.10) and (3.11), but with $h_i e_i$ replacing the $h e_i$. Let $c = h_2/h_1$. Then one can show that the resulting coefficients are the transition probabilities and interpolation interval for a Markov chain which is locally consistent with $x(\cdot)$ provided that

$$a_{11} - a_{12}/c > 0, \quad a_{22} - ca_{12} > 0. \quad (3.16)$$

As shown in [80, Section 5], any value of c satisfying $2 < c < 2.5$ will yield a locally consistent chain. In the next section, it will be shown how to relax (3.12) further with the use of transitions to non-neighboring states.

Simplifying the $\Delta t^h(x, \alpha), p^h(x, y|\alpha)$. As noted in Section 5.2, it is sometimes useful to eliminate the dependence of the interpolation intervals on the control parameter α . We can use the procedures discussed in Section 5.2. Suppose that (3.12) holds and that

$$\min_x \left[\sum_i a_{ii}(x) - \sum_{i,j:i \neq j} |a_{ij}(x)|/2 \right] > 0. \quad (3.17)$$

Define a new interpolation interval (not depending on the control parameter)

$$\Delta t^h(x) = \frac{h^2}{[\sum_i a_{ii}(x) - \sum_{i,j:i \neq j} |a_{ij}(x)|/2]}. \quad (3.18)$$

Then the interpolation intervals of (3.14) and (3.18) differ by $O(h^3)$, and with the new interval, we still have local consistency with (3.1).

Recall that local consistency is needed to guarantee that the optimal value functions for the Markov chain approximations will converge to the optimal value function for (3.1) as $h \rightarrow 0$. Different locally consistent approximations might have rather different properties for “non-infinitesimal” values of h . Thus, one must exercise care with the use of any simplifications such as those discussed here.

To eliminate the control dependence from the denominator $Q^h(x, \alpha)$ of the transition functions, we can proceed as in the comments on Example 4 which were made in Section 5.2, and we briefly discuss the idea. Define $\bar{Q}^h(x) = \max_{\alpha \in \mathcal{U}} Q^h(x, \alpha)$. Define the *new* $p^h(x, y|\alpha)$, to be denoted by $\bar{p}^h(x, y|\alpha)$, as the old ones were defined in (3.15), but with $\bar{Q}^h(x)$ replacing the $Q^h(x, \alpha)$ there, and define the *new* interpolation interval by

$$\Delta \bar{t}^h(x) = h^2 / \bar{Q}^h(x).$$

The values of the transition probability which were just defined might sum (over y) to less than unity for some values of x and α . To compensate for this, we allow each state x to communicate with itself with the probability

$$\bar{p}^h(x, x|\alpha) = 1 - \sum_{y:y \neq x} \bar{p}^h(x, y|\alpha). \quad (3.19)$$

The new transition probabilities and interpolation interval are also locally consistent with (3.1).

5.3.2 A TWO DIMENSIONAL EXAMPLE: SPLITTING THE OPERATORS

For illustrative purposes, we repeat some of the calculations of Subsection 5.2.3 for the two dimensional example:

$$\begin{aligned} dx_1 &= \bar{b}_1(x)dt, \\ dx_2 &= \bar{b}_2(x)dt + u(x)dt + \sigma(x)dw. \end{aligned} \quad (3.20)$$

Writing the operator \mathcal{L}^α defined by (3.2) in the split up form suggested by (2.9), we can write (3.3) as

$$\begin{aligned} \mathcal{L}^{u(x)} W(x, u) &= W_{x_1}(x, u)\bar{b}_1(x) + W_{x_2}(x)\bar{b}_2(x) \\ &\quad + W_{x_2}(x, u)u(x) + \frac{W_{x_2 x_2}(x, u)\sigma^2(x)}{2}. \end{aligned}$$

Let the control space be $\mathcal{U} = [-1, 1]$. Now use (3.4) for the second derivative $W_{x_2 x_2}(x, u)$, and use the one sided approximation (3.5) for $W_{x_1}(x, u)$. Approximate each of the $W_{x_2}(x, u)$ terms separately. For each of these

$W_{x_2}(x, u)$ terms, use either the one sided or the two sided (symmetric) difference approximation, depending on the magnitudes of the coefficients. In particular, let $\sigma^2(x) \geq 1$ for all x , and use the one sided approximation for $W_{x_2}(x)\bar{b}_2(x)$. Then the two sided approximation can be used for $W_{x_2}(x, u)\alpha$ and we get for $h \leq 1$

$$\begin{aligned} p^h(x, x \pm e_1 h | \alpha) &= \frac{h\bar{b}_1^\pm(x)}{Q^h(x)}, \\ p^h(x, x \pm e_2 h | \alpha) &= \frac{\sigma^2(x)/2 + h\bar{b}_2^\pm(x) \pm h\alpha/2}{Q^h(x)}, \\ \Delta t^h(x) &= \frac{h^2}{Q^h(x)}, \\ Q^h(x) &= \sigma^2(x) + h|\bar{b}_1(x)| + h|\bar{b}_2(x)|. \end{aligned} \quad (3.21)$$

If a one sided difference approximation is used also to approximate the term $W_{x_2}(x, u)\alpha$, then we get

$$\begin{aligned} p^h(x, x + e_1 h | \alpha) &= \frac{h\bar{b}_1^\pm(x)}{Q^h(x, \alpha)}, \\ p^h(x, x \pm e_2 h | \alpha) &= \frac{\sigma^2(x)/2 + h\bar{b}_2^\pm(x) + h\alpha^\pm}{Q^h(x, \alpha)}, \\ \Delta t^h(x) &= \frac{h^2}{Q^h(x, \alpha)}. \\ Q^h(x, \alpha) &= \sigma^2(x) + h|\bar{b}_1(x)| + h|\bar{b}_2(x)| + h|\alpha|, \end{aligned} \quad (3.22)$$

5.4 A Direct Construction of the Approximating Markov Chain

In this section, it will be shown how to construct a locally consistent chain without having to go through the details of the finite difference approximation. The method illustrates the considerable freedom that we have in selecting the appropriate approximating chains. We restrict the description to the case of the model (3.1), but the changes required for other cases (e.g., the singular control problem dealt with in Chapter 8 or the jump diffusion case dealt with in Section 5.6 below) should be obvious. The finite difference method has the advantage of being essentially automatic and requires little effort. It does have some shortcomings. For example, (3.12) needs to hold, and this can cause a problem if the covariance matrix $a(x)$ is too “skewed.” But as pointed out in Section 5.3, (3.12) can be relaxed if we

use a different difference interval in each coordinate direction or rotate the coordinates, or by a combination of these two methods. Other methods, such as “non-local” difference approximations also help. The method to be discussed in this section includes all of the above possibilities, and takes advantage of the intuition which we gained in our discussion of the finite difference method. Three examples will be given and then the general procedure will be described. We will try to utilize the information which we have concerning the forms of the transition probabilities which is provided by (1.22), (2.1), (2.15), (3.7), and (3.15).

Essentially, the method decomposes the increments of $\Delta\xi_n^h$ into components, each of which is due to a different part of the dynamical effects, then it gets a locally consistent transition probability for each component, and combines them appropriately. We first give three examples and then the general method. The method is quite powerful. It provides a good and often simple way to view the construction of the approximating Markov chain. Indeed, one should not rely too heavily on direct finite difference type methods, unless clearly convenient, without understanding their relationship to the approach of this section.

5.4.1 AN INTRODUCTORY EXAMPLE

Example 1. We start with an examination of the transition probabilities in (3.7). Write the transition probabilities in (3.7) as

$$\begin{aligned} P_{x,n}^{h,\alpha}\{\Delta\xi_n^h = \pm e_i h\} &= \frac{a_{ii}(x)}{2} \times \text{normalization} \\ &\quad + h b_i^\pm(x, \alpha) \times \text{normalization}, \end{aligned} \tag{4.1}$$

where the normalization guarantees that the sum of the transition probabilities is unity for each x, α . In fact, the normalization is just the familiar expression $1/Q^h(x, \alpha)$ given by (3.7). The “separated” form of the right hand side in (4.1) implies that we can split the transition into disjoint possibilities, a transition due to “noise” [which yields the term $a_{ii}(x)$] and one due to “drift” [which yields the drift term $b_i^\pm(x, \alpha)$]. This separation is the key to the direct method. One gets a locally consistent transition probability for each part of the system separately; i.e., for the two systems $dx = \sigma(x)dw$ and $dx = b(x, \alpha)dt$. Then the separate forms are combined with the appropriate weights and normalization factor.

A Decomposition of the Transition Probability. Suppose that we have two sets of transition probabilities and interpolation intervals

$$(p_a^h(x, y|\alpha), \Delta t_a^h(x, \alpha)), \quad (p_b^h(x, y|\alpha), \Delta t_b^h(x, \alpha)),$$

which are locally consistent with the systems defined by

$$dx = \sigma(x)dw, \tag{4.2a}$$

$$dx = b(x, u)dt, \quad (4.2b)$$

respectively. By the definition of local consistency, we have

$$\begin{aligned} \sum_y (y - x)(y - x)' p_a^h(x, y | \alpha) &= a(x) \Delta t_a^h(x, \alpha) + o(\Delta t_a^h(x, \alpha)), \\ \sum_y (y - x) p_a^h(x, y | \alpha) &= o(\Delta t_a^h(x, \alpha)) \end{aligned} \quad (4.3)$$

and

$$\begin{aligned} \sum_y (y - x)(y - x)' p_b^h(x, y | \alpha) &= o(\Delta t_b^h(x, \alpha)), \\ \sum_y (y - x) p_b^h(x, y | \alpha) &= b(x, \alpha) \Delta t_b^h(x, \alpha) + o(\Delta t_b^h(x, \alpha)). \end{aligned} \quad (4.4)$$

We next show how to combine (4.3) and (4.4) to get the desired transition probabilities and interpolation intervals $p^h(x, y | \alpha), \Delta t^h(x, \alpha)$ in (3.7). This is done via a “coin toss.” Choose system (4.2a) [i.e., the transition probabilities $p_a^h(x, y | \alpha)$ giving (4.3)] with probability $P^h(a|x, \alpha)$, and the system (4.2b) [i.e., the transition probabilities giving (4.4)] with probability $P^h(b|x, \alpha) = 1 - P^h(a|x, \alpha)$, where $P^h(a|x, \alpha)$ is to be determined. The transition probabilities determined by the coin toss will be the $p^h(x, y | \alpha)$. Dropping the $o(\cdot)$ terms, local consistency with (3.1) requires that

$$\begin{aligned} \sum_y (y - x)(y - x)' p_a^h(x, y | \alpha) P^h(a|x, \alpha) &= a(x) \Delta t_a^h(x, \alpha) P^h(a|x, \alpha) \\ &= a(x) \Delta t^h(x, \alpha), \\ \sum_y (y - x) p_b^h(x, y | \alpha) P^h(b|x, \alpha) &= b(x, \alpha) \Delta t_b^h(x, \alpha) P^h(b|x, \alpha) \\ &= b(x, \alpha) \Delta t^h(x, \alpha). \end{aligned} \quad (4.5)$$

Equations (4.5) imply that

$$P^h(a|x, \alpha) = \frac{\Delta t_b^h(x, \alpha)}{\Delta t_b^h(x, \alpha) + \Delta t_a^h(x, \alpha)}.$$

This implies that

$$\begin{aligned} p^h(x, y | \alpha) &= [p_a^h(x, y | \alpha) \Delta t_b^h(x, \alpha) + p_b^h(x, y | \alpha) \Delta t_a^h(x, \alpha)] \\ &\quad \times \text{normalization}, \end{aligned} \quad (4.6)$$

$$\Delta t^h(x, \alpha) = \frac{\Delta t_a^h(x, \alpha) \Delta t_b^h(x, \alpha)}{\Delta t_a^h(x, \alpha) + \Delta t_b^h(x, \alpha)}. \quad (4.7)$$

Because the normalization equals $1/(\Delta t_a^h(x, \alpha) + \Delta t_b^h(x, \alpha))$, we have

$$p^h(x, y|\alpha) = \frac{p_a^h(x, y|\alpha)\Delta t_b^h(x, \alpha) + p_b^h(x, y|\alpha)\Delta t_a^h(x, \alpha)}{\Delta t_a^h(x, \alpha) + \Delta t_b^h(x, \alpha)}.$$

In (3.22), the transition probabilities were written as certain ratios. Such forms are common. Suppose that we can write

$$p_a^h(x, y|\alpha) = \frac{n_a^h(x, y|\alpha)}{Q_a^h(x, \alpha)}, \quad p_b^h(x, y|\alpha) = \frac{n_b^h(x, y|\alpha)}{Q_b^h(x, \alpha)}.$$

By suitably scaling $n_a^h(x, y|\alpha)$ we can suppose that

$$\sum_y (y - x)(y - x)' n_a^h(x, y|\alpha) = h^2 a(x)$$

which yields that

$$\Delta t_a^h(x, \alpha) = \frac{h^2}{Q_a^h(x, \alpha)}.$$

An analogous calculation yields

$$\Delta t_b^h(x, \alpha) = \frac{h}{Q_b^h(x, \alpha)}.$$

Define $Q^h(x, \alpha) = Q_a^h(x, \alpha) + hQ_b^h(x, \alpha)$. Then (4.6) and (4.7) can be written as

$$p^h(x, y|\alpha) = \frac{n_a^h(x, y|\alpha) + hn_b^h(x, y|\alpha)}{Q^h(x, \alpha)} \quad (4.8)$$

and

$$\Delta t^h(x, \alpha) = \frac{h^2}{Q^h(x, \alpha)}. \quad (4.9)$$

To gain confidence in the formulas (4.6) to (4.9), let us see how (3.7) can be recovered. Define

$$Q_a^h(x, \alpha) = \sum_i a_{ii}(x),$$

$$Q_b^h(x, \alpha) = \sum_i |b_i(x, \alpha)|.$$

We can use the locally consistent [with (4.2a) and (4.2b), respectively] transition probabilities and interpolation intervals

$$\begin{aligned} p_a^h(x, x \pm he_i|\alpha) &= \frac{a_{ii}(x)}{2Q_a^h(x, \alpha)}, & \Delta t_a^h(x, \alpha) &= \frac{h^2}{Q_a^h(x, \alpha)}, \\ p_b^h(x, x \pm he_i|\alpha) &= \frac{b_i^\pm(x, \alpha)}{Q_b^h(x, \alpha)}, & \Delta t_b^h(x, \alpha) &= \frac{h}{Q_b^h(x, \alpha)}. \end{aligned} \quad (4.10)$$

Using (4.10) in (4.6) and (4.7) yields

$$\begin{aligned} p^h(x, x \pm he_i | \alpha) &= \left[\frac{a_{ii}(x)/2}{Q_a^h(x, \alpha)} \frac{h}{Q_b^h(x, \alpha)} + \frac{b_i^\pm(x, \alpha)}{Q_b^h(x, \alpha)} \frac{h^2}{Q_a^h(x, \alpha)} \right] \\ &\quad \times \text{normalization} \\ &= \left[\frac{a_{ii}(x)}{2} + b_i^\pm(x, \alpha)h \right] \times \text{normalization} \end{aligned}$$

and

$$\Delta t^h(x, \alpha) = \frac{h^2}{Q_a^h(x, \alpha) + hQ_b^h(x, \alpha)},$$

which are precisely the forms given by (3.7) or, equivalently, by (4.8) and (4.9). The general case follows the same lines.

It is generally the case that the interpolation intervals have the representations in (4.10), for appropriate $Q_i^h(x, \alpha)$. These representations imply the form of the interpolation interval (4.9), where the drift component is weighed by h in calculating the normalization, and the form of the transition probabilities in (4.8).

Two special cases which illustrate variations of this procedure will be described in the next two subsections.

5.4.2 EXAMPLE 2. A DEGENERATE COVARIANCE MATRIX

Consider the two dimensional case, where the system is defined by (3.1), but the covariance $a(\cdot)$ is totally degenerate and takes the value $a(x) = a = q^2vv'$, where v is a column vector and q a real number. That is, the system satisfies

$$dx = b(x, u)dt + qvdw, \quad v = (\beta_1, \beta_2), \quad \beta_i > 0. \quad (4.11)$$

Thus, the effect of the noise is to “push” the system in the directions $\pm v$. See the illustration of the terms in Figure 5.1a, where we let $\beta_1 = 1$.

The degenerate structure of the noise covariance matrix suggests that the part of the transitions of any approximating Markov chain which approximates the effects of the “noise” would move the chain in the directions $\pm v$. We next pursue the decomposition approach of Example 1. Let the state space S_h be that indicated by the extension of the grid in Figure 5.1b. The “boxes” are such that the diagonals are in the directions $\pm v$ and of magnitude vh . Given $h > 0$, the grid has spacing h in the e_1 direction and $\beta_2 h$ in the e_2 direction.

One set of transition probabilities for a locally consistent chain for the component of (4.11) which is represented by

$$dx = qvdw \quad (4.12)$$

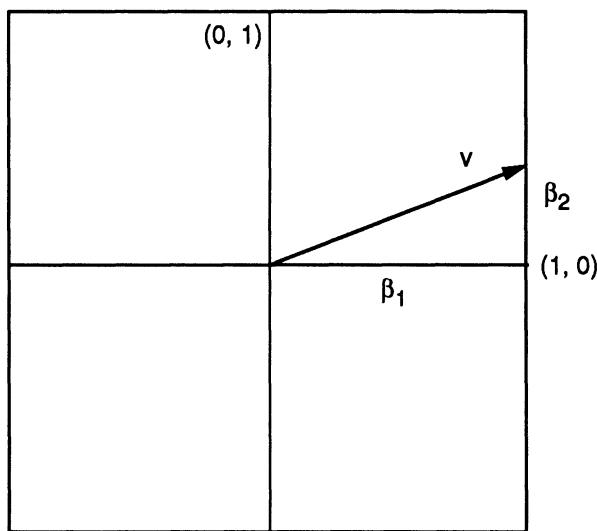


Figure 5.1a. The noise direction. Example 2.

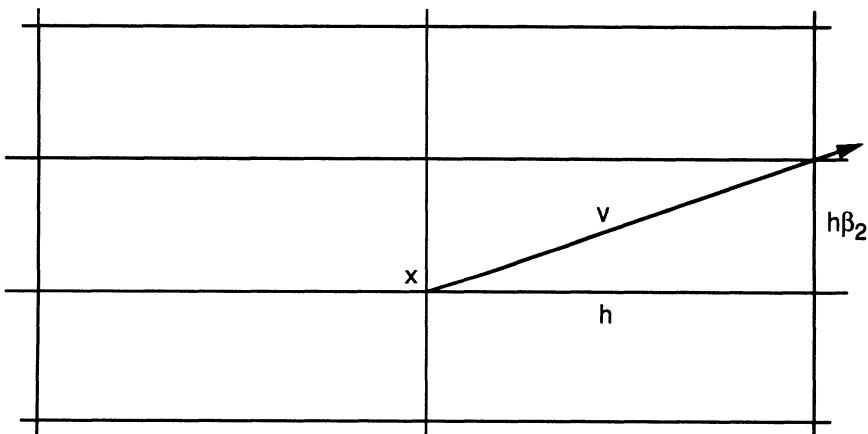


Figure 5.1b. The state space. Example 2.

is $p_a^h(x, x \pm hv|\alpha) = 1/2$. With these transition probabilities, the covariance of the state transition can be written as

$$\sum_y (y - x)(y - x)' p_a^h(x, y|\alpha) = vv' h^2 = q^2 vv' \frac{h^2}{q^2} = a(x) \frac{h^2}{q^2}.$$

Thus, $p_a^h(x, y|\alpha)$ together with the interpolation interval $\Delta t_a^h(x, \alpha) = h^2/q^2$, is locally consistent with (4.12). One possibility for the transition probability for the approximation to (4.2b) is

$$p_b^h(x, x \pm \beta_i e_i h|\alpha) = \frac{b_i^\pm(x, \alpha)}{\beta_i} \times \text{normalization}, \quad (4.13)$$

where the normalization is $1/ [|b_1(x, \alpha)/\beta_1| + |b_2(x, \alpha)/\beta_2|] = 1/Q_b^h(x, \alpha)$. The β_i in the denominator in (4.13) is needed to get local consistency. The local consistency of (4.13) with (4.2b) is shown by the calculation

$$\begin{aligned} \sum_y (y - x)p_b^h(x, y|\alpha) &= h \left[\frac{b_1(x, \alpha)}{b_2(x, \alpha)} \right] \times \text{normalization} \\ &= b(x, \alpha) \frac{h}{|b_1(x, \alpha)/\beta_1| + |b_2(x, \alpha)/\beta_2|} \\ &= b(x, \alpha) \Delta t_b^h(x, \alpha), \end{aligned}$$

$$\sum_y (y - x)(y - x)' p_b^h(x, y|\alpha) = o(\Delta t_b^h(x, \alpha)).$$

Now, we combine the above “partial” transition probabilities. Following the guide of Example 1, note that the interpolation intervals have the form of (4.10) with $Q_a^h(x, \alpha) = q^2$. Thus, as in (4.8) and (4.9), we weigh the contribution of the diffusion and drift component to get

$$p^h(x, x \pm hv|\alpha) = \frac{q^2}{2} \frac{1}{Q^h(x, \alpha)},$$

$$p^h(x, x \pm h\beta_i e_i|\alpha) = \frac{hb_i^\pm(x, \alpha)}{\beta_i} \frac{1}{Q^h(x, \alpha)},$$

where

$$Q^h(x, \alpha) = q^2 + \frac{h|b_1(x, \alpha)|}{\beta_1} + \frac{h|b_2(x, \alpha)|}{\beta_2},$$

$$\Delta t^h(x, \alpha) = \frac{h^2}{Q^h(x, \alpha)}.$$

Local consistency with the process (4.1) of this example follows by an argument of the type used in the discussion in Example 1 and can also be readily checked directly.

Extensions. It is not necessary that we use transitions to $x \pm h\beta_i e_i$ to get the approximations for the drift term. Depending on the mean directions, one could use the “corner” points in lieu of some of the $x \pm h\beta_i e_i$. The transition probabilities just derived are analogous to what one would get with the finite difference method if a one sided difference approximation were used for the first derivative terms. This is evident because we are approximating the noise and drift terms separately here, so each needs to yield a “partial” transition probability. However, once the transition probabilities $p^h(x, y|\alpha)$ are determined, we might be able to modify them to get a result analogous to that for the finite difference method with the symmetric difference used for the first derivatives. Simply replace the $b_i^\pm(x, \alpha)/\beta_i$ by

$$\pm b_i(x, \alpha)/2\beta_i.$$

We only need to check whether the resulting expressions for the transition probabilities are all non-negative. This same procedure could also be used for the finite difference method and would transform the expressions obtained with one sided difference approximation into the expressions corresponding to the use of the symmetric difference.

5.4.3 EXAMPLE 3

We next deal with a special two dimensional case with a non-degenerate covariance. Refer to Figure 5.2a for an illustration of the terms and the specification of the values of β_i . Let there be vectors (as in the figure) v_i of the form $v_1 = (\beta_{11}, \beta_{12})$ and $v_2 = (-\beta_{21}, \beta_{22})$, where we set $\beta_{11} = \beta_{22} = 1$. Assume that there are positive real numbers q_i such that $a(x) = a = q_1^2 v_1 v_1' + q_2^2 v_2 v_2'$. Thus, we can write the process as

$$dx = b(x, u)dt + q_1 v_1 dw_1 + q_2 v_2 dw_2, \quad (4.14)$$

where the $w_i(\cdot)$ are mutually independent one dimensional Wiener processes.

We follow the development of the previous two examples by first decomposing the transition probability into component parts, some of which are due to the noise effects and some to the drift, and then combining the parts. Refer to Figure 5.2b for a description of the state space and transitions. For this example, the part of the approximating chain which concerns the noise term will take a state $x \in S_h$ into other states in the directions $\pm v_1$ or $\pm v_2$. In particular, the transitions concerned with the component of the covariance which is represented by $q_i^2 v_i v_i'$ should move the chain in the directions $\pm v_i$. The development of the transition probabilities for the subsystems in Examples 1 and 2 can be nested. In the present case, we can deal with the subsystems $dx = q_1 v_1 dw_1$ and $dx = q_2 v_2 dw_2$ separately, and then combine the results. Using the method of Example 2 yields the following formulas:

$$p^h(x, x \pm v_1 h | \alpha) = (q_1^2/2) \times \text{normalization constant},$$

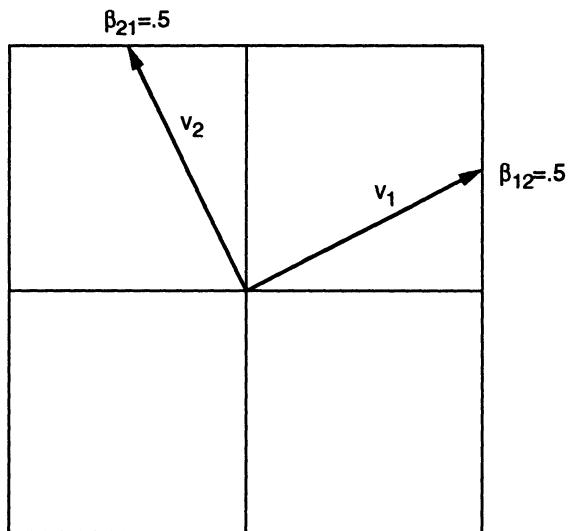


Figure 5.2a. The principle directions of the noise.

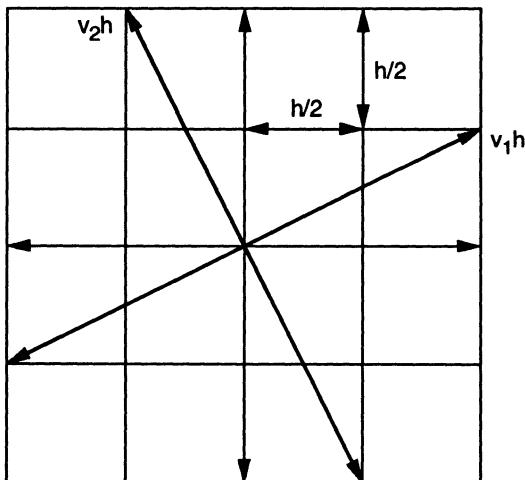


Figure 5.2b. Transition directions. Example 3.

$$\begin{aligned} p^h(x, x \pm v_2 h | \alpha) &= (q_2^2 / 2) \times \text{normalization constant}, \\ p^h(x, x \pm e_i h | \alpha) &= h b_i^\pm(x, \alpha) \times \text{normalization constant}. \end{aligned}$$

In order that the transition probabilities which were just defined add to unity, the normalization constant needs to be

$$Q^h(x, \alpha) = q_1^2 + q_2^2 + h|b_1(x, \alpha)| + h|b_2(x, \alpha)|.$$

If the interpolation time interval is defined by $h^2/Q^h(x, \alpha)$, then the chain is locally consistent with the systems model of this example. Because the transitions are not only to the nearest neighbors, more memory is required.

5.4.4 THE GENERAL METHOD

The decomposition scheme used in the above examples can be described for a general problem as follows. For each point x , let there be a set of vectors $M(x) = \{v_i(x), i \leq m(x)\}$, where $m(x)$ is uniformly bounded in x . For $x \in S_h$, define the set $Y^h(x) = \{y : y = x + hv_i(x), i \leq m(x)\}$, and suppose that $Y^h(x) \subset S_h$. Local consistency with (3.1) implies that

$$\begin{aligned} \sum_{i \in M(x)} p^h(x, x + hv_i(x) | \alpha) hv_i(x) &= b(x, \alpha) \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)), \\ \sum_{i \in M(x)} p^h(x, x + hv_i(x) | \alpha) hv_i(x) (hv'_i(x)) \\ &= a(x) \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)), \end{aligned} \tag{4.15}$$

where $p^h(x, y | \alpha)$ are the transition probabilities and $\Delta t^h(x, \alpha) \rightarrow 0$, as $h \rightarrow 0$.

A convenient way of getting the transition probabilities follows the approach of the above examples by treating the noise and drift components of the transition probability separately and then adding them up with the appropriate weight and normalization factor. Suppose that $q_i^1(x)$ and $q_i^0(x, \alpha)$ are non-negative (see the remarks on extensions below) and satisfy

$$b(x, \alpha) = \sum_{i \in M(x)} q_i^0(x, \alpha) v_i(x), \tag{4.16a}$$

$$a(x) = \sum_{i \in M(x)} q_i^1(x) v_i(x) v'_i(x), \tag{4.16b}$$

$$\sum_{i \in M(x)} q_i^1(x) v_i(x) = 0. \tag{4.17}$$

Define

$$Q^h(x, \alpha) = \sum_{i \in M(x)} [h q_i^0(x, \alpha) + q_i^1(x)]. \tag{4.18}$$

If the covariance $a(x)$ depends on the control, then so would the $q_i^1(\cdot)$.

Define the interpolation interval

$$\Delta t^h(x, \alpha) = \frac{h^2}{Q^h(x, \alpha)} \quad (4.19)$$

and suppose that it goes to zero as $h \rightarrow 0$. Then (4.19) and the transition probabilities defined by

$$p^h(x, x + hv_i(x)|\alpha) = \frac{hq_i^0(x, \alpha) + q_i^1(x)}{Q^h(x, \alpha)} \quad (4.20)$$

are locally consistent with (3.1).

This method contains Examples 2 and 3 of this section. It reduces to that of Section 5.3, when the $\{v_i(x)\}$ is the set of vectors $\{\pm e_i, e_i \pm e_j, -e_i \pm e_j, i \neq j, i, j\}$, and the values $q_j^0(x, \alpha) = b_i^\pm(x, \alpha)$ are used with the vectors $v_j(x)$ of the form $\pm e_i$ and the $q_j^1(x)$ are determined in an obvious way from (3.15).

Comment. (4.17) is used to guarantee that the $q_i^1(x)$ do not contribute to the mean direction. If $q_i^0(x, \alpha)$ are set equal to zero, then (4.17) and (4.16b) imply that (4.19) and (4.20) are locally consistent with (4.2a). The comments made at the end of Example 2 concerning modifications of the contributions of the “drift” components hold here also. We note also that the $q_i^0(x, \alpha)$ need not actually all be positive. It is required only that

$$q_i^1(x) + hq_i^0(x, \alpha) \geq 0, \quad Q^h(x, \alpha) > 0. \quad (4.21)$$

Another Example: Correlated Noise. The following example illustrates a useful decomposition. The system is two dimensional and

$$dx = b(x)dt + dw,$$

where $w(\cdot)$ can be written in terms of three mutually independent Wiener processes as

$$w(t) = \begin{pmatrix} w_1(t) \\ w_2(t) \end{pmatrix} + vw_3(t).$$

With this decomposition, and the grid spacing depending on the direction, we can easily get a locally consistent approximation.

5.5 Variable Grids

Up to this point, the major objective was to get a locally consistent approximating Markov chain. It is not always possible to get a chain which

is locally consistent at all points of S_h for all h . Loosely speaking, the criterion for local consistency can be relaxed on a set of points if the limit process spends negligible time on arbitrarily small neighborhoods of that set and the possible values that the coefficients might take on that set do not affect the limit process. This point will be discussed further in Chapter 10. The probabilistic interpretation comes to our aid here also and allows us to treat a relaxation of the consistency and continuity conditions which might be difficult to achieve via the analytic methods of numerical analysis. The problem occurs in practice in situations such as the following. Consider a problem where the set G of interest for the control problem is divided into disjoint subsets $\{G^1, \dots\}$, and we wish to have a different “level” of approximation on each G^i . For example, a different spacing might be used for the approximating grid on each subset. Suppose that the state space or approximating grid is S_h and that we can find a suitable locally consistent chain on the parts of S_h which are interior to each G^i . Due to the discontinuities in the grid sizes, it might not be possible to get local consistency at all points on the boundaries between the G^i .

Let us consider a particular two dimensional example, illustrated in Figure 5.3a. We continue to use the model (3.1). The set G is the entire two dimensional space, and G^1 is the “southeast” subset enclosed by the dashed line boundary. The boundary is divided into two disjoint segments, denoted by A_0 and A_1 . Suppose that, for reasons concerning the desired numerical accuracy, we require a finer grid on G^1 than on the rest of G . On G^1 , we set the spacing interval to be $h/2$; on the rest of G we let it be h . We need not use a square grid, but it simplifies the discussion. The set S_h is the collection of all points on the grids. Suppose that we can get a locally consistent approximating chain at all points not on the boundary $\partial G^1 = A_0 \cup A_1$. Owing to the regularity of the grids interior to both G^1 and $G - G^1$, this could be done by any of the methods of the previous sections, under appropriate conditions on the covariance matrix. The same methods can be used to get a locally consistent transition probability on the boundary points which are in the coarse grid. But there might be problems at the other points on the boundary because these do not connect to their neighbors in the same “symmetric” way that the other points do. The main difficulties as well as their resolution can be seen via a particular example.

A Two Dimensional Example. Let us consider the degenerate two dimensional case given by (5.1), where the $b_i(x, \alpha)$ are bounded and continuous:

$$\begin{aligned} dx_1 &= b_1(x, \alpha)dt, \\ dx_2 &= b_2(x, \alpha) + \sigma dw, \quad \sigma > 0. \end{aligned} \tag{5.1}$$

The two boundary segments will be examined separately. We will first show that a locally consistent transition probability can be constructed at the point y_0 on A_1 , and hence for all the points on A_1 . We will use the direct method of Section 5.4, and deal with the drift and noise terms separately,

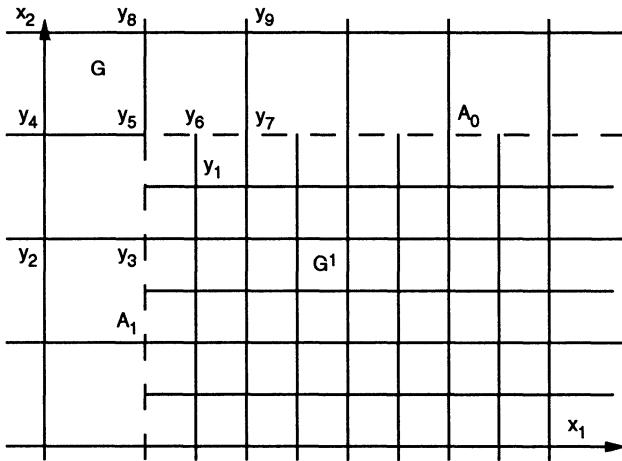


Figure 5.3a. A variable grid.

and then combine them with the appropriate weights and normalization. In particular, the formulas (4.16) and (4.20) will be used.

The Calculation at y_0 . Define the vectors $v_i(y_0), i = 1, \dots, 5$, by

$$hv_i(y_0) = y_i - y_0.$$

These are represented in Figure 5.3b. In particular,

$$v_1(y_0) = e_1/2,$$

$$v_5(y_0) = -v_3(y_0) = e_2/2,$$

$$v_4(y_0) = e_2/2 - e_1, \quad v_2(y_0) = -e_2/2 - e_1.$$

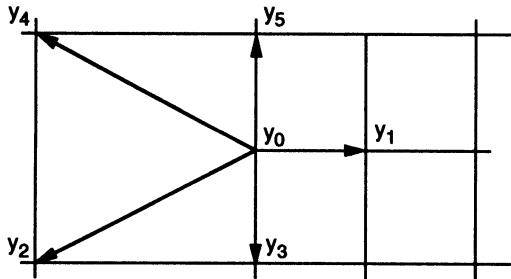
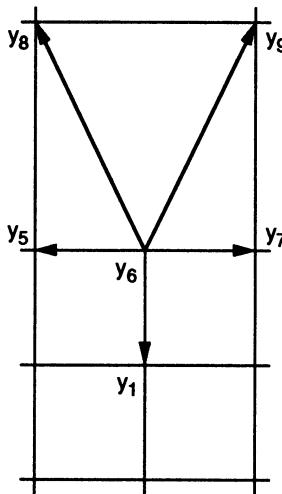
It is easy to find $q_j^0(x, \alpha)$ which yield (4.16a): Use

$$\begin{aligned} q_5^0(y_0, \alpha) &= 2b_2^+(y_0, \alpha), \\ q_3^0(y_0, \alpha) &= 2b_2^-(y_0, \alpha), \\ q_1^0(y_0, \alpha) &= 2b_1^+(y_0, \alpha), \\ q_2^0(y_0, \alpha) &= b_1^-(y_0, \alpha)/2, \\ q_4^0(y_0, \alpha) &= b_1^-(y_0, \alpha)/2. \end{aligned}$$

We can also write

$$a(y_0) = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2 \end{pmatrix} = 2\sigma^2[v_5(y_0)v'_5(y_0) + v_3(y_0)v'_3(y_0)]$$

Thus, we can use $q_5^1(y_0) = q_3^1(y_0) = 2\sigma^2$, and then (4.19) and (4.20) are locally consistent with (5.1) at y_0 .

Figure 5.3b. Transitions from y_0 .Figure 5.3c. Transitions from y_6 .

The Calculation at y_6 . We will show that it is not possible to get a locally consistent transition probability at y_6 . Nevertheless, one can get close enough so that the limit processes and value function are the desired ones. Following the procedure of Section 5.4, as done above, define the vectors $v_i(y_6)$, $i = 1, 5, 7, 8, 9$, by

$$hv_i(y_6) = y_i - y_6.$$

The possible transitions are shown in Figure 5.3c. It is possible to get a representation of the drift $b(y_6, \alpha)$ in terms of these $v_i(y_6)$ such that (4.16a) holds. The problem arises with the representation of the covariance matrix

$$a = \sigma^2 e_2 e_2' \quad (5.2)$$

of the type needed for (4.16b). The difficulty is due to the fact that none of the $v_i(y_6)$ equal $-v_1(y_6)$. The results of Theorem 10.5.3 imply that the

desired convergence will occur if we get a transition probability which is locally consistent everywhere except on A_0 , where it satisfies

$$E_{x,n}^{h,\alpha} \Delta \xi_n^h = O(h),$$

$$c_2 h^2 I \geq \text{cov}_{x,n}^{h,\alpha} \Delta \xi_n^h \geq c_1 h^2 I, \quad c_i > 0. \quad (5.3)$$

Consider an “approximation” to (5.2) of the form

$$\bar{a} = q_1^1(y_6)v_1(y_6)v'_1(y_6) + q_8^1(y_6)[v_8(y_6)v'_8(y_6) + v_9(y_6)v'_9(y_6)], \quad (5.4)$$

where $q_1^1(y_6)$ and $q_8^1(y_6)$ are positive. In order not to introduce a very large value of $E_{y_6,n}^{h,\alpha} \Delta \xi_n^h$, we require that the mean change implied by (5.4) is zero. Thus, we require that

$$q_1^1(y_6)v_1(y_6) + q_8^1(y_6)[v_8(y_6) + v_9(y_6)] = 0. \quad (5.5)$$

This is a critical condition and implies that $q_1^1(y_6) = 4q_8^1(y_6)$. Choose

$$q_1^1(y_6, \alpha) = 4\sigma^2/3, \quad q_8^1(y_6, \alpha) = q_9^1(y_6, \alpha) = \sigma^2/3.$$

Then

$$\bar{a} = \frac{\sigma^2}{3} [4v_1(y_6)v'_1(y_6) + v_8(y_6)v'_8(y_6) + v_9(y_6)v'_9(y_6)], \quad (5.6)$$

or

$$\bar{a} = \sigma^2 \begin{bmatrix} 1/6 & 0 \\ 0 & 1 \end{bmatrix}.$$

Summary. The transition probabilities and interpolation interval just constructed are locally consistent with (5.1) at all points except at the points on A_0 which are in the fine grid but not on the coarse grid. At these points there is local consistency, except for the added “noise term” in the horizontal direction. Asymptotically, the behavior on A_0 does not affect the form of the limits for two reasons: First, the fraction of time that the limit process will spend in a small neighborhood of A_0 vanishes as the size of that neighborhood goes to zero. Second, the behavior of the approximating chain on the boundary does not affect the limit processes (as $h \rightarrow 0$), due to the fact that $\sigma \neq 0$ and

$$\text{var}_{x,n}^{h,\alpha} (\Delta \xi_n^h)_2 \geq k_1 h^2, \quad (5.7)$$

where $k_1 > 0$ and $(\Delta \xi_n^h)_2$ is the component in the vertical direction. In fact, the properties in the last sentence are the most important, provided that the “drift” $|E_{x,n}^{h,\alpha} \Delta \xi_n^h|/\Delta t_n^h$ is bounded on the boundary. Otherwise, the form of the approximation used on A_0 is not important.

5.6 Jump Diffusion Processes

5.6.1 THE JUMP DIFFUSION PROCESS MODEL: RECAPITULATION

The previous sections have been concerned with the problem of getting a locally consistent Markov chain approximation for the diffusion model (3.1). If such a locally consistent chain is available, then the extension to a locally consistent chain for the jump diffusion of Section 1.5 is straightforward. First, we review some properties of the jump diffusion process. Let us recall the jump diffusion model (1.5.6):

$$dx = b(x, u)dt + \sigma(x)dw + dJ. \quad (6.1)$$

The jump term was represented in Section 1.5 in terms of a function $q(\cdot)$ and a Poisson measure $N(\cdot)$ of intensity $\lambda dt \times \Pi(d\rho)$, where $\Pi(\cdot)$ has compact support Γ . The function $q(\cdot)$ is assumed to be bounded and measurable and is continuous in x for each value of ρ . In terms of the Poisson measure, the jump term has the representation

$$J(t) = \int_0^t \int_{\Gamma} q(x(s^-), \rho) N(ds d\rho).$$

The use of the Poisson measure is just a convenient way of keeping track of the jump times and values. We will next recapitulate some of the basic facts and implications of Section 1.5 which will be useful for constructing the approximating Markov chain and for the convergence theorems.

There is an intuitively simple, but equivalent, way to define the process (6.1) and the jump term, by working with the jump times and values directly. To do this, let ν_n , $n \geq 1$, denote the time of the n^{th} jump, and set $\nu_0 = 0$. Let $\{\nu_{n+1} - \nu_n, \rho_n, n < \infty\}$ be mutually independent random variables with $\nu_{n+1} - \nu_n$ being exponentially distributed with mean value $1/\lambda$, and let the ρ_n have the common distribution $\Pi(\cdot)$. In addition, for each n , let $\{\nu_{i+1} - \nu_i, \rho_i, i \geq n\}$ be independent of $\{x(s), s < \nu_n, \nu_{i+1} - \nu_i, \rho_i, i < n\}$. There are such $\{\nu_n, \rho_n\}$ such that we can write the n^{th} jump of the process $x(\cdot)$ as $q(x(\nu_n^-), \rho_n)$, and the jump term as

$$J(t) = \sum_{\nu_n \leq t} q(x(\nu_n^-), \rho_n). \quad (6.2)$$

We can define the Poisson measure in terms of these quantities as follows, where H is a measurable subset of Γ : Define $N(t, H) = N([0, t] \times H)$ where

$$N(t, \Gamma) = \max\{n : \nu_n \leq t\} = \sum_n I_{\{\nu_n \leq t\}},$$

$$N(t, H) = \sum_n I_{\{\nu_n \leq t\}} I_{\{\rho_n \in H\}}. \quad (6.3)$$

The term $N(t, H)$ is just the number of jumps with values in the set H on the interval $[0, t]$, and $N(t, \Gamma) =$ number of jumps on the interval $[0, t]$.

A Convention Concerning “Zero” Jumps. Let $\tilde{\rho}$ be a random variable with the distribution $\Pi(\cdot)$. It is conceivable that $q(x, \tilde{\rho}) = 0$ with a positive probability for some x . Equivalently, it is possible that $q(x(\nu_n^-), \rho_n) = 0$ with a positive probability for some n . In this case, the actual physical process $x(\cdot)$ does not jump at ν_n even though the “driving” process $N(\cdot)$ does. For the sake of uniformity of expression, we still say that there is a jump at that time, but with value zero. The jumps are *defined* by the times ν_n .

Local Properties of the Jumps of (6.1). Because $\nu_{n+1} - \nu_n$ is exponentially distributed, we can write

$$P\{x(\cdot) \text{ jumps on } [t, t + \Delta] | x(s), w(s), N(s), s \leq t\} = \lambda\Delta + o(\Delta). \quad (6.4)$$

Define

$$\bar{\Pi}(x, H) = \Pi\{\rho : q(x, \rho) \in H\}.$$

By the independence properties and the definition of ρ_n , for $H \subset \Gamma$ we have

$$\begin{aligned} P\{x(t) - x(t^-) \in H | \text{jump at } t, x(t^-) = x, w(s), x(s), N(s), s < t\} \\ = \Pi\{\rho : q(x(t^-), \rho) \in H\} = \bar{\Pi}(x(t^-), H). \end{aligned} \quad (6.5)$$

5.6.2 CONSTRUCTING THE APPROXIMATING MARKOV CHAIN

It is implied by the above discussion that the jump diffusion $x(\cdot)$ satisfying (6.1) can be viewed as a process which evolves as a diffusion (3.1), with jumps which occur at random according to the rate defined by (6.4). If we wish to delete the jumps of magnitude “zero,” then we can use the state dependent jump rate

$$\lambda(x) = \lambda\Pi\{\rho : q(x, \rho) \neq 0\}.$$

We will, in fact, always allow zero jumps and use the jump rate λ . Given that the n^{th} jump occurs at time ν_n , we construct its values according to the conditional probability law (6.5) or, equivalently, write it as $q(x(\nu_n^-), \rho_n)$, and then let the process continue as a diffusion (3.1) until the time of the next jump. The locally consistent approximating Markov chain can be constructed in an analogous way, and we now proceed to give the details.

Definition of Local Consistency. The desired approximating Markov chain $\{\xi_n^h, n < \infty\}$ needs to mimic the local behavior of (6.1) in a way that is analogous to what was required by (4.1.3) for the diffusion model. The only difference between the cases of Section 4.1 [i.e., (3.1)] and our case (6.1) is the presence of the jumps. Let $q_h(\cdot)$ be a bounded (uniformly in h) measurable function such that

$$|q_h(x, \rho) - q(x, \rho)| \rightarrow 0,$$

uniformly in x in any compact set for each $\rho \in \Gamma$, and which satisfies $x + q_h(x, \rho) \in S_h$ for $x \in S_h$. The $q_h(x, \rho)$ will be used to get the approximations to the jump terms of (6.1). The condition that $x + q_h(x, \rho)$ takes points in S_h to points in S_h for each possible value $\rho \in \Gamma$ is needed because S_h is the state space of the chain.

A controlled Markov chain $\{\xi_n^h, n < \infty\}$ is said to be *locally consistent* with (6.1) if there is an interpolation interval $\Delta t^h(x, \alpha)$ which goes to zero as $h \rightarrow 0$ uniformly in x and α , and such that:

- (a) There is a transition probability $p_D^h(x, y|\alpha)$ which with an interpolation interval $\Delta t^h(x, \alpha)$ is locally consistent with (3.1) in the sense that (4.1.3) holds;
- (b) There is $\delta^h(x, \alpha) = o(\Delta t^h(x, \alpha))$ such that the one-step transition probability $p^h(x, y|\alpha)$ for the chain can be represented in the factored form:

$$\begin{aligned} p^h(x, y|\alpha) &= (1 - \lambda \Delta t^h(x, \alpha) - \delta^h(x, \alpha)) p_D^h(x, y|\alpha) \\ &\quad + (\lambda \Delta t^h(x, \alpha) + \delta^h(x, \alpha)) \Pi\{\rho : q_h(x, \rho) = y - x\}. \end{aligned} \quad (6.6)$$

An Interpretation of (6.6). The rule (6.6) is equivalent to the following procedure. Let $u^h = \{u_n^h, n < \infty\}$ be an admissible control sequence. Let the time be n and let $\xi_n^h = x, u_n^h = \alpha$. Flip a coin with the probability of Heads being $(1 - \lambda \Delta t^h(x, \alpha) - \delta^h(x, \alpha))$. At time n , this probability depends only on ξ_n^h and the current control, and not otherwise on the past. If Heads occurs, then the next state ξ_{n+1}^h is determined via use of the “diffusion” transition probability $p_D^h(x, y|\alpha)$. If Tails occurs, then the new state can be represented as $x + q_h(x, \tilde{\rho})$, where $\tilde{\rho}$ has the distribution $\Pi(\cdot)$ and is “independent of the past.” Thus, as the chain evolves, there is a sequence of coin tosses, and corresponding Heads and Tails. Let v_n^h denote the time of occurrence of the n^{th} Tail. There is a sequence of random variables $\{\rho_n, n < \infty\}$ where ρ_n has the distribution $\Pi(\cdot)$ and is “independent of the past”

$$\{\xi_i^h, u_i^h, i \leq v_n^h, \rho_j, v_j^h, j < n, v_n^h\}$$

and such that the next state of the chain is defined by

$$\xi_{v_n^h+1}^h - \xi_{v_n^h}^h = q_h(\xi_{v_n^h}^h, \rho_n). \quad (6.7)$$

The Dynamic Programming Equation. The definition of local consistency which was used above seems to be the simplest from the point of view of the writing of the dynamic programming equations. Let G be a given set and suppose that the cost function for the chain is

$$W^h(x, u) = E_x^u \left[\sum_{n=0}^{N_h-1} e^{-\beta t_n^h} k(\xi_n^h, u_n^h) \Delta t_n^h + e^{-\beta t_{N_h}^h} g(\xi_{N_h}^h) \right],$$

where $\beta > 0$ and N_h is the first exit time from the set $G_h^0 = S_h \cap G^0$. Then the dynamic programming equation is, for $x \in G_h^0$,

$$\begin{aligned} V^h(x) = \min_{\alpha \in \mathcal{U}} \left[& e^{-\beta \Delta t^h(x, \alpha)} (1 - \lambda \Delta t^h(x, \alpha) - \delta^h(x, \alpha)) \sum_y p_D^h(x, y | \alpha) V^h(y) \right. \\ & + e^{-\beta \Delta t^h(x, \alpha)} (\lambda \Delta t^h(x, \alpha) + \delta^h(x, \alpha)) \int_{\Gamma} V^h(x + q_h(x, \rho)) \Pi(d\rho) \\ & \left. + k(x, \alpha) \Delta t^h(x, \alpha) \right], \end{aligned}$$

with the boundary conditions $V^h(x) = g(x)$, $x \notin G_h^0$.

This equation might seem formidable, but in applications it generally takes a simple form. For example, consider the special two dimensional case where (6.1) takes the form

$$dx = b(x, u)dt + \sigma(x)dw + \begin{pmatrix} 0 \\ dJ \end{pmatrix},$$

where the jumps in $J(\cdot)$ are either ± 1 , each with probability $1/2$. Let $1/h$ be an integer, and suppose that the points in S_h are h units apart in the e_2 direction. Then the integral in the above expression is just

$$\frac{1}{2} [V^h(x + e_2) + V^h(x - e_2)].$$

Of course, any suitable approximation to the exponentials in the equation for $V^h(x)$ can be used, and $\delta^h(x, \alpha)$ can be set to zero if we wish.

The Interpolated Process $\psi^h(\cdot)$. The continuous parameter Markov process interpolation is defined exactly as in Section 4.3, and we continue to use the terminology of that section. Recall the definition of the jump times τ_n^h given above (4.3.1). We need to distinguish the jump times of $\psi^h(\cdot)$ which are due to the approximation to the jumps of $J(\cdot)$ from the other jump times. Define $\nu_n^h = \tau_{v_n^h+1}^h$ to be the moments of change of state of $\psi^h(\cdot)$ which correspond to the v_n^h for the discrete parameter chain. By (6.7)

$$\psi^h(\nu_n^h) - \psi^h(\nu_n^h^-) = q_h(\psi^h(\nu_n^h^-), \rho_n). \quad (6.8)$$

See Figure 5.4 for a concrete illustration of the terms.

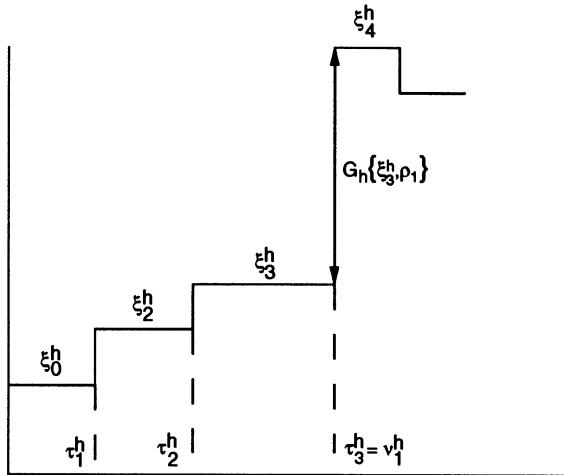


Figure 5.4. The continuous time interpolation.

5.6.3 A CONVENIENT REPRESENTATION OF $\{\xi_n^h, n < \infty\}$ AND $\psi^h(\cdot)$

It will be shown next that $\psi^h(\cdot)$ can be represented in the form (4.3.9) with a “jump term” added. This representation will be quite useful in the proofs of the convergence theorems of Chapter 10. Recall the Head-Tail terminology introduced below (6.6). Let $\xi_0^h = x$, and suppose that T_n^h and H_n^h , respectively, denote the events that Tails and Heads, respectively, occurred at the n^{th} step of the discrete time chain. Let E_n^h denote the expectation conditioned on the “data up to time n ,” i.e., on $\{\xi_i^h, u_i^h, H_i^h, i \leq n; v_j^h, \rho_j : v_j^h < n\} = \mathcal{D}_n^h$. We can write

$$\begin{aligned}\xi_n^h &= x + \sum_{i=0}^{n-1} \Delta \xi_i^h I_{H_i^h} + \sum_{i=0}^{n-1} \Delta \xi_i^h I_{T_i^h} \\ &= x + \sum_{i=0}^{n-1} E_i^h \Delta \xi_i^h I_{H_i^h} + \sum_{i=0}^{n-1} [\Delta \xi_i^h - E_i^h \Delta \xi_i^h] I_{H_i^h} + \sum_{i: v_i^h < n} q_h(\xi_{v_i^h}^h, \rho_i).\end{aligned}$$

Using the values of $E_i^h \Delta \xi_i^h$ from Section 4.3, the last equation can be written as

$$\xi_n^h = x + \sum_{i=0}^{n-1} [b(\xi_i^h, u_i^h) \Delta t_i^h + o(\Delta t_i^h)] I_{H_i^h} + M_n^h + J_n^h, \quad (6.9)$$

where M_n^h and J_n^h are the middle and right hand sums, respectively, in the previous equation. Due to the centering of its terms about the conditional expectation, given the “past,” M_n^h is a martingale with respect to

the filtration determined by \mathcal{D}_n^h and has the quadratic variation

$$\sum_{i=0}^{n-1} [a(\xi_i^h) \Delta t_i^h + o(\Delta t_i^h)] I_{H_i^h}. \quad (6.10)$$

For each t ,

$$E[\text{number of } n : \nu_n^h \leq t] \rightarrow \lambda t$$

as $h \rightarrow 0$. This implies that we can drop the $I_{H_i^h}$ in (6.9) and (6.10) with no effect in the limit. Using this reasoning, and the reasoning which led to the representation (4.3.9), we can write

$$\psi^h(t) = x + \int_0^t b(\psi^h(s), u^h(s)) ds + M^h(t) + J^h(t) + \delta_1^h(t), \quad (6.11)$$

where $E \sup_{s \leq t} |\delta_1^h(s)| \rightarrow 0$ as $h \rightarrow 0$. In (6.11), we define $u^h(\cdot)$ as the interpolation of the admissible control sequence $\{u_n^h, n < \infty\}$ which was defined below (4.3.1') and

$$J^h(t) = \sum_{n: \nu_n^h \leq t} q_h(\psi^h(\nu_n^h), \rho_n). \quad (6.12)$$

$M^h(\cdot)$ is a martingale whose discontinuities go to zero as $h \rightarrow 0$, and with quadratic variation

$$\int_0^t a(\psi^h(s)) ds + \delta_2^h(t),$$

where $E \sup_{s \leq t} |\delta_2^h(s)| \rightarrow 0$ as $h \rightarrow 0$.

5.6.4 A FINITE DIFFERENCE METHOD FOR CONSTRUCTING THE APPROXIMATING CHAIN

When a Markov chain which is locally consistent with (3.1) can be found by the finite difference method, then the finite difference method can be used to get a chain which is locally consistent with (6.1). There is no need to go through the procedure because the method outlined above, which starts with a chain which represents the diffusion and then adds the effects of the jump terms, covers any such case. So only a few comments will be made.

The differential operator \mathcal{L}^α of (6.1) with the control fixed at the value α is defined by (see Section 1.5)

$$\begin{aligned} \mathcal{L}^\alpha f(x) &= f'_x(x)b(x, \alpha) + \frac{1}{2} \sum_{ij} f_{x_i x_j}(x)a_{ij}(x) \\ &\quad + \lambda \int [f(x + q(x, \rho)) - f(x)] \Pi(d\rho). \end{aligned} \quad (6.13)$$

As in Section 5.3, the finite difference approach starts with an approximation to the equation

$$\mathcal{L}^{u(x)} W(x, u) + k(x, u(x)) = 0,$$

using appropriately chosen finite difference approximations to the derivatives. Any of the approximations in Section 5.3 can be tried. Let $W^h(x, u)$ denote the solution to the finite difference approximation.

In order to approximate the integral in (6.13), we start by getting a distribution function $\bar{\Pi}^h(x, \cdot)$ which converges in distribution to $\bar{\Pi}(x, \cdot)$ uniformly in x as $h \rightarrow 0$ and which is concentrated on a finite number of points $\delta x_j^h(x)$, $j = 1, 2, \dots$, such that for $x \in S_h$ we have $x + \delta x_j^h(x) \in S_h$. Then use the approximation

$$\lambda \sum_j [f(x + \delta x_j^h(x)) - f(x)] \bar{\Pi}^h(x, \delta x_j)$$

for the integral. This approach is equivalent to the use of $q_h(\cdot)$ above.

The procedure is completed as it was in Section 5.3. Collect terms and divide all terms by the coefficient of $W^h(x, u)$. The transition probabilities $p^h(x, y|\alpha)$ are the coefficients of the $W^h(y, u)$, $y \neq x$, in the finite difference equation, provided that they are all non-negative. The interpolation interval is the coefficient of $k(x)$, provided that it is positive and goes to zero as $h \rightarrow 0$.

5.7 Approximations for Reflecting Boundaries

5.7.1 GENERAL DISCUSSION

Suppose that the controlled process is of interest in a compact set G which is the closure of its interior G^0 . The approximations of the controlled processes in Sections 5.1 to 5.6 have been concerned with getting Markov chains which are locally consistent with (3.1) or (6.1) at all points x in a state space S_h , and the behavior on a boundary was not considered. If the original control process (3.1) or (6.1) is stopped as soon as the sample path leaves the open set G^0 , then for the control problem for the approximating chain one simply stops the chain the first time that it leaves $G_h^0 = G^0 \cap S_h$. For part of the problems discussed in Section 1.4 and in Chapters 3 and 8, the boundary is reflecting or constraining. For these cases, we need to reflect or project the approximating chain back into the set G whenever it leaves G , in a way that is consistent with the set of allowed reflection directions. Essentially all that we need when $\xi_n^h = x$ is not in $G_h = G \cap S_h$ is that the conditional mean change $E_{x,n}^{h,\alpha} \Delta \xi_n^h$ approximate the desired reflection direction. This can be readily done. It will be seen that, as with the construction of the locally consistent chain itself, there are numerous

possibilities, and, generally speaking, any intuitively reasonable method can be used.

The Reflected Diffusion Model. We begin by considering the reflected diffusion model discussed in Section 1.4, and comment later on the adjustments needed for the reflected jump diffusion. Let us recapitulate the reflecting diffusion model and the associated conditions of Section 1.4. Thus, for each $x \in \partial G$, we are given a set of direction vectors $r(x)$ whose members are of unit length. In many cases, the set contains only a single vector, and then we abuse the terminology and refer to that singleton as $r(x)$ also.

The model of the reflected diffusion given in Section 1.4 is that of the Skorokhod Problem

$$x(t) = x + \int_0^t b(x(s), u(s))ds + \int_0^t \sigma(x(s))dw(s) + z(t), \quad (7.1)$$

where the “reflecting” term $z(\cdot)$ is continuous and keeps the process $x(\cdot)$ from leaving G . It is required to satisfy

$$|z|(t) = \int_0^t I_{\partial G}(x(s))d|z|(s), \quad z(t) = \int_0^t \gamma(s)d|z|(s), \quad (7.2)$$

where $\gamma(s) \in r(x(s))$ almost surely (ω, s) with respect to the random measure induced by $|z|(\cdot)$.

We next list the assumptions on $r(\cdot)$ that will be used. Following the statements of the assumptions is a discussion of their significance and references to the literature.

The following conditions are assumed to hold. These conditions will be used in the convergence theorems of Chapter 11.

(i) For each x , the positive cone generated by the vectors in $r(x)$ is convex.

(ii) The set G can be constructed as the intersection of a finite number of “smooth” sets in the following way. There are a finite number of continuously differentiable functions $g_i(\cdot)$ and sets $G_i = \{x : g_i(x) \leq 0\}$ whose boundaries are $\partial G_i = \{x : g_i(x) = 0\}$ and such that $G = \cap_i G_i$. The sets G and each G_i are assumed to be the closure of their interiors.

(iii) Suppose that for a given $x \in \partial G$ there is a single index $i = i(x)$ such that $g_i(x) = 0$, and let $n(x)$ denote the interior normal to $\partial G_{i(x)}$ at x . Then the inner product of all the vectors in $r(x)$ with $n(x)$ is positive.

(iv) Define the index set $I(x) = \{i : x \in \partial G_i\}$. $I(x)$ is upper semicontinuous in the sense that if $x \in \partial G$, there is $\delta > 0$ such that $|x - y| < \delta$ implies that $I(y) \subset I(x)$. Next, suppose that $x \in \partial G$ lies in the intersection of more than one boundary; i.e., $I(x)$ has the form $I(x) = \{i(1), \dots, i(k)\}$ for some $k > 1$. Let $N(x)$ denote the convex hull of the interior normals $n_{i(1)}, \dots, n_{i(k)}$ to $\partial G_{i(1)}, \dots, \partial G_{i(k)}$ at x . Let there be some vector $v \in N(x)$, such that $v'v > 0$ for all $v \in r(x)$.

In (7.1) and (7.2), $r(x)$ needed to be defined only for $x \in \partial G$. For our purposes of “approximation,” the definition needs to be extended so that $r(x)$ is defined and satisfies an appropriate “upper semicontinuity” condition in an appropriate “outer” neighborhood of the boundary. More precisely, assume the following.

(v) Let $N(G) \supset G$ be a neighborhood of G . There exists an extension of $r(\cdot)$ to $\overline{N(G)} - G^0$ which is *upper semicontinuous* in the following sense: Let $x_n \notin G^0$. If $x_n \rightarrow x \in \partial G$ and if $\gamma_n \rightarrow \gamma$ with $\gamma_n \in r(x_n)$, then $\gamma \in r(x)$.

The set $\overline{N(G)} - G^0$ will contain the “discretized” reflecting boundary ∂G_h^+ defined in the next subsection.

Remark. The theory of reflecting diffusion processes is currently an active area of research. To meet the demands of various application areas, a wide variety of assumptions on G and $r(\cdot)$ have been considered in the literature. It seems likely that this trend will continue, with even more unusual assumptions on the reflection term becoming commonplace. On the other hand, the weakest assumptions that are required so that weak sense uniqueness and existence hold are still far from clear. For these reasons two criteria have been applied in formulating the assumptions used in this book. The first criterion requires that the assumptions be flexible enough to accommodate as many sets of assumptions that are of current interest as possible. At the same time, the assumptions used must be strong enough that the approximations to reflecting diffusions that occur in the numerical schemes can be shown to converge. The set of conditions given above satisfy this second requirement, and yet are flexible enough that they simultaneously cover a number of quite different cases of interest. In particular, they include those which arise in the “heavy traffic” models of Chapter 8, where the reflection directions are discontinuous and multivalued at the “corners”. In most current applications, the set $r(x)$ contains only one direction, except possibly at the “corners” of ∂G .

In the remainder of this remark we will try to motivate each of the assumptions and also cite the relevant literature regarding existence and uniqueness results.

Conditions (i) and (v) are standard in the literature, although they may be implied rather than explicitly stated. They are related to the fact that reflecting diffusions often appear as weak limits. For example, in the “heavy traffic” models of Chapter 8 the set G is usually the non-negative orthant in some Euclidean space, and $r(x)$ is independent of x and single valued on the relative interior of each “face” of G . The definition of $r(x)$ is then extended to all of ∂G by the sort of “semicontinuity” assumed in (v) together with the convexity assumed in (i). Assumption (ii) describes the regularity of the boundary. It is formulated in such a way that it covers both the fairly standard case in which ∂G is a smooth manifold as well as the more novel setting in which G is a polytope. The formulation of conditions (i), (ii) and (v) follows [35], which considers questions of strong existence and unique-

ness. Related works that also deal with strong existence and uniqueness of solutions are [57, 92, 106, 117].

The most interesting conditions are (iii) and its extension in (iv). Basically, these assumptions guarantee that the reflection term must always point “into” the domain G . They allow a key estimate on the reflection term of the process in terms of the drift and diffusion components. This estimate will be exploited in Chapter 11 to prove the weak convergence of processes that is needed for the convergence proofs for the numerical schemes. This type of assumption has been used previously in proving existence in [36]. For the special case when G is the r -dimensional non-negative orthant in \mathbb{R}^r and $r(x)$ is independent of x and single valued on the relative interior of each “face” of ∂G , the condition (iv) is equivalent to the “completely- S ” condition used in [119, 118]. In these papers the condition has been used to prove weak existence and uniqueness for the special case of a reflecting Wiener process (i.e. $b(\cdot, \cdot) = 0$ and $\sigma(\cdot) = I$).

As pointed out in Chapter 1, the term “constrained process” is often more appropriate than the term “reflected process.” In many of the applications in which reflecting diffusion models arise, the reflection term is actually a constraining term. For example, in the so-called “heavy traffic” models discussed in Chapter 8, the reflection term is what keeps the buffers of certain queues from either becoming negative or overflowing. This is more of a constraint than a reflection, in the physical sense. Reflecting boundaries are often artificially introduced in order to get a bounded region in which the numerical approximation can be carried out. Many control problems are originally defined in an unbounded space, which is inconvenient for numerical purposes. One tries to truncate the space in such a way that the essential features of the control problem are retained. We can do this by simply stopping the process when it first leaves some fixed and suitably large set. One must then introduce some boundary condition or cost on the stopping set. The objective is to choose a boundary cost that is believed to be close to whatever the cost would be at these points for the untruncated process. In particular, we do not want the chosen boundary and boundary condition to seriously distort the optimal control at points not near that boundary. An alternative is to introduce an appropriate reflecting boundary.

5.7.2 LOCALLY CONSISTENT APPROXIMATIONS ON THE BOUNDARY

Recall the definition $G_h = G \cap S_h$. Assume that the transition probabilities and interpolation interval $p^h(x, y|\alpha)$, $\Delta t^h(x, \alpha)$, respectively, of a Markov chain which is locally consistent with (3.1) have already been defined for all points in S_h . Let the distance between communicating states of this

chain be bounded above and below by some positive constant times h . The aim of the reflection or constraint is to keep the process in the set G , if it ever attempts to leave it. We will use ∂G_h^+ to denote the “reflecting boundary” for the approximating chain. The set ∂G_h^+ is always disjoint from G_h . This is often a convenience in programming, although one can redo the formulation so that the “reflecting boundary” is simply “near” ∂G whether on the inside or outside. The transition probabilities at the states in ∂G_h^+ are chosen so as to “mimic” the behavior of the reflection for (7.1) and (7.2). The reflection direction is not controlled here, although there are some comments on the algorithm when such boundary controls are allowed at the end of Chapter 7. Thus, we may use $p^h(x, y)$ to denote the transition function for points $x \in \partial G_h^+$. The reflecting boundary is defined in a natural way as follows. ∂G_h^+ includes all points y in $S_h - G_h$ for which $p^h(x, y|\alpha) > 0$ for some $x \in G_h$ and $\alpha \in \mathcal{U}$. It also is “closed” in the sense that it includes all points y such that $p^h(x, y) > 0$ for $x \in \partial G_h^+$. We suppose that the “radius” of ∂G_h^+ goes to zero as $h \rightarrow 0$ in the sense that

$$\lim_{h \rightarrow 0} \sup_{x \in \partial G_h^+} \text{dist}(x, G) = 0.$$

Let $\{\xi_n^h, n < \infty\}$ be a Markov process with the transition probabilities $p^h(x, y)$ on ∂G_h^+ and $p^h(x, y|\alpha)$ on G_h . We say that the transition function $p^h(x, y)$ is *locally consistent with the reflection directions* $r(\cdot)$ if there are $\epsilon_1 > 0$, $c_1 > 0$ and $c_2(h) \rightarrow 0$ as $h \rightarrow 0$, such that for all $x \in \partial G_h^+$ and all h ,

$$E_{x,n}^{h,\alpha}(\xi_{n+1}^h - \xi_n^h) \in \{\theta\gamma + o(h) : c_2(h) \geq \theta \geq c_1 h, \gamma \in r(x)\}, \quad (7.3a)$$

$$\text{cov}_{x,n}^{h,\alpha}(\xi_{n+1}^h - \xi_n^h) = O(h^2), \quad (7.3b)$$

$$p^h(x, G_h) \geq \epsilon_1, \text{ all } h \text{ and } x \in \partial G_h^+. \quad (7.3c)$$

Condition (7.3c) is convenient, but it is sufficient to let it hold for the k -step transition probabilities, where k is an integer not depending on x . Because the chain is not controlled on ∂G_h^+ , the α in the $E_{x,n}^{h,\alpha}$ is redundant. Thus the mean direction of the increment is an “admissible reflection direction” plus a “small” error. If a Markov chain is locally consistent with (3.1) in G and is also locally consistent with the reflection directions $r(x)$, as defined above, then we say that the chain is *locally consistent with the reflected diffusion* (7.1), (7.2). See Figure 5.5 for an illustration, where G is the shaded area and ∂G_h^+ is the set of dots, and we suppose that the “dots” communicate only with each other or with points in G_h and are the only points accessible from G_h .

The conditions (7.3) are needed for the convergence theorem in Chapter 11. For any fixed value of h , we would try to choose the $p^h(x, y)$ for $x \in \partial G_h^+$ so that the behavior of the chain on ∂G_h^+ copies as closely as possible the

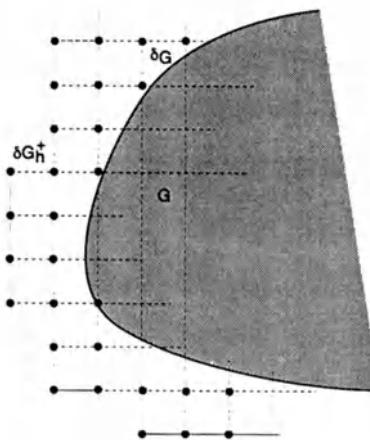


Figure 5.5. The boundary approximation.

behavior of the physical model on ∂G . This is particularly helpful when there is nonuniqueness of the directions. See the comments at the end of Example 2 below.

The Interpolated Process $\psi^h(\cdot)$. For $x \in \partial G_h^+$, we define the interpolation interval $\Delta t^h(x)$ to be zero. The reason for this is the correspondence of the set ∂G_h^+ with the reflecting boundary ∂G , and the fact that the role of the reflection is to merely keep the process from leaving the desired state space G . Indeed, the “instantaneous” character is inherent in the definition of the reflected diffusion (7.1), (7.2).

Define τ_n^h as in Section 4.3, but use the interpolation interval $\Delta t^h(x) = 0$ for $x \in \partial G_h^+$. Thus, if $\xi_n^h \in \partial G_h^+$ (i.e., n is the time of a “reflection” step), then $\tau_{n+1}^h = \tau_n^h$. Define the continuous time parameter Markov chain interpolation $\psi^h(\cdot)$ by (4.3.1'). Note that the expression (4.3.1) cannot be used to define $\psi^h(\cdot)$, because (4.3.1) is multi-valued at those n which are reflection steps. An alternative to (4.3.1') is: $\psi^h(\tau_n^h) = \xi_{m(n)}^h$, where $m(n) = \max\{i : \tau_i^h = \tau_n^h\}$. The values of the states at the moments of the reflection steps do not appear in the definition of the interpolation. In this sense these states are “instantaneous,” and the reflection is instantaneous for the continuous parameter interpolation.

The reflecting states can actually be removed from the problem formulation by using the multistep transition functions which eliminate them. It is often useful to retain them to simplify the coding of the computational algorithms for the solution of the dynamic programming equations of, e.g., Section 5.8 below, as well as to facilitate the convergence proofs.

5.7.3 THE CONTINUOUS PARAMETER MARKOV CHAIN INTERPOLATION

For use in Chapter 11, we now give a representation of the continuous parameter interpolation which is analogous to (6.11) or (4.3.9). Let n be a reflection step and define Δz_n^h by $E_n^h \Delta \xi_n^h = \Delta z_n^h$. For programming simplicity, one normally chooses a function $\Delta z(\cdot)$ such that Δz_n^h defined by $\Delta z_n^h = \Delta z(\xi_n^h)$ satisfies (7.3). Define $\Delta \tilde{z}_n^h$ by $\Delta \xi_n^h = \Delta z_n^h + \Delta \tilde{z}_n^h$. Define $\Delta z_n^h = \Delta \tilde{z}_n^h = 0$ if n is not a reflection step. Define the continuous parameter interpolations

$$z^h(t) = \sum_{n:\tau_{n+1}^h \leq t} \Delta z_n^h, \quad \tilde{z}^h(t) = \sum_{n:\tau_{n+1}^h \leq t} \Delta \tilde{z}_n^h.$$

Then (6.11) (or (4.3.9) if there is no jump term) can be extended as

$$\psi^h(t) = x + \int_0^t b(\psi^h(s), u^h(s))ds + M^h(t) + J^h(t) + z^h(t) + \tilde{z}^h(t) + \delta_1^h(t), \quad (7.4)$$

where all the terms are as in (6.11), except for the just defined reflection terms.

It turns out that $\tilde{z}^h(t) \rightarrow 0$ in mean square. This implies that in the limit only the mean reflection directions count. The effects of the perturbations about the mean go to zero. As a first step in the proof note that because this process is martingale (when sampled at the moments of change), we can write

$$\begin{aligned} E \sup_{m \leq n} \left| \sum_{i=0}^{n-1} \Delta \tilde{z}_i^h \right|^2 &= O(h^2) E(\text{number of reflection steps on } [0, n]) \\ &\leq O(h) E \left| \sum_{i=0}^{n-1} \Delta z_i^h \right|. \end{aligned} \quad (7.5)$$

The proof will be completed in Chapter 11.

5.7.4 EXAMPLES

Example 1. Getting a transition function which is locally consistent with the boundary reflection is often quite straightforward, as will be illustrated by these simple examples. First consider the case where G is the closed right half plane in $\mathbb{R}^2 : G = \{(x_1, x_2) : x_1 \geq 0\}$, which is the shaded area illustrated in Figure 5.6. The reflection direction, denoted by $r(x)$, is the vector of unit length which has direction $(2, 1)$. Let the state space S_h be the regular grid \mathbb{R}_h^2 . Then $\partial G_h^+ = \{(-h, ih) : -\infty < i < \infty\}$. In this case it is very easy to find $p^h(x, y)$ satisfying (7.3). The choice which appears simplest is obviously $p^h(x, x + (h, 0)) = p^h(x, x + (h, h)) = 1/2$. Here, we have obtained the proper average direction by *randomizing* among two

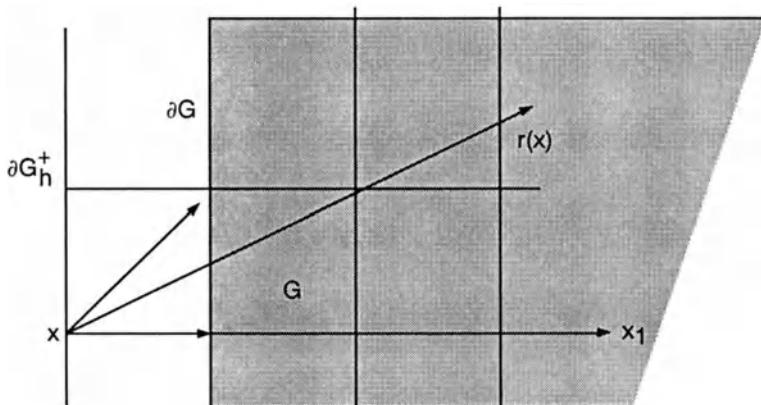


Figure 5.6. Directions and transitions for Example 1.

“allowed” directions. The effects of the perturbations about the mean value will disappear as $h \rightarrow 0$. Similar randomizations will be used in Chapter 8 to calculate the controls for singular control problems.

Example 2. We next consider a problem for the same set G in Figure 5.6, but now the reflection direction is the unit vector in the direction $(2, 3)$. See Figures 5.7a,b. Here, several possibilities are of interest. Clearly, this direction cannot be achieved as a convex combination of the vectors $(1, 0)$ and $(1, 1)$. There are several alternative constructions. For example, $(2, 3)$ is in the convex cone generated by the vectors $(1, 1)$ and $(1, 2)$, and in fact we can take $p^h(x, x + (h, h)) = p^h(x, x + (h, 2h)) = 1/2$, as in Figure 5.7a. An alternative is to exploit the possibility of transitions between states in ∂G_h^+ . For example, we can take $p^h(x, x + (h, h)) = 2/3$ and $p^h(x, x + (0, h)) = 1/3$, as in Figure 5.7b. It is not clear which alternative is preferable.

A slight variation of this problem appears in Figure 5.8, where G is the closed northeast quadrant, and the two sides have different directions of reflection, with r_1 on the horizontal axis and r_2 on the vertical. The set $r(0)$ is composed of the convex cone generated by these directions. Thus, there are several possibilities for the reflection directions at the points $(-h, 0), (-h, -h), (0, -h) \in \partial G_h^+$. The choices indicated in the figure are reasonable for this problem. The physical origin of the problem often suggests appropriate directions in such cases of nonuniqueness. This example is typical of those arising as “heavy traffic limits,” as in Chapter 8. For that case, the reflections serve the purposes of maintaining the non-negativity of the contents of the buffers and also keeping them from exceeding their capacities.

Example 3. Next consider the two dimensional example illustrated in

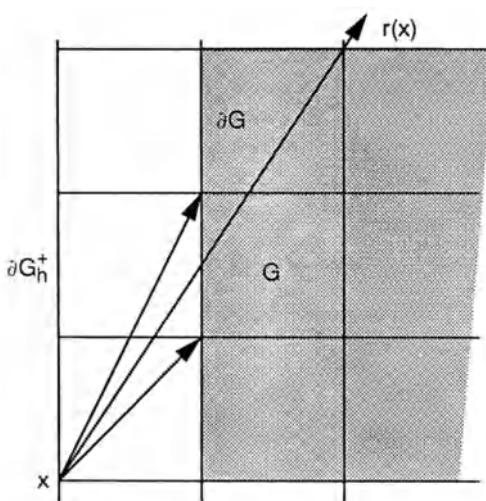


Figure 5.7a. Example 2, choice 1.

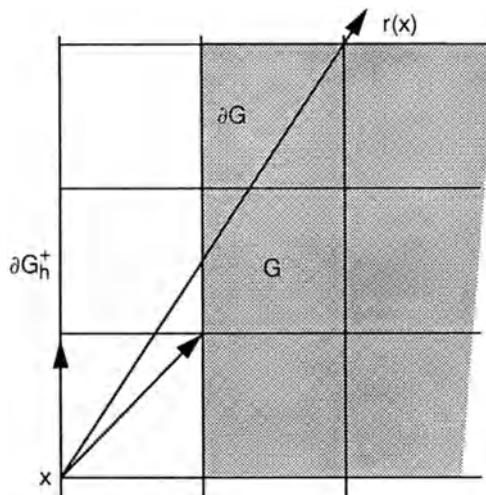


Figure 5.7b. Example 2, choice 2.

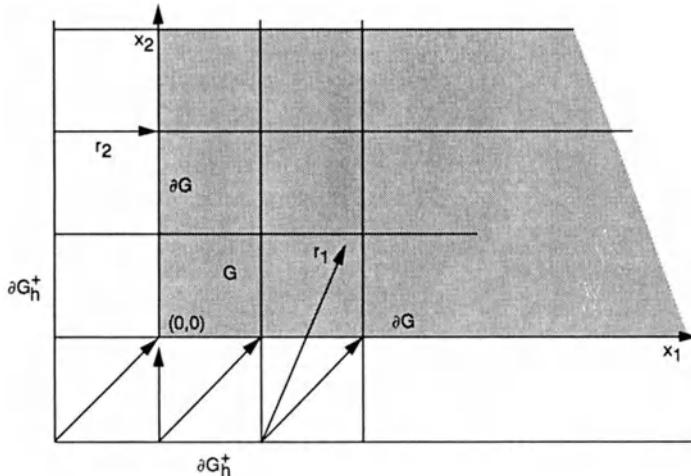


Figure 5.8. Example 3 with a corner.

Figure 5.9, where the boundary is curved and the directions depend on x . Suppose that $r(\cdot)$ is single valued. Consider the point $x \in \partial G_h^+$. Going in a direction $r(x)$ from x , we meet line connecting the two points y_1 and y_2 in S_h at the point $y_0 \notin S_h$. To meet the condition (7.3), define the transition function

$$p^h(x, y_1) = \frac{y_{21} - y_{01}}{h}, \quad p^h(x, y_2) = \frac{y_{01} - y_{11}}{h},$$

where y_{ij} is the j^{th} component of the vector y_i . It is straightforward to define ∂G_h^+ such that the “randomizing” procedure just described can be carried out for all points in ∂G_h^+ .

5.7.5 THE REFLECTED JUMP DIFFUSION

This is treated in the same way as was the reflected diffusion. We need only decide what to do if the process $x(\cdot)$ leaves the set G because of a jump. For the reflecting boundary problem, we suppose that $x + q(x, \rho) \in G$ for $x \in G$ and $\rho \in \Lambda$. Without serious restriction, we suppose that $x + q_h(x, \rho) \in G$ for $x \in G$ and $\rho \in \Lambda$.

5.8 Dynamic Programming Equations

5.8.1 OPTIMAL STOPPING

In this section, we will rewrite the dynamic programming equations for the problems in Chapter 2 in the notation of the controlled Markov chain approximations of this chapter. These will be needed in the discussion of numerical methods in Chapter 6. The functions $k(\cdot)$ and $g(\cdot)$ are assumed to

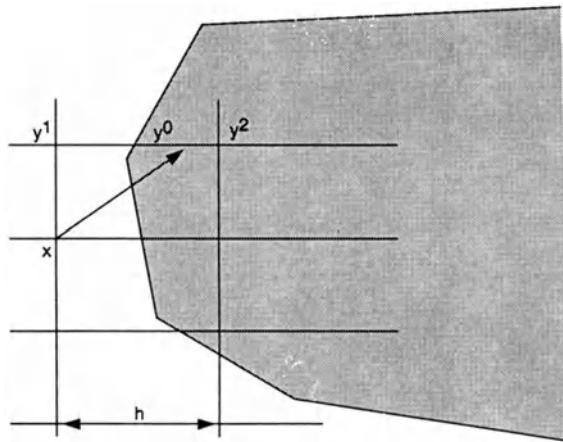


Figure 5.9. Example 3.

be bounded. We first treat the optimal stopping problem. The underlying model is (3.1.1) or (3.1.2) or, equivalently, (3.1) or (6.1) but with the control $u(\cdot)$ dropped. For a stopping time τ and a discount factor $\beta > 0$, let the cost for $x(\cdot)$ be the discounted form of (3.2.1):

$$W(x, \tau) = E_x \int_0^\tau e^{-\beta s} k(x(s)) ds + E_x e^{-\beta \tau} g(x(\tau)). \quad (8.1)$$

Here, the only control is the choice of the stopping time.

Let $W^h(x, N)$ denote the cost function for the approximating chain when the stopping time is N . For the Markov chain approximation, an appropriate cost is

$$W^h(x, N) = E_x \left[\sum_{n=0}^{N-1} e^{-\beta t_n^h} k(\xi_n^h) \Delta t_n^h + e^{-\beta t_N^h} g(\xi_N^h) \right], \quad (8.2)$$

where we recall the notation $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$ and $\Delta t_n^h = \Delta t^h(\xi_n^h)$. From Section 2.2, the dynamic programming equation for the infima of the costs is

$$V^h(x) = \min \left[e^{-\beta \Delta t^h(x)} \sum_y p^h(x, y) V^h(y) + k(x) \Delta t^h(x), g(x) \right]. \quad (8.3)$$

Due to the discounting, all the costs and equations are well defined. Any acceptable approximation to $e^{-\beta \Delta t^h(x)}$ can be used; for example, if h is small, then one can use either of

$$1 - \beta \Delta t^h(x), \quad \frac{1}{1 + \beta \Delta t^h(x)}.$$

Generally, for numerical purposes the process will be confined to some compact set G . Suppose that we will be *required* to stop on *first exit* from G , with a stopping cost $g(x)$, if we have not decided to stop before that time. To formulate this case, define

$$\tau' = \inf\{t : x(t) \notin G\}. \quad (8.4)$$

Thus we require that the stopping times for $x(\cdot)$ be no larger than τ' . In this case, (8.3) holds for $x \in G_h = G \cap S_h$; otherwise,

$$V^h(x) = g(x), \quad x \notin G_h. \quad (8.5)$$

Undiscounted Optimal Stopping. Continue with the setup of the last paragraph, but set $\beta = 0$, and refer to Section 2.2.2. Neither the cost for an arbitrary stopping time nor the dynamic programming equations are necessarily well defined, and special conditions are needed. Suppose that the stopping times must satisfy the condition below (8.4) and similarly for the approximating chains. Then, formally, (8.3) can be written as

$$V(x) = \begin{cases} \min \left[\sum_y p^h(x, y) V^h(y) + k(x) \Delta t^h(x), g(x) \right], & x \in G_h \\ g(x), & x \notin G_h. \end{cases} \quad (8.6)$$

The dynamic programming equation is well defined if the analogue of the conditions given in Section 2.2.2 hold. In particular, if (a) there is $k_0 > 0$, such that $k(x) \geq k_0$ all x , or (b) the mean time to the obligatory stopping set is bounded for each initial condition x . A case where neither condition holds is dealt with in the “shape from shading” example in Chapter 13.

Reflecting Boundary. Next, suppose that the process $x(\cdot)$ is a “reflecting” jump diffusion of the type dealt with in Section 5.7, where the “instantaneous” reflection constrains $x(\cdot)$ to stay in a given compact set G . The cost is still given by (8.1). Suppose that there is no obligatory stopping set and that $\beta > 0$. Let ∂G_h^+ denote the “reflecting boundary” for the approximating chain, as in Section 5.7. Then the dynamic programming equation is (8.3) for $x \in G_h$, and otherwise

$$V^h(x) = \sum_y p^h(x, y) V^h(y), \quad x \in \partial G_h^+. \quad (8.7)$$

Because the reflection is “instantaneous,” there is no discount factor in (8.7). It is possible to eliminate the reflecting states from the equation (8.3), as discussed below (8.23), but from the computational point of view it is generally more convenient to keep them. If $\beta = 0$, then the dynamic

programming equation is well defined under condition (a) in the above paragraph. In Chapter 8, there is a brief discussion of an example where a cost is associated with the reflection.

The Interpolated Process $\psi^h(\cdot)$. Recall the continuous parameter Markov chain interpolation $\psi^h(\cdot)$ from Section 4.3 or Subsection 5.7.3. For a stopping time τ for $\psi^h(\cdot)$, an appropriate analogue of the cost function (8.1) is

$$W^h(x, \tau) = E_x \left[\int_0^\tau e^{-\beta s} k(\psi^h(s)) ds + e^{-\beta \tau} g(\psi^h(\tau)) \right]. \quad (8.8)$$

By the results in Section 4.3, (8.8) is also the cost for the discrete parameter chain (with $\tau = \tau_N^h$ being the interpolated time which corresponds to N) if the discount factor $e^{-\beta \Delta t^h(x)}$ is approximated by $1/[1+\beta \Delta t^h(x)]$. Otherwise it is an approximation, with an error which goes to zero as $h \rightarrow 0$.

5.8.2 CONTROL UNTIL EXIT FROM A COMPACT SET

We next treat a discounted cost problem with discount factor $\beta \geq 0$, where the control stops on first exit τ from the interior G^0 of the compact set G . Also, suppose that there is no reflecting boundary. Suppose that the functions $k(x, \cdot)$ and $p^h(x, y | \cdot)$ are continuous in $\alpha \in \mathcal{U}$ for each x, y . Let $u = \{u_n^h, n < \infty\}$ be an admissible control sequence and define N_h to be the first exit time of the chain from the set $G_h^0 = S_h \cap G^0$. Also, define ∂G_h^0 to be the “boundary” set of states which are reachable from states in G^0 in one step under some control action. An appropriate cost for the chain is:

$$W^h(x, u) = E_x^u \left[\sum_{n=0}^{N_h-1} e^{-\beta t_n^h} k(\xi_n^h, u_n^h) \Delta t_n^h + e^{-\beta t_{N_h}^h} g(\xi_{N_h}^h) \right]. \quad (8.9)$$

Let $u_n^h = u(\xi_n^h)$ for a feedback control $u(\cdot)$. If the sum is well defined and finite for each x under $u(\cdot)$, then $W^h(x, u)$ satisfies the equation

$$W^h(x, u) = \sum_y e^{-\beta \Delta t^h(x, u(x))} p^h(x, y | u(x)) W^h(y, u) + k(x, u(x)) \Delta t^h(x, u(x)) \quad (8.10)$$

for $x \in G_h^0$, and with the boundary condition

$$W^h(x, u) = g(x), \quad \text{for } x \in \partial G_h. \quad (8.11)$$

The dynamic programming equation for the optimal value function is

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_y e^{-\beta \Delta t^h(x, \alpha)} p^h(x, y | \alpha) V^h(y) + k(x, \alpha) \Delta t^h(x, \alpha) \right] \quad (8.12)$$

for $x \in G_h^0$ with the boundary condition (8.11). If the interpolation intervals are small, then it might be convenient to replace the exponentials in (8.10) and (8.12) by $1 - \beta\Delta t^h(x, \alpha)$ or by $1/[1 + \beta\Delta t^h(x, \alpha)]$.

An appropriate cost function for the continuous parameter interpolation $\psi^h(\cdot)$ is

$$W^h(x, u^h) = E_x^{u^h} \left[\int_0^{\tau_h} e^{-\beta s} k(\psi^h(s), u^h(s)) ds + e^{-\beta \tau_h} g(\psi^h(\tau_h)) \right], \quad (8.13)$$

where τ_h is the first escape time of $\psi^h(\cdot)$ from G^0 , and $u^h(\cdot)$ is the continuous parameter interpolation of $\{u_n^h, n < \infty\}$. The remarks which were made in the previous subsection concerning the equivalence of this cost with that for the discrete parameter chain hold here also.

A More Compact Form of the Equations. For use in Chapter 6, it is convenient to write the above equations in a more compact matrix-vector form analogous to (2.2.5), (2.3.3), or (2.4.4). Let $u(\cdot)$ be a feedback control. Define the vectors $W^h(u) = \{W^h(x, u), x \in G_h^0\}$ and $V^h = \{V^h(x), x \in G_h^0\}$. Define the cost vector $C^h(u) = \{C^h(x, u), x \in G_h^0\}$ with components

$$C^h(x, u) = k(x, u(x))\Delta t^h(x, u(x)) + e^{-\beta\Delta t^h(x, u(x))} \sum_{y \in \partial G_h} p^h(x, y|u(x))g(y) \quad (8.14)$$

for $x \in G_h^0$. Define the matrix $R^h(u) = \{r^h(x, y|u(x)); x, y \in G_h^0\}$ where

$$r^h(x, y|\alpha) = e^{-\beta\Delta t^h(x, \alpha)} p^h(x, y|\alpha) \quad (8.15)$$

for $x \in G_h^0$. Then we can write the equation for the cost (8.10) and the dynamic programming equation (8.12) as

$$W^h(u) = R^h(u)W^h(u) + C^h(u), \quad (8.16)$$

$$V^h = \min_{u(x) \in \mathcal{U}} [R^h(u)V^h + C^h(u)]. \quad (8.17)$$

The minimum in the vector valued expression (8.17) is understood to be taken component by component.

5.8.3 REFLECTING BOUNDARY

Now consider the problem in the above subsection, where the boundary is “instantaneously” reflecting as in Section 5.7, and the transition probabilities for the reflecting states do not depend on the control. The approximating chain is assumed to be locally consistent with the diffusion (3.1) or the jump diffusion (6.1) in $G_h = S_h \cap G$, and with the reflection directions $r(x)$ on ∂G_h^+ , where the reflecting boundary is disjoint from G_h . The interpolation intervals $\Delta t^h(x)$ equal zero for points in ∂G_h^+ . An obligatory stopping

set and associated stopping cost can be added and will be commented on briefly below.

Suppose that the cost for (7.1) is

$$E_x^u \int_0^\infty e^{-\beta t} [k(x(t), u(t))dt + c'(x(t))dz(t)], \quad (8.18)$$

where $c(\cdot)$ is bounded and continuous and $c'(x)\gamma \geq 0$, for any $\gamma \in r(x), x \in \partial G$.

Let $u = \{u_n^h, n < \infty\}$ be an admissible control sequence. Write the “mean” increment when the state $\xi_n^h = x$ is in ∂G_h^+ as $\Delta z(x) = E_{x,n}^{h,\alpha} \Delta \xi_n^h$. Then, for this unstopped reflection problem, an appropriate analogue for the chain of the cost function (8.18) is

$$W^h(x, u) = E_x^u \left\{ \sum_{n=0}^{\infty} e^{-\beta t_n^h} [k(\xi_n^h, u_n^h) \Delta t_n^h + c'(\xi_n^h) \Delta z^h(\xi_n^h)] \right\}. \quad (8.19)$$

Let $u_n^h = u(\xi_n^h)$ for a feedback control $u(\cdot)$. Then, if the sum in (8.19) is well defined, $W^h(x, u)$ satisfies (8.10) for $x \in G_h$. For $x \in \partial G_h^+$, we have

$$W^h(x, u) = \sum_y p^h(x, y) W^h(y, u) + c'(x) \Delta z^h(x). \quad (8.20)$$

The equation for the optimal value function is (8.12) for $x \in G_h$, and for $x \in \partial G_h^+$ it is

$$V^h(x) = \sum_y p^h(x, y) V^h(y) + c'(x) \Delta z^h(x). \quad (8.21)$$

The “reflection” properties of the approximating Markov chain models the “instantaneous” reflection of the underlying diffusion or jump diffusion model. The interpolation time interval for the reflection steps equals zero, consistent with the instantaneous character of the reflection for (7.1), (7.2). Hence, there is no discounting associated with that step.

Suppose that $x(\cdot)$ is a reflecting diffusion process, but we allow the interpolation interval for the reflecting states to be nonzero and of the order of the step size. Then the limit processes might be “sticky” on the boundaries; i.e., spend positive time on the boundaries. An approximation for such a case was discussed in [75], but we omit it here because it seems to occur rarely in applications at this time.

A Vector Form of the Equations for the Cost. The formulas (8.16) and (8.17) hold with appropriate redefinitions. Redefine the vector $W^h(u) = \{W^h(x, u), x \in G_h \cup \partial G_h^+\}$ and similarly redefine the vector V^h . For feedback $u(\cdot)$, redefine $R^h(u)$ to have the components $\{r^h(x, y|u(x)), x, y \in G_h \cup \partial G_h^+\}$, where (8.15) holds for $x \in G_h$, and for $x \in \partial G_h^+$ use

$$r^h(x, y|\alpha) = p^h(x, y). \quad (8.22)$$

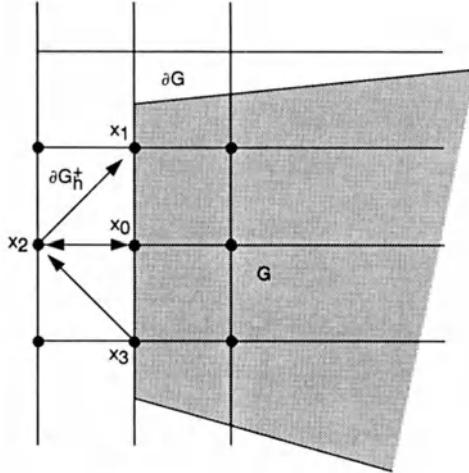


Figure 5.10. Eliminating a reflection state.

Finally, redefine the vector $C^h(u)$ with components $\{C^h(x, u), x \in G_h \cup \partial G_h^+\}$ where

$$C^h(x, u) = k(x, u(x))\Delta t^h(x, u(x)), \quad x \in G_h \quad (8.23)$$

with $C^h(x, u) = c'\Delta z^h(x)$ for $x \in \partial G_h^+$. Then, (8.16) holds for the given control $u(\cdot)$ if the cost $W^h(x, u)$ is well defined and bounded for each initial state, and (8.17) holds if the minimum cost is well defined and bounded for each initial state.

Eliminating the Reflection States. The states in the reflection set ∂G_h^+ can be eliminated from the state space if desired, because their transition probabilities do not depend on the control and their interpolation intervals are zero. But, from the point of view of programming convenience (the automatic and simple generation of the transition probabilities and intervals) it seems to be simpler to keep these states. We will illustrate the method of eliminating these states for one particular case, and this should make the general idea clear. Refer to Figure 5.10. State x_2 is a reflection state. It communicates only with states x_0, x_1 , and it is reached only from states x_0, x_3 . To eliminate state x_2 , we redefine the transition probability from x_0 to x_1 to be $p^h(x_0, x_1|\alpha) + p^h(x_0, x_2|\alpha)p^h(x_2, x_1)$. Analogously, define $p^h(x_3, x_0|\alpha)$ and $p^h(x_3, x_1|\alpha)$. Suppose that states in ∂G_h^+ communicate only to states in G_h . Then, in general, for $x, y \notin \partial G_h^+$, use the redefinitions

$$p^h(x, y|\alpha) \rightarrow p^h(x, y|\alpha) + \sum_{z \in \partial G_h^+} p^h(x, z|\alpha)p^h(z, y).$$

Both Reflecting and Stopping Sets. The dynamic programming equations can easily be derived for the case where part of the boundary is

reflecting, but where there is also an obligatory stopping set with an associated stopping cost $g(\cdot)$. Write ∂G_h^+ for the discretization of the reflecting boundary, as before. Divide the set G_h into disjoint subsets, G_h^0 and ∂G_h^0 , where the latter represents a discretization of the stopping set. Suppose that the reflecting boundary does not communicate with ∂G_h^0 , the absorbing boundary. Redefine the vectors $W^h(u) = \{W^h(x, u), x \in G_h^0 \cup \partial G_h^+\}$ and $V^h(u) = \{V^h(x, u), x \in G_h^0 \cup \partial G_h^+\}$. The components of the vector $C^h(u)$ and matrix $R^h(u)$ are defined by (8.14) and (8.15), respectively, for $x \in G_h^0$, and by (8.22) and below (8.23) for $x \in \partial G_h^+$. Then, if the costs are well defined and finite, (8.16) and (8.17) continue to hold.

6

Computational Methods for Controlled Markov Chains

The chapter presents many of the basic ideas which are in current use for the solution of the dynamic programming equations for the optimal control and value function for the approximating Markov chain models. We concentrate on methods for problems which are of interest over a potentially unbounded time interval. Numerical methods for the ergodic problem will be discussed in Chapter 7, and are simple modifications of the ideas of this chapter. Some approaches to the numerical problem for the finite time problem will be discussed in Chapter 12.

The basic problem and the equations to be solved are stated in Section 6.1. Section 6.2 treats two types of classical methods: the approximation in policy space method and the approximation in value space method. These methods or combinations of them have been used because the early days of stochastic control theory, and their various combinations underlie all the other methods which are to be discussed. The first approach can be viewed as a “descent” method in the space of control policies. The second method calculates an optimal n -step control and value function and then lets n go to infinity. The Jacobi and Gauss-Seidel relaxation (iterative) methods are then discussed. These are fundamental iterative methods which are used with either the approximation in policy space or the approximation in value space approach. When the control problem has a discounted cost, then one can improve the performance of the iterations via use of the bounds given in Section 6.3. The so-called accelerated Gauss-Seidel methods are described in Section 6.4. This modification generally yields much faster convergence, and this is borne out by the numerical data which is presented.

The possible advantages of parallel processing are as interesting for the control problem as for any of the other types of problems for which it has been considered. Although it is a vast topic, we confine our remarks to the discussion of several approaches to domain decomposition in Section 6.5. The size of the state spaces which occur in the approximating Markov chain models can be quite large and one would like to do as much of the computation as possible in a smaller state space. There are several ways to approach this. Section 6.6 deals with a so-called adaptive state aggregation method, where there is an effort to “correct” the current estimate of the solution by estimating a projection of an error on an “aggregated” state

space of much lower dimension. This is an intriguing approach, although at the present level of development for the type of problem which arises from our approximations, it is not as reliable as the multigrid method. Section 6.7 discusses the basic idea of grid refinement, where one first gets a rough solution on a coarser state space, and then continues the computation on the desired “finer” state space, but with a good initial condition obtained from the “coarse” state space solution. Section 6.8 outlines the basic multigrid or variable grid idea, which has been so successful for the numerical solution of many types of partial differential equations. This method can be used in conjunction with all of the methods discussed previously. With appropriate adaptations of the other methods, the result is effective and robust and experimental evidence suggests that it might be the best currently available. There are comments on the numerical properties throughout, and more such comments appear in Chapter 7. In Section 6.9, the computational problem is set up as a linear programming problem. The connections between approximation in policy space and the simplex algorithm are explored and it is shown that the dual equation is simply the dynamic programming equation. Some useful information on computation and some interesting examples are in [13,15,107,120]. An expert system which incorporates some forms of the Markov chain approximation method is discussed in [16,17].

6.1 The Problem Formulation

The cost functionals or optimal cost functionals in Section 5.8 can all be written in the following forms:

$$W^h(x, u) = \sum_y r^h(x, y|u(x))W^h(y, u) + C^h(x, u(x)),$$

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_y r^h(x, y|\alpha)V^h(y) + C^h(x, \alpha) \right],$$

for appropriate $r^h(x, y|\alpha)$ and $C^h(x, \alpha)$. The vector forms are

$$W^h(u) = R^h(u)W^h(u) + C^h(u) \quad (1.1)$$

and

$$V^h = \min_{u(x) \in \mathcal{U}} [R^h(u)V^h(u) + C^h(u)], \quad (1.2)$$

respectively.

This chapter is concerned with a variety of methods which have been found useful for the solution of these equations. Many of the methods come from or are suggested by methods used for the solution of the systems of

linear equations which arise in the discretization of PDE's, as, for example, in [113].

because we will generally be concerned with the solution of equations [such as (1.1) and (1.2)] for fixed values of h , in order to simplify the notation *we will refer to the state space of the Markov chain simply as S and rewrite these equations, with the superscript h deleted, as*

$$W(u) = R(u)W(u) + C(u), \quad (1.3)$$

$$V = \min_{u(x) \in \mathcal{U}} [R(u)V + C(u)], \quad (1.4)$$

where for a feedback control $u(\cdot)$, $R(u) = \{r(x, y|u(x)), x, y \in S\}$ is a stochastic or substochastic matrix. When working with the variable grid methods in Sections 6.7 and 6.8, we will reintroduce the h superscript, so that we can keep track of the actual and relative grid sizes. The following assumptions will be used to assure that the equations (1.3) and (1.4) are meaningful. Our aim is to outline the main computational techniques of current interest, and not to get the best convergence results for iterative methods for controlled Markov chains.

A1.1. *$r(x, y|\alpha), C(x, \alpha)$ are continuous functions of α for each x and y in S .*

A1.2. (i) *There is at least one admissible feedback control $u_0(\cdot)$ such that $R(u_0)$ is a contraction, and the infima of the costs over all admissible controls is bounded from below.* (ii) *$R(u)$ is a contraction for any feedback control $u(\cdot)$ for which the associated cost is bounded.*

A1.3. *If the cost associated with the use of the feedback controls $u^1(\cdot), \dots, u^n(\cdot), \dots$ in sequence, is bounded, then*

$$R(u^1) \cdots R(u^n) \xrightarrow{n} 0.$$

Richardson Extrapolation. Suppose that the solution to the original problem for the diffusion is $V(x)$ and that for the Markov chain approximation it is $V^h(x)$. Suppose that the solutions are related by (for appropriate values of x)

$$V^h(x) = V(x) + V_1(x)h + o(h)$$

for a bounded sequence $V^h(x)$. Then for small enough h , one can get a more accurate solution by using

$$V(x) \approx 2V^{h/2}(x) - V^h(x).$$

While there is no current theory to support this practice in general for all of the problems of interest in this book, it generally yields good results,

if reasonably accurate solutions for small enough parameters h and $h/2$ are available. The scheme is known as Richardson extrapolation [29]. The general problem of selecting the approximation for best accuracy and the appropriate extrapolations needs much further work.

6.2 Classical Iterative Methods: Approximation in Policy and Value Space

In this section, we discuss the two classical methods. The first, called *approximation in policy space* was introduced by Bellman, then extended by Howard [59] to simple Markov chain models, and later extended to quite general models. It is somewhat analogous to a gradient procedure in the space of controls. See the discussion concerning the relations between linear programming and approximation in policy space in Section 6.9. The second method, called *approximation in value space* calculates the optimal cost and control as the limits, as time n goes to infinity, of the optimal costs and controls for the problems that are of interest over a finite time interval $[0, n]$ only. This method can also be interpreted as a fixed point iteration for (1.2).

6.2.1 APPROXIMATION IN POLICY SPACE

Theorem 2.1. *Assume (A1.1) and (A1.2). Then there is a unique solution to (1.4), and it is the infimum of the cost functions over all time independent feedback controls. Let $u_0(\cdot)$ be an admissible feedback control such that the cost $W(u_0)$ is bounded. For $n \geq 1$, define the sequence of feedback controls $u_n(\cdot)$ and costs $W(u_n)$ recursively by (1.3) together with the formula*

$$u_{n+1}(x) = \arg \min_{\alpha \in \mathcal{U}} \left[\sum_y r(x, y | \alpha) W(y, u_n) + C(x, \alpha) \right]. \quad (2.1)$$

Then $W(u_n) \rightarrow V$.

Under the additional condition (A1.3), V is the infimum of the costs over all admissible control sequences.

Proof. To prove the uniqueness of the solution to (1.4), suppose that there are two solutions \bar{V} and \tilde{V} , with minimizing controls denoted by $\bar{u}(\cdot)$ and $\tilde{u}(\cdot)$, resp. Recall that all inequalities between vectors are taken component by component. Then we can write

$$\bar{V} = R^n(\bar{u})\bar{V} + \sum_{i=0}^{n-1} R^i(\bar{u})C(\bar{u}). \quad (2.2)$$

The right hand side sum in (2.2), which is obtained by iterating (1.4), is bounded. Hence, by (A1.2) $R(\bar{u})$ is a contraction, and similarly for $R(\tilde{u})$. Consequently, the right sum of (2.2) converges to the cost function $W(\bar{u}) = \bar{V}$. We can write

$$\begin{aligned}\bar{V} &= \min_{u(x) \in \mathcal{U}} [R(u)\bar{V} + C(u)] \\ &= R(\bar{u})\bar{V} + C(\bar{u}) \leq R(\tilde{u})\bar{V} + C(\tilde{u}),\end{aligned}\tag{2.3a}$$

$$\begin{aligned}\tilde{V} &= \min_{u(x) \in \mathcal{U}} [R(u)\tilde{V} + C(u)] \\ &= R(\tilde{u})\tilde{V} + C(\tilde{u}) \leq R(\bar{u})\tilde{V} + C(\bar{u}).\end{aligned}\tag{2.3b}$$

The equalities and inequalities (2.3) yield

$$R(\bar{u})(\bar{V} - \tilde{V}) \leq \bar{V} - \tilde{V} \leq R(\tilde{u})(\bar{V} - \tilde{V}).$$

Iterating this inequality and using the contraction properties of the $R(\bar{u})$ and $R(\tilde{u})$ implies that $\bar{V} = \tilde{V}$.

By a contraction type of argument such as the one just used, it can be shown that if a solution to (1.4) exists, then it is the infimum of the costs $W(u)$ for feedback controls which do not depend on time. Under (A1.3), it can be shown that any solution to (1.4) is also the infimum over all admissible controls, and we omit the details.

We now prove that $W(u_n) \rightarrow V$. By the definition of $W(u_n)$ in (1.3) and the minimizing operation in (2.1), we have

$$\begin{aligned}W(u_n) &= R(u_n)W(u_n) + C(u_n) \\ &\geq R(u_{n+1})W(u_n) + C(u_{n+1}) \\ &\geq R^m(u_{n+1})W(u_n) + \sum_{i=0}^{m-1} R^i(u_{n+1})C(u_{n+1}).\end{aligned}\tag{2.4}$$

As before, (A1.2) implies that $R(u_{n+1})$ is a contraction. Hence, as $m \rightarrow \infty$, the first term on the far right hand side goes to zero and the sum converges to the cost $W(u_{n+1})$. Thus $W(u_{n+1}) \leq W(u_n)$ and there is a vector \bar{W} such that $W(u_n) \downarrow \bar{W}$. The inequality

$$W(u_n) \geq \min_{u(x) \in \mathcal{U}} [R(u)W(u_n) + C(u)]$$

implies that

$$\bar{W} \geq \min_{u(x) \in \mathcal{U}} [R(u)\bar{W} + C(u)].\tag{2.5}$$

Conversely, the inequality

$$W(u_{n+1}) \leq R(u_{n+1})W(u_n) + C(u_{n+1}) = \min_{u(x) \in \mathcal{U}} [R(u)W(u_n) + C(u)]$$

implies that

$$\bar{W} \leq \min_{u(x) \in \mathcal{U}} [R(u)\bar{W} + C(u)].\tag{2.6}$$

Inequalities (2.5) and (2.6) imply that \bar{W} satisfies (1.4). ■

6.2.2 APPROXIMATION IN VALUE SPACE

The Jacobi Iteration.

Theorem 2.2. Let $u(\cdot)$ be an admissible feedback control such that $R(u)$ is a contraction. Then for any initial vector W_0 , the sequence W_n defined by

$$W_{n+1}(x, u) = \sum_y r(x, y|u(x))W_n(y, u) + C(x, u(x)) \quad (2.7)$$

converges to $W(u)$, the unique solution to (1.3). Assume (A1.1)-(A1.3). Then for any vector V_0 , the sequence recursively defined by

$$V_{n+1} = \min_{u(x) \in \mathcal{U}} [R(u)V_n + C(u)] \quad (2.8)$$

converges to V , the unique solution to (1.4). In detail, (2.8) is

$$V_{n+1}(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_y r(x, y|\alpha)V_n(y) + C(x, \alpha) \right]. \quad (2.9)$$

V_n is the minimal cost for an n -step problem with terminal cost vector V_0 .

Proof. The convergence of $W_n(u)$ is just a consequence of the fact that $R(u)$ is a contraction. It will be shown next that V_n is the minimal n -step cost for the control problem with terminal cost V_0 . Let $u^n(\cdot)$ be minimizing in (2.8) at step n . We can write

$$\begin{aligned} V_1 &= R(u^1)V_0 + C(u^1), \\ V_n &= R(u^n) \cdots R(u^1)V_0 + \sum_{i=1}^n R(u^n) \cdots R(u^{i+1})C(u^i), \end{aligned}$$

which is the n -step cost for the policy which uses $u^i(\cdot)$ when there are still i steps to go and with terminal cost V_0 . In the above expression, the empty product $\prod_{i=1}^0$ is defined to be the identity matrix. The minimizing property in (2.8) yields that for any other admissible feedback control sequence $\{\tilde{u}^n(\cdot)\}$, we have

$$V_{n+1} \leq R(\tilde{u}^{n+1})V_n + C(\tilde{u}^{n+1}).$$

Iterating the last inequality yields

$$V_{n+1} \leq R(\tilde{u}^{n+1}) \cdots R(\tilde{u}^1)V_0 + \sum_{i=1}^{n+1} R(\tilde{u}^{n+1}) \cdots R(\tilde{u}^{i+1})C(\tilde{u}^i)$$

which is the cost for an $n+1$ -step process under the controls $\{\tilde{u}^i(\cdot)\}$ and terminal cost V_0 . Thus, V_n is indeed the asserted minimal n -step cost.

According to Theorem 2.1, there is a unique solution to (1.4). Let $\bar{u}(\cdot)$ be a minimizer in (1.4). Then

$$R(u^{n+1})V_n + C(u^{n+1}) = V_{n+1} \leq R(\bar{u})V_n + C(\bar{u}),$$

$$R(\bar{u})V + C(\bar{u}) = V \leq R(u^{n+1})V + C(u^{n+1}),$$

which implies that

$$R(\bar{u})(V - V_n) \leq (V - V_{n+1}) \leq R(u^{n+1})(V - V_n).$$

By iterating this latter set of inequalities,

$$R^{n+1}(\bar{u})(V - V_0) \leq (V - V_{n+1}) \leq R(u^{n+1}) \cdots R(u^1)(V - V_0).$$

The boundedness of $\{V_n\}$ follows by the contraction property of $R(\bar{u})$. Thus, by (A1.3) we have

$$R(u^n) \cdots R(u^1) \xrightarrow{n} 0,$$

and $V_n \xrightarrow{n} V$. ■

The Gauss-Seidel or Successive Approximation Iteration. The iteration in (2.7) or (2.8) calculates all the values of the components of the $(n+1)^{\text{st}}$ iterates $W_{n+1}(u)$ and V_{n+1} , resp., directly from the components of $W_n(u)$ and V_n , resp. The newly calculated values $W_{n+1}(x, u)$ and $V_{n+1}(x)$ are not used until they are available for all $x \in S$. If the computation uses successive substitutions, in the sense that each newly calculated term $W_{n+1}(x, u)$ or $V_{n+1}(x)$ is immediately substituted for the $W_n(x)$ or $V_n(x)$, resp., in the formulas (2.7) or (2.8), the procedure is then referred to as the *Gauss-Seidel procedure*. In particular, let us order the states in the state space S in some way, and let the inequality sign $<$ denote the ordering. Then we have the following algorithm.

Theorem 2.3. *Let $u(\cdot)$ be an admissible feedback control for which $R(u)$ is a contraction. For any given W_0 , define W_n recursively by*

$$W_{n+1}(x, u) = \sum_{y < x} r(x, y|u(x))W_{n+1}(y, u) \quad (2.10)$$

$$+ \sum_{y \geq x} r(x, y|u(x))W_n(y, u) + C(x, u(x)).$$

Then W_n converges to the unique solution to (1.3).

Assume (A1.1)-(A1.3). For any V_0 , define V_n recursively by

$$V_{n+1}(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_{y < x} r(x, y|\alpha)V_{n+1}(y) + \sum_{y \geq x} r(x, y|\alpha)V_n(y) + C(x, \alpha) \right]. \quad (2.11)$$

Then V_n converges to the unique solution of (1.4).

Remark. It is useful to note that there are $\tilde{r}(x, y|u)$ and $\tilde{C}(x, u)$ such that (2.10) can be written in the form

$$W_{n+1}(x, u) = \sum_y \tilde{r}(x, y|u) W_n(y, u) + \tilde{C}(x, u). \quad (2.12)$$

Let $x(1)$ denote the lowest state in the ordering. Then the terms $\tilde{r}(x, y|u)$ and $\tilde{C}(x, u)$ in (2.12) are defined in the following recursive way:

$$\tilde{r}(x(1), y|u) = r(x(1), y|u(x)),$$

$$\tilde{r}(x, y|u) = r(x, y|u(x)) + \sum_{z < x} r(x, z|u(x)) \tilde{r}(z, y|u), \quad y \geq x \geq x(1), \quad (2.13)$$

$$\tilde{r}(x, y|u) = \sum_{z < x} r(x, z|u(x)) \tilde{r}(z, y|u), \quad x > y \geq x(1),$$

$$\tilde{C}(x, u) = \sum_{y < x} r(x, y|u(x)) \tilde{C}(y, u) + C(x, u(x)). \quad (2.14)$$

The proofs of the theorem and of the representation (2.12)-(2.14) are in [74] or [84], which also contain discussions of the numerical properties relative to the Jacobi procedure, and the second reference is the first place where the stochastic applications appeared. The Gauss-Seidel method of Theorem 2.3 is never inferior to the Jacobi method of Theorem 2.2, and it requires less storage space. Some numerical properties and comparisons will be presented below. Combinations of the Gauss-Seidel and Jacobi procedures are useful in implementations on parallel processors, as noted in Section 6.5 below. Iterations such as (2.7)-(2.11) are sometimes referred to as *relaxations* and this term will be used interchangeably with the term iteration.

6.2.3 COMBINED APPROXIMATION IN POLICY SPACE AND APPROXIMATION IN VALUE SPACE

The approximation in policy space, as stated in Theorem 2.1, requires the solution to (1.3) for each successive control $u_n(\cdot)$. Because the state spaces S tend to have many points (perhaps in the tens of thousands or even larger in some cases), obtaining a good approximation to the solution to (1.3) for each control $u_n(\cdot)$ might be an onerous task. On the other hand, the approximation in value space method of Theorem 2.2 or 2.3 might require much calculation at each step in order to get the minima. Also, the “backwards iteration” in value space method does not allow the use of multigrid or aggregation type methods, which have been found to be quite powerful in many cases.

Typical computational methods use the approximation in policy space as a basis, but get only an approximation to the values $W(u_n)$ at each step. For example, let $u_n(\cdot)$ be the current candidate for the optimal control, and let \tilde{W}_n be an approximation to $W(u_n)$. Calculate the next candidate $u_{n+1}(\cdot)$ for the optimal control via (2.1) with \tilde{W}_n replacing $W(u_n)$. Then, starting with the initial guess \tilde{W}_n , obtain an approximation \tilde{W}_{n+1} to the solution $W(u_{n+1})$ to

$$W = R(u_{n+1})W + C(u_{n+1}).$$

For example, one might use a sequence of say N Gauss-Seidel relaxations, starting with initial value \tilde{W}_n and define \tilde{W}_{n+1} to be the value obtained at the end of that sequence. Then repeat the procedure until some criterion of convergence is satisfied. The extreme cases $N = 1$ and $N = \infty$ are, of course, the approximation in value and approximation in policy space methods, resp. Such a procedure emphasizes the necessity for examining efficient methods for the solution of systems of linear equations, such as the acceleration, multigrid or aggregation methods introduced in the following sections and appropriate choices among these methods are generally preferred to the use of the unaltered Gauss-Seidel relaxation.

Suppose that one considers the computing time required to reach a given error in the estimate of the V or in the optimal control versus the number N of relaxations used between policy updates. The details of the graph would depend on the difficulty of getting the minimum in (2.1). But in general, as N increases it first drops and ultimately increases. It has been observed in a variety of two dimensional problems where the time required to get the minimum in (2.1) for all $x \in S$ was of the order of two or three times the time required for a single relaxation that the best value varied between five and fifteen, which is not a large range. Generally, some experimentation is required. The procedure which uses several Jacobi relaxations between each policy update is known as the *modified policy iteration algorithm* and convergence proofs for the discounted problem are in [99].

6.2.4 THE GAUSS-SEIDEL METHOD: PREFERRED ORDERINGS OF THE STATES

The Jacobi procedure of Theorem 2.1 does not depend on the way the states are ordered. However, the performance of the Gauss-Seidel procedure does depend on the ordering. If the diffusion effects dominate those of the drift terms in the original process from which the Markov chain approximation was derived and the diffusion is non-degenerate, then the orderings in successive iterations might be alternated in a fashion such as illustrated in Figures 6.1a,b. If the diffusion is degenerate and there is a dominant flow direction due to the effects of the drift term, then the performance is improved if these effects are accounted for in the ordering. Only

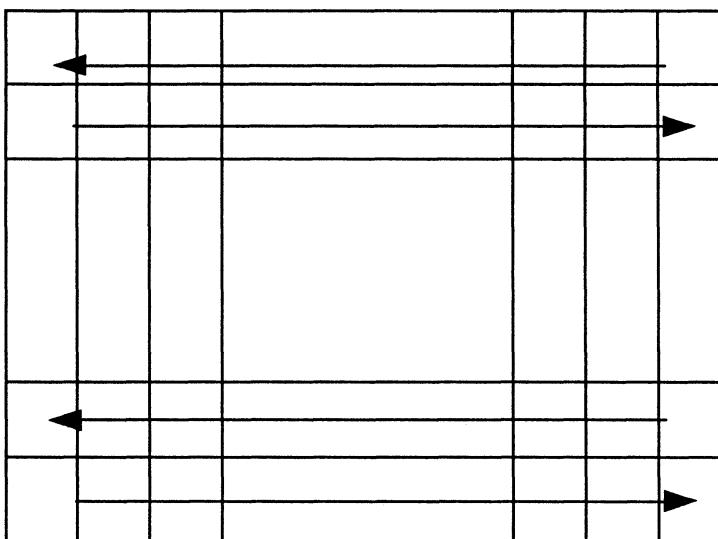


Figure 6.1a. Directions of iteration.

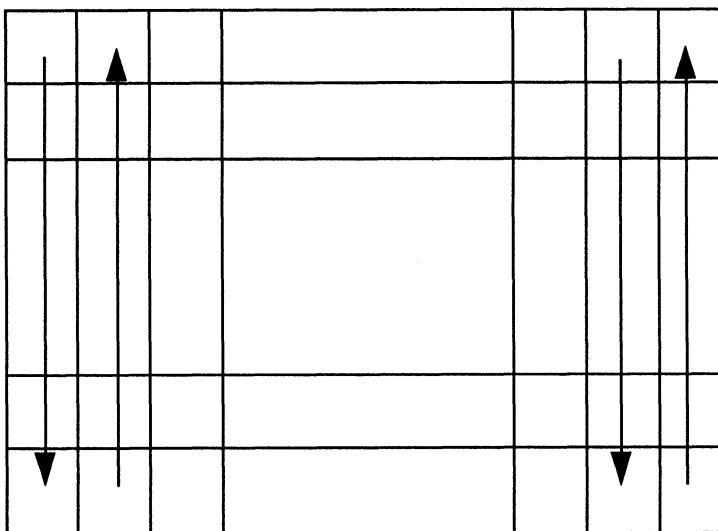


Figure 6.1b. Directions of iteration.

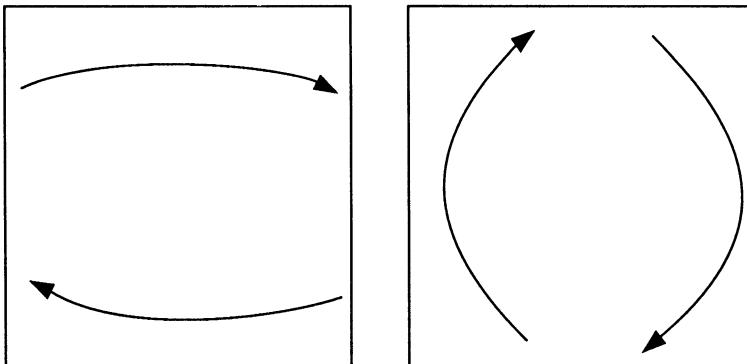


Figure 6.2a,b. Directions of iteration.

a few comments will be made in order to give the general idea. Consider the special case

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= b_2(x, u(x))dt + \sigma dw, \end{aligned}$$

and refer to Figures 6.2a,b. For $x_2 > 0$, x_1 increases, and conversely for $x_2 < 0$. Thus the general mean flow is qualitatively as indicated in Figure 6.2a.

It should be noted that in the early stages of either the approximation in policy space or the approximation in value space method, the actual control which appears can vary quite a bit from iteration to iteration. It is therefore preferable to order the states in a way that does not depend on the control.

The following rule seems to be useful, where applicable: *Order the states to “tighten” the connection with the absorbing states, going against the “tendency of the flow” where possible.* For the above example, if the outer boundary is absorbing, then a reasonable ordering is shown in Figure 6.2c. Here we ignored the component of flow in the “vertical” direction. If the general tendency [due to the effects of the $b_2(x, u(x))$ term] were as in Figure 6.2b, then one might alternate between the orderings in Figures 6.2c and 6.2d.

More detail concerning the mathematics of the role of the orderings in the convergence is in [74]. Some intuitive justification for the recommended orderings can be seen from the degenerate example, where $\sigma = 0$. Suppose that the outer boundary is absorbing. If it is possible to order the states in such a way that the iteration is against the flow, then convergence will be obtained in one sweep. This is equivalent to the approach to solving the Hamilton-Jacobi equation by integrating backwards along the characteristic curves. Following this analogy, we can see why the method works well when there is an absorbing boundary and when it is possible to order the

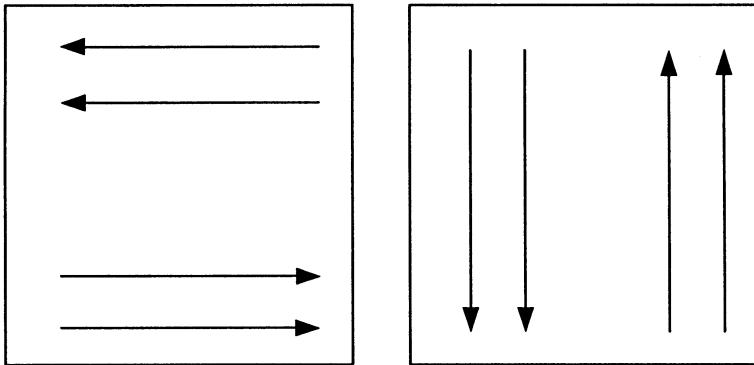


Figure 6.2c,d. Directions of iteration.

states such that the iteration is against the tendency of the flow and moves backwards from the absorbing set.

6.3 Error Bounds for Discounted Problems

6.3.1 THE JACOBI ITERATION

When the problem of interest has a discounted cost function, and the Jacobi or Gauss-Seidel procedure is used to solve (1.3) or (1.4), one can obtain additional estimates of the actual solutions to these equations. These estimates can be used to improve the values given by the current iterates. First, the Jacobi case will be dealt with, and the control will be fixed (as will be the case when we need to estimate $W(u_n)$ for a cycle of the iteration in policy space method).

Suppose that (1.3) can be written in the form

$$W(u) = \gamma R_0(u)W(u) + C(u), \quad (3.1)$$

where $R_0(u)$ is a stochastic or substochastic matrix and $0 < \gamma < 1$. For example, if (1.3) or (1.4) arose from the discounted problem (5.8.10), then (reintroducing the h superscript for the moment), $r^h(x, y|u(x)) = e^{-\beta \Delta t^h(x, u(x))} p^h(x, y|u(x))$. Define γ by

$$\gamma = \max_x e^{-\beta \Delta t^h(x, u(x))}.$$

Then (dropping the h again), (5.8.10) can be written as (3.1) with an appropriate matrix $R_0(u)$, whose row sums are no greater than one.

Let the vector W_n denote the n^{th} estimate of the solution of (3.1). Define the “Jacobi” relaxation $T_{J,u}(X, C)$ by $T_{J,u}(X, C) = \gamma R_0(u)X + C(u)$. Then $T_{J,u}(W_n, C(u))$ is an “updated” estimate of $W(u)$. Next, some estimates of $W(u)$ will be stated and they will allow us to improve this latest estimate

$T_{J,u}(W_n, C(u))$. Define

$$\begin{aligned}\delta_{\max} &= \max_x [T_{J,u}(W_n, C(u))(x) - W_n(x)], \\ \delta_{\min} &= \min_x [T_{J,u}(W_n, C(u))(x) - W_n(x)].\end{aligned}$$

We can now state [10, pp. 190–192], the following theorem.

Theorem 3.1. *Under the above conditions, we have*

$$W(x, u) \leq T_{J,u}(W_n, C(u))(x) + \frac{\gamma}{1-\gamma} \delta_{\max} \leq W_n(x) + \frac{\delta_{\max}}{1-\gamma}, \quad (3.2)$$

$$W(x, u) \geq T_{J,u}(W_n, C(u))(x) + \frac{\gamma}{1-\gamma} \delta_{\min} \geq W_n(x) + \frac{\delta_{\min}}{1-\gamma}. \quad (3.3)$$

Applying Theorem 3.1 to the Improvement of the Current Trial Solution W_n to (1.3). To improve the current estimate W_n of the solution to (3.1), choose any vector W_{n+1} satisfying

$$T_{J,u}(W_n, C(u))(x) + \frac{\gamma}{1-\gamma} \delta_{\min} \leq W_{n+1}(x) \quad (3.4)$$

$$\leq T_{J,u}(W_n, C(u))(x) + \frac{\gamma}{1-\gamma} \delta_{\max}.$$

For example, for each x choose the value which is at the midpoint of the range in (3.4).

6.3.2 THE GAUSS-SEIDEL PROCEDURE

The estimates of Theorem 3.1 can also be used to improve the performance of the Gauss-Seidel iteration for (3.1). To do this, we need to represent the Gauss-Seidel iteration for (3.1) in a form analogous to (2.12) for some appropriate matrix $\tilde{R}(u) = \{\tilde{r}_0(x, y|u(x)), x, y \in S\}$. Given the n^{th} estimate W_n of the solution to (3.1), the Gauss-Seidel procedure updates the estimate via the iteration

$$T_{GS,u}(W_n, C(u))(x) = \quad (3.5)$$

$$\gamma \left[\sum_{y < x} r_0(x, y|u(x)) W_n(y) + \sum_{y \geq x} r_0(x, y|u(x)) W_n(y) \right] + C(x, u(x)).$$

Referring to (2.13) and (2.14), rewrite the iteration (3.5) in the form

$$T_{GS,u}(W_n, C(u))(x) = \gamma \sum_y \tilde{r}_0(x, y|u) W_n(y) + \tilde{C}(x, u), \quad (3.6)$$

where we sequentially (in x) define

$$\begin{aligned}\tilde{r}_0(x, y|u) &= r_0(x, y|u(x)) + \gamma \sum_{z < x} r_0(x, z|u(x))\tilde{r}_0(z, y|u), \quad y \geq x, \\ &= \gamma \sum_{z < x} r_0(x, z|u(x))\tilde{r}_0(z, y|u), \quad x > y,\end{aligned}\tag{3.7}$$

$$\tilde{C}(x, u) = C(x, u(x)) + \gamma \sum_{y < x} r_0(x, y|u(x))\tilde{C}(y, u).\tag{3.8}$$

Now write (3.6) in the vector form

$$T_{GS,u}(W_n, C(u)) = \gamma \tilde{R}_0(u)W_n + \tilde{C}(u).\tag{3.9}$$

With this representation, Theorem 3.1 can be applied directly, by replacing the $T_{J,u}$ by $T_{GS,u}$. The $\tilde{R}_0(u)$ or $\tilde{C}(u)$ need not be explicitly computed, only (3.5) and the Gauss-Seidel equivalents of δ_{\max} and δ_{\min} need be computed.

A very similar improvement procedure can be given for the iteration in value space forms (2.9) and (2.11). See [10, pp. 190–192], for the formulas.

6.4 Accelerated Jacobi and Gauss-Seidel Methods

6.4.1 THE ACCELERATED AND WEIGHTED ALGORITHMS

Consider the linear problem (1.3), where the matrix $R(u)$ is a contraction. Again, let us order the states in some way, so that the Gauss-Seidel method is well defined. Let Ω denote the diagonal matrix with the given entries $\omega(x) > 0$, $x \in S$. Given W_0 , an initial estimate of the solution to (1.3), let W_n denote the sequence of estimates of the solution which is obtained by the following recursion

$$\begin{aligned}\tilde{W}_n &= R(u)W_n + C(u) \\ W_{n+1} &= \Omega \tilde{W}_n + (I - \Omega)W_n.\end{aligned}\tag{4.1}$$

If $\omega(x) > 1$ for all x , then this procedure is known as the *accelerated Jacobi* (AJ) method, where the $\omega(x)$ are the acceleration parameters. If $\omega(x) < 1$, then it is called the *weighted Jacobi procedure* [14].

There are two forms of the accelerated Gauss-Seidel method, depending on whether the “acceleration” is done at the end of a Gauss-Seidel sweep or continuously during the sweep. Let $\omega(x) > 1$. The first, which we call the semi-accelerated Gauss-Seidel method (SAGS) is defined by the recursion

$$\tilde{W}_n(x) = \left[\sum_{y < x} r(x, y|u(x))\tilde{W}_n(y) + \sum_{y \geq x} r(x, y|u(x))W_n(y) + C(x, u(x)), \right]$$

$$W_{n+1} = \Omega \tilde{W}_n + (I - \Omega) W_n. \quad (4.2)$$

The full *accelerated Gauss-Seidel method* (AGS) is defined by

$$\begin{aligned} W_{n+1}(x) &= \omega(x) \left[\sum_{y < x} r(x, y|u(x)) W_{n+1}(y) \right. \\ &\quad \left. + \sum_{y \geq x} r(x, y|u(x)) W_n(y) + C(x, u(x)) \right] + (1 - \omega(x)) W_n(x). \end{aligned} \quad (4.3)$$

In applications, one generally lets $\omega(x) = \omega \geq 1$, a constant, and we will do this henceforth. The algorithm (4.3) is also referred to as a *successive overrelaxation method* (SOR) [52, Section 10.1.4].

For appropriate values of the acceleration parameter ω , the accelerated Gauss-Seidel methods are preferable to the original Gauss-Seidel method (where $\omega = 1$). For the case (4.1), the acceleration is equivalent to a shift in the eigenvalues of the matrix $R(u)$. To see this most clearly, let us rewrite (4.1) in vector form: Define the matrix $\Omega = \omega I$, where I is the identity matrix, write the iteration (4.1), and define the matrix $R_\omega(u)$ by

$$W_{n+1} = [\omega R(u) + (1 - \omega)I] W_n + \omega C(u) = R_\omega(u) W_n + \omega C(u). \quad (4.4)$$

There is a similar expression for (4.2), where $R(u)$ is replaced by the $\tilde{R}(u)$ defined in (2.1.3). If λ is an eigenvalue of $R(u)$, then $\omega\lambda + (1 - \omega)$ is an eigenvalue of $R_\omega(u)$. The idea is to choose a value of ω such that the spectral radius of $R_\omega(u)$ is less than that of $R(u)$. The accelerated Jacobi procedure is not generally useful, but the accelerated Gauss-Seidel is. Examples will be given below. It will be seen below that the AGS is actually the best procedure among these choices, in the sense that it gives the smallest spectral radius with appropriate choices of the parameter ω .

Generally some experimentation is needed to get appropriate values of the factor ω . This is often easier than it might seem at first sight. For the approximation in policy space method, one solves a sequence of problems of the type (1.3). For many problems, it has been experimentally observed that the range of possible ω (those for which the algorithm is stable) as well as the optimal values of ω do not change much from one control to another. This point will also be pursued further in the example of Subsection 6.4.3. Also, in doing numerical optimization, one frequently wants to solve a family of optimization problems with closely related cost functions or system dynamics, in order to get controls which are reasonably robust with respect to possible variations in these quantities. Experimentation with the value of ω in the early stages of such as procedure generally yields useful values for the rest of the procedure. Accelerated procedures have also been used with the nonlinear iteration in value space algorithm (1.4). See [84], which first introduced these methods for the computation of optimal controls via (1.4).

6.4.2 NUMERICAL COMPARISONS BETWEEN THE BASIC AND ACCELERATED PROCEDURES

General Comments and Comparisons. Some of the salient features of and relations between the Jacobi and Gauss-Seidel procedures and their accelerated versions can be seen by means of a simple example. The Markov chain in this example is a locally consistent approximation of the one dimensional system

$$dx = cxdt + dw, \quad c > 0.$$

The set G is the unit interval $[0, 1]$, and the endpoints 0 and 1 are instantaneously reflecting. $1/h$ is assumed to be an integer. Let the state space of the approximating chain, which includes the reflecting endpoints, be $\{0, h, \dots, 1-h, 1\}$. Because the endpoints are instantaneously reflecting, they communicate with their nearest neighbors with probability one and “zero delay.” Thus, they can be eliminated from the problem. We will do this and denote the reduced state space by $S = \{h, \dots, 1-h\}$. Let $\lambda \in (0, 1)$ denote a discount factor and P the matrix of transition probabilities, where $p(x, x \pm h) = (1 \pm hc)/2$, except for the transition of the endpoints h and $1-h$ to themselves. Let $hc \leq 1$. For some given vector C , we wish to solve the equation

$$W = \lambda PW + C. \quad (4.5)$$

Recalling the discussion in Section 5.2, we write (4.5) in the “normalized” form

$$W(x) = \lambda \sum_{y \neq x} \frac{p(x, y)}{1 - \lambda p(x, x)} W(y) + \frac{C(x)}{1 - \lambda p(x, x)}, \quad (4.6)$$

which we rewrite again with the obvious definitions of P_N and C_N as

$$W = \lambda P_N W + C_N.$$

In this example, $p(x, x) = 0$ unless x equals h or $1-h$.

The Gauss-Seidel and Jacobi methods and their variations can be used with either the normalized equation (4.6) or the unnormalized equation (4.5). The various transition matrices are tabulated below for the case of $c = 0$, $h = 1/5$, $\lambda = .995$. Thus, in Tables 4.1 to 4.4 the state space has four points $\{.2, .4, .6, .8\}$. For the Gauss-Seidel case, the sweep is from left to right. The general comparisons below hold true for any value of c for which the $p(x, y)$ are non-negative. The subscript N denotes that the normalized transition probabilities are used.

$$\begin{array}{cccc} .4975 & .4975 & 0 \\ .4975 & 0 & .4975 & 0 \\ 0 & .4975 & 0 & .4975 \\ 0 & 0 & .4975 & .4975 \end{array}$$

Table 4.1. The matrix λP .

0	.9900	0	0
.4975	0	.4975	0
0	.4975	0	.4975
0	0	.9900	0

Table 4.2. The matrix λP_N .

.4975	.4975	0	0
.2475	.2475	.4975	0
.1231	.1231	.2475	.4975
.0613	.0613	.1231	.7450

Table 4.3. The matrix for the Gauss-Seidel iteration using λP .

0	.9900	0	0
0	.4925	.4975	0
0	.2450	.2475	.4975
0	.2426	.2450	.4925

Table 4.4. The matrix for the Gauss-Seidel iteration using λP_N .

Because the case at hand, where the states are ordered on the line, is so special, we need to be careful concerning any generalizations that might be made. However certain general features will be apparent. With the Gauss-Seidel forms, the state transitions are not just to nearest neighbors, but to all the points to which the neighbors are connected, in the direction from which the sweep comes.

The eigenvalues of the matrix λP are in the interval $[-.995, .995]$, those of λP_N are in $[-.9933, .9933]$, those of the matrix in Table 4.3 in $[0, .9920]$, and those of the matrix in Table 4.4 in the interval $[0, .9867]$. (They are all real valued.) The eigenvalues of the matrices used for the Jacobi iterations, λP and λP_N , are real and symmetric about the origin in this case. The “shift of eigenvalues” argument made in connection with (4.4) above or (4.7) below, implies that the use of (4.1) is not an improvement over the Jacobi method if $\omega > 1$. The “one sided” distribution of the eigenvalues for the Gauss-Seidel case suggests that the methods (4.2) and (4.3) might yield significant improvement over the basic methods.

To see this point more clearly, consider the same example as above, but with the smaller value of the difference interval $h = 1/20$. The results are tabulated in the tables below, where the subscript N denotes that the normalized matrix P_N is used as the basis of the calculation of the matrix for the Gauss-Seidel or accelerated cases, as appropriate. For this case, the spectral radius of the Jacobi matrix λP is 0.9950, that for the normalized Jacobi matrix λP_N is 0.9947, and that for the normalized Gauss-Seidel case

is 0.9895. For the accelerated procedure, we have the following spectral radii for the listed cases

ω	1.7	1.5
SAGS _N	.9821	.9842
AGS	.9650	.9766
AGS _N	.9455	.9699

Table 4.5. Spectral radii for the accelerated procedures.

The best procedure is the full AGS_N, although the acceleration obtained here with AGS_N is greater than one would normally get for problems in higher dimensions. Data for a two dimensional case in [83] support this preference for AGS in the case of the nonlinear iteration in value space algorithm also.

Comments on the Choice of the Weight ω . The AGS method is particularly hard to analyze because it is nonlinear in the parameter ω . To get some idea of a bound on the range of the parameter, consider the AJ or SAGS, where the matrix in the iteration can be written as $\omega\tilde{R} + (1 - \omega)I$, for some stochastic or substochastic matrix \tilde{R} . Suppose that the spectrum of \tilde{R} is on the interval $[-a, b]$, where a and b are non-negative. Then, by the “eigenvalue shift” argument, the best weight would equalize the absolute value of the shifted values of b and $-a$. This yields

$$\omega b + (1 - \omega) = \omega a - (1 - \omega),$$

which yields

$$\omega = \frac{2}{2 - (b - a)}. \quad (4.7)$$

Thus for larger $(b - a)$, use a larger value of ω .

6.4.3 EXAMPLE

Suppose that the approximation in policy space algorithm of Theorem 2.1 is used to solve (1.4). Then a sequence of solutions $u_n(\cdot)$ is generated, and one wants to get an estimate of the solution to (1.3) for each such control. We comment next on experimental observations of the sequence of optimal acceleration parameters. The Markov chain is obtained as an approximation to the following system where $x = (x_1, x_2)$ and $|u| \leq 1$:

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= (b_2(x) + u)dt + dw. \end{aligned} \quad (4.8)$$

The state space is defined to be the square $G = \{x : |x_i| \leq B, i = 1, 2\}$ centered about the origin. We let the outer boundary of the square be

reflecting and stop the process at the first time τ that the strip $\{x : |x_1| \leq \delta < B\}$ is hit. Let the cost be $W(x, u) = E_x^u \int_0^\tau [1 + |u(s)|] ds$. To start the approximation in policy space method, suppose that the initial control $u_0(\cdot)$ is identically zero, get an approximation to the cost $W(u_0)$, and continue as in Theorem 2.1.

Suppose that we use the AGS method to get approximate solutions to the linear equations (1.3) for the costs $W(u_n)$. The range of acceleration parameters for which the algorithm is stable and the best acceleration parameter were observed to decrease slightly as n increases, but not by much. For $n = 0$, the best weight is slightly greater than 1.3. This would be unstable for the optimal control, for which the best parameter is about 1.25. Even when accelerated, the convergence of the AGS methods for (1.3) was observed to be slower for the initial control than for the optimal control. Let $\bar{u}(\cdot)$ denote the optimal control. Then to solve (1.3) to within a given precision takes more than five times longer with $R(u_0)$ than with $R(\bar{u})$. These observations hold true for various functions $b_2(\cdot)$. The conclusions are problem-dependent, but analogous patterns appear frequently.

For this example, the process is absorbed faster for the optimal control than for the initial members of the sequence of controls generated by the policy iteration method. Intuitively, faster absorption means a smaller spectral radius. For other cases, the best acceleration parameter increases slightly as n increases.

The method of Section 6.7 first gets a good approximation to the solution of the optimization problem on a coarse grid (state space), and then uses that approximation as the initial condition for the solution on a finer grid. We observe experimentally for problems such as the one above that the optimal acceleration parameter (for the optimal controls) increases slightly as the grid is refined. This seems to be another consequence of the relative speed of absorption or killing of the process, which is faster on the coarser grid. Thus, good values of the weight for the coarse grid would provide a good start for determining a good weight for the finer grid.

6.5 Domain Decomposition and Implementation on Parallel Processors

Exploiting the possibilities of parallel processing is obviously important for any numerical method. We mention only three standard schemes of the domain decomposition type, in order to show some of the the possibilities. Such methods are interesting and undoubtably useful. The reader should keep in mind that the main computational problems for the optimal control problem concern either high dimensional problems (dimension greater than three) for which little has been done to date, problems with complicated boundary conditions, or problems where great accuracy is required.

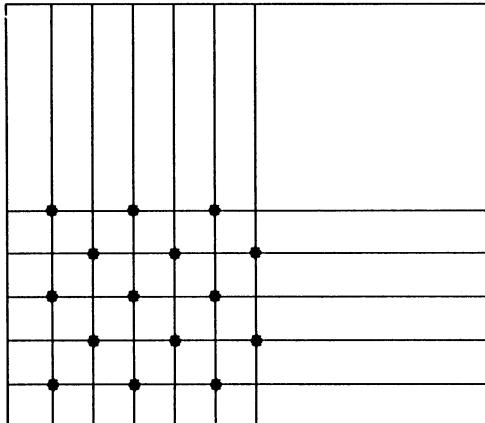


Figure 6.3. Red-black Gauss-Seidel.

Parallelization methods for the multigrid method of Section 6.8 are a topic of active research, and the reader can find useful ideas in the proceedings of many conferences devoted to the subject [94].

In principle, the computations (2.7) for the Jacobi method can be done *simultaneously* for each of the states, because the values of the $(n+1)^{st}$ iterate $W_{n+1}(x, u)$ for all $x \in S$ are computed directly from the values of the n^{th} iterate $W_n(y, u)$, $y \in S$. On the other hand, the computation of $W_{n+1}(x, u)$ in (2.10) for each state x depends on the values of the $(n+1)^{st}$ iterate $W_{n+1}(y, u)$ for the states $y < x$. Thus a simultaneous calculation of all $W_{n+1}(x, u)$ via a parallel processor is not possible. Many intermediate variations have been designed with the aim of preserving some of the advantages of the Gauss-Seidel method, but allowing the use of some parallel processing. One can, in fact, mix the Gauss-Seidel and Jacobi procedures in rather arbitrary ways, with the state space being divided into disjoint groups of points, with a Gauss-Seidel procedure used within the groups and the groups connected via the Jacobi procedure. Such is the case with the three examples below. One well known scheme, called the *red-black Gauss-Seidel method* (for obvious reasons!) will be described next [14]. In fact, with an appropriate ordering of the states, this method is actually a Gauss-Seidel method.

Red-Black Gauss-Seidel. Refer to Figure 6.3 where the square state space for the two dimensional problem of interest is defined. For simplicity, suppose that the boundary is absorbing and let the state space consist of the points of intersection of the grid lines in the figure. The dark points are to be called “black,” and the others (excluding the boundary) “red.” Suppose that the transitions of the Markov chain from a black point are either to the red points or to the boundary (but not to other black points), and that the transitions from the red points are either to the black points

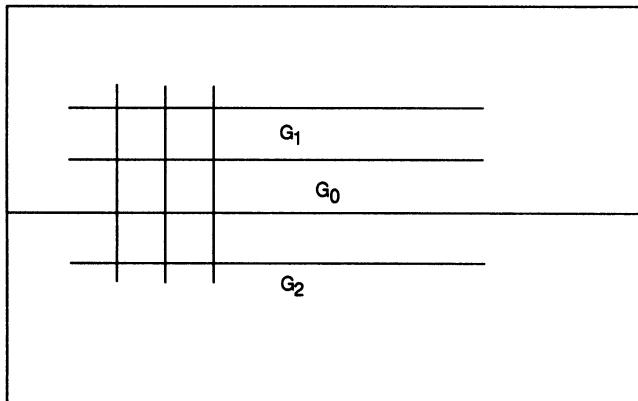


Figure 6.4. Domain decomposition (1).

or to the boundary (but not to other red points). This would be the case if the transition probabilities were obtained by the finite difference method of Sections 5.1-5.3 applied to a diffusion (no jump term) model.

Suppose that we wish to solve (1.3). The procedure for each iteration is as follows: Let $W_n(u)$ be the n^{th} estimate of the solution to (1.3). First iterate on all of the black points, using a Jacobi relaxation (the Gauss-Seidel procedure would give the same result, because the black states do not communicate to other black states), obtaining $W_{n+1}(x, u)$ at the black points x . Then using these newly computed values, iterate on all the red points with a Jacobi procedure to get $W_{n+1}(x, u)$ for the red points x . The procedure has divided the state space into two groups, each group having roughly half of the states, and the computation within each group can be implemented in a parallel fashion.

Domain Decomposition (1). Refer to Figure 6.4, where the domain G is divided into the disjoint sets G_1, G_2 and the connecting line G_0 .

Let the state space S be the intersection of the lines in the figure and suppose that under the given transition probability, the states communicate only to the nearest neighbors. Thus, the points on G_0 communicate with points in both G_1 and G_2 . The points in G_1 (resp., G_2) communicate only with points in G_0 and G_1 (resp., G^0 and G_2). A standard decomposition technique updates $W_n(u)$ via a Gauss-Seidel or Jacobi method (or accelerated method) in each domain G_1 and G_2 separately and simultaneously. It then uses either a Jacobi or a Gauss-Seidel procedure to update the values at the points on the connecting set G_0 .

In principle, the domain can be subdivided into as many sections as desired, at the expense of increasing overhead.

Domain Decomposition (2). Refer to Figure 6.5, where the state space

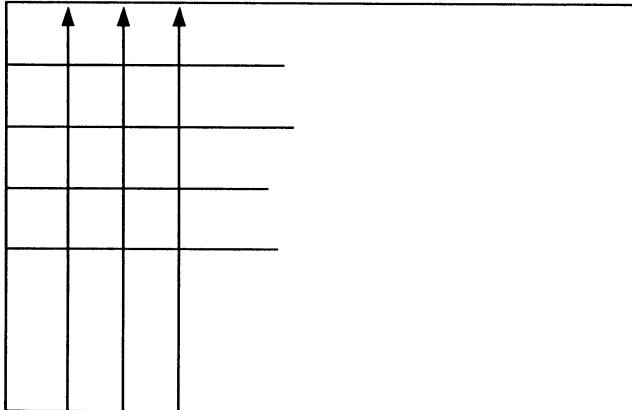


Figure 6.5. Domain decomposition (2).

is as in the above two examples, and the states communicate only to their nearest neighbors as well. A separate processor is assigned to update the states on each of the “columns” in the figure. Proceed as follows. Let the n^{th} estimate $W_n(u)$ be given as above. The memory of the i^{th} processor contains the values of $W_n(x, u)$ for states x which are in the $(i - 1)^{st}$, i^{th} , and $(i + 1)^{st}$ column. It then obtains $W_{n+1}(x, u)$ for the states x in the i^{th} column by a combined Jacobi and Gauss-Seidel type procedure which uses $W_n(y, u)$ for y in the $(i - 1)^{st}$ and $(i + 1)^{st}$ column, and a successive substitution for y in the i^{th} column. After this computation is done for all columns the new values for each column are then transferred to the processors for the neighboring columns.

For a system such as (4.8) where the “diffusion” or “noise” implies a strong connection in the x_2 direction between the states, the separation of the state space into vertical lines (the x_2 direction) would seem to be preferable to a separation into horizontal lines, although we do not have supporting data for this conjecture.

6.6 A State Aggregation Method

The possibility of doing part of the computations for the solution of (1.3) or (1.4) in a state space of much reduced size is appealing. One approach (Section 6.7 below) uses a coarse grid or state space to get an initial estimate of the solution. But even this method requires that a system of linear equations (although of reduced size) be solved, at least approximately, so that there is still the general question of efficient methods of solution of such equations. Very useful alternative methods are forms of the variable or multigrid method of Section 6.8. In this section, we discuss an method the so-called *adaptive state aggregation method* which has been found to

be quite useful on problems where the state space can be “nearly” decomposed. Although such decompositions are not common for our type of problem, the method is intriguing. The essential idea is the adaptive lumping together of states into a small number of groups, and then “adjusting” the current trial solution via a side calculation on the reduced state space. These aggregation steps are alternated with a sequence of relaxations (say, of the Gauss-Seidel or accelerated Gauss-Seidel type) for the solution of (1.3). The method is said to be adaptive because the aggregation groupings which are used in each aggregation step depend on the values of certain “residuals.” We will follow the development in [11] for the solution of (1.3), but with a change in notation. The reader is referred to the reference for further information. The reference deals with both the discounted and the average cost per unit time problem. For the latter case, see Chapter 7. In this section, we will treat the case where $R(u)$ is a contraction. The method is actually an adaptation of a projection method for improving the estimates of solutions to general linear equations when the matrix is invertible [18].

The actual computational algorithm consists of one or more aggregation steps used between a sequence of either Gauss-Seidel or Jacobi steps, accelerated or not. In the form in which the algorithm is stated, there is no proof of convergence. The convergence proof in [11] requires an increasing number of Gauss-Seidel type relaxations between the successive aggregation steps, although the method performed well on the examples presented in the reference, where only a fixed finite number of intermediate (Jacobi) relaxations were used.

6.6.1 OUTLINE OF THE AGGREGATION METHOD

Let $u(\cdot)$ be a control such that $R(u)$ is a contraction, and let $d = |S|$ denote the number of points in the state space S . Define the “Jacobi” relaxation operator $T_{J,u}$ by $T_{J,u}(X, C(u)) = R(u)X + C(u)$. Let W_n denote the n^{th} estimate of the solution to (1.3). For some integer $q \geq 1$, let $\{S_j, j \leq q\}$ be a partition of the state space S . The integer q is the number of aggregation groups, and the S_j are the aggregation groups. Generally, q is fairly small and we have not used values greater than eight, which was usually several orders of magnitude smaller than the size of the original state space S . Define the “membership” matrix (with d rows and q columns) $H = \{h(x, j), x \in S, j \leq q\}$, where $h(x, j) = 1$ if the state x is in the set S_j and is zero otherwise.

Let $W(u)$ denote the solution to (1.3). The error in the current estimate can be represented in the form: There is a vector z and a vector v which is orthogonal to the columns of H such that

$$W(u) - W_n = Hz + v. \quad (6.1)$$

The basic idea of the aggregation method is to get a “good” aggregation

matrix H , then estimate the value of z , and then use either $W_n + Hz$ or $T_{J,u}(W_n + Hz, C(u))$ as the next estimate of the solution to (1.3). The Gauss-Seidel relaxation can be used in lieu of $T_{J,u}$.

There are two immediate problems in working with (6.1), even if H is given. First, the solution $W(u)$ is not known, and second, we need a computationally convenient way of getting a good approximation to z . The entire procedure, as it will be described now, is essentially heuristic. But it was useful on many of the problems on which it has been tried.

On Estimating z . To get an approximation to z , start by eliminating the unknown $W(u)$ by writing $W(u) - W_n$ in terms of $W_n + T_{J,u}(W_n, C(u))$. From the expressions

$$W_n - T_{J,u}(W_n, C(u)) = W_n - R(u)W_n - C(u),$$

$$W(u) - R(u)W(u) - C(u) = 0,$$

one gets

$$(I - R(u))(W_n - W(u)) = W_n - T_{J,u}(W_n, C(u)).$$

Using (6.1) in this expression yields

$$(I - R(u))Hz + (I - R(u))v = T_{J,u}(W_n, C(u)) - W_n.$$

Premultiplying the last equation by H' gives

$$H'(T_{J,u}(W_n, C(u)) - W_n) = H'(I - R(u))Hz + H'(I - R(u))v. \quad (6.2)$$

At this point, one makes the essential *assumption* that the “error” vector v is small relative to Hz . The error v is ignored in calculating z , and z is defined by

$$H'(T_{J,u}(W_n, C(u)) - W_n) = H'(I - R(u))Hz. \quad (6.3)$$

Since $R(u)$ is a contraction and is a stochastic or substochastic matrix, $H'(I - R(u))H$ is invertible and we can write (6.3) as

$$z = [H'(I - R(u))H]^{-1}H'(T_{J,u}(W_n, C(u)) - W_n). \quad (6.4)$$

The “corrected” value $W_n + Hz$ can be used as the “improved” estimate of the solution. But because we have already calculated $T_{J,u}(W_n, C(u))$, it makes sense to use it to define the new estimate of the solution to (3.1) as

$$W_{n+1} = T_{J,u}(W_n + Hz, C(u)) = T_{J,u}(W_n, C(u)) + R(u)Hz. \quad (6.5)$$

This aggregation step is then followed by a sequence of (say) Gauss-Seidel or accelerated Gauss-Seidel relaxations, then another aggregation step is taken, and so on.

Interpretation of $(H'H)^{-1}H'R(u)H \equiv \{\bar{r}(i,j), i, j = 1, \dots, q\}$. This square matrix is actually a stochastic or substochastic matrix, whose entries can be interpreted as the degenerate transition probabilities of a Markov chain whose state space consists of the set of aggregation groups themselves. The degeneracy is implied by the contraction property of $R(u)$. In particular, $(H'H)^{-1}$ is a diagonal matrix whose diagonal entries are $n_1^{-1}, \dots, n_q^{-1}$, where n_i is the size of the i^{th} aggregation group and

$$\bar{r}(i,j) = \sum_{x \in S_i} \sum_{y \in S_j} r(x,y|u(x))/n_i. \quad (6.6)$$

Computation of the Inverse in (6.4). The evaluation of the inverse in (6.4) is fairly straightforward. The entries in the term $H'R(u)H$ are computed by (6.6) with the n_i deleted. Commonly, the matrix $R(u)$ is sparse, as is usual for matrices obtained as discretizations of elliptic type operators. See Chapter 5 for typical methods of construction. Typically, $H'(I - R(u))H$ is of a small enough order so that prepackaged methods for getting the inverse can be used and account for a negligible part of the total computation. With the problems that we have used it on, the entire aggregation step takes on the order of two or three times the time required for a single Gauss-Seidel relaxation.

An Adaptive Method for Getting H . For the type of Markov chains which arise as approximations to diffusion type processes, the state space S cannot usually be decomposed into disjoint groups such that the states within each group communicate (with a “high” probability) only to members of the same group. Given q , the method to be used to get the aggregation groups involves finding a partition $\{S_j, j \leq q\}$ of the state space such that the components of $T_{J,u}(W_n, C(u)) - W_n$ in each group are as close as possible. As will be seen below, under a certain assumption, this minimizes the values of the next residual $T_{J,u}(W_{n+1}, C(u)) - W_{n+1}$. The partition might be different each time the aggregation step is used. The following method, suggested in [11] has been found to be useful in the problems to which it was applied.

Define the projection Π by

$$\Pi = H(H'H)^{-1}H'.$$

Π is a projection onto the span of the columns of H . Suppose that we compute $W_{n+1}(u)$ by (6.5) and follow that by a Jacobi iteration. [A Gauss-Seidel iteration can also be used, but we are concerned here only with developing the logic for defining appropriate aggregation groups.] Then $W_{n+2} = T_{J,u}(W_{n+1}, C(u))$. It can be shown [11, equation (18)] that

$$T_{J,u}(W_{n+1}, C(u)) - W_{n+1} = (I - \Pi)(T_{J,u}(W_n, C(u)) - W_n) + (I - \Pi)R(u)Hz. \quad (6.7)$$

The method which will be used for choosing the aggregation groups ignores the right hand term in (6.7). The original motivation for this supposes that the state space is divided into groups which “essentially” communicate within themselves. Suppose, for the moment, that this is the case and that H is the membership matrix for this set of groups. Then the right hand term in (6.7) will be small, and it will be zero if the groups communicate only within themselves, since then $R(u)H$ is in the span of the columns of H . Of course, there will be no such “essentially self contained” groups of states for the chains which we use, but the approach still seems to be worth considering.

The membership matrix H to be used will be that which minimizes the first term on the right of (6.7) in a minimax sense [over the possible values of $R(u)$]. In order to see what one needs to do, first consider the following special cases. First, suppose that there is a real number ϵ_1 such that $T_{J,u}(W_n, C(u))(x) - W_n(x) = \epsilon_1$ for all x . Let there be a single aggregation group. Thus, $H = (1, \dots, 1) = \mathbf{e}$, a column vector. It follows that $(I - \Pi)(T_{J,u}(W_n, C(u)) - W_n) = 0$; i.e., the first term on the right of (6.7) is zero.

For the second example, suppose that there are real numbers $\epsilon_1, \dots, \epsilon_q$ such that $T_{J,u}(W_n, C(u)) - W_n$ equals ϵ_1 for the first n_1 states, ..., and ϵ_q for the last n_q states. Let H be the membership matrix with q columns and where the first n_1 states are in group 1, ..., and the last n_q states are in group q . Then $(I - \Pi)(T_{J,u}(W_n, C(u)) - W_n) = 0$, and the first term on the right of (6.7) is zero.

In general, we define H as follows: Define

$$\delta_{\min} = \min_{x \in S} \{T_{J,u}(W_n, C(u))(x) - W_n(x)\},$$

$$\delta_{\max} = \max_{x \in S} \{T_{J,u}(W_n, C(u))(x) - W_n(x)\}.$$

Define $\bar{\delta} = [\delta_{\max} - \delta_{\min}]$. Define

$$S_i = \left\{ x : T_{J,u}(W_n, C(u))(x) - W_n(x) \in \left[\delta_{\min} + \frac{(i-1)\bar{\delta}}{q}, \delta_{\min} + \frac{i\bar{\delta}}{q} \right] \right\}$$

with the interval for $i = q$ also being closed on the right. Additional motivation and discussion is in the main reference.

6.6.2 EXAMPLE AND RELATION TO THE PROCEDURE OF THEOREM 3.1

Consider the simplest case, where we wish to solve the special form of (3.1) given by $W = \gamma PW + C$, where P is either a stochastic or a substochastic matrix, $\gamma \in (0, 1)$, and the dimension of W is d . Suppose that we use a single aggregation class (i.e., $q = 1$). Thus, $H = \mathbf{e} = (1, \dots, 1)$, a column vector.

Then $H'H = d$, a scalar, $R(u)$ equals γP , and $H'(I - \gamma P)H = (1 - \gamma)d$, a scalar. Then the “correction” Hz equals

$$Hz = \frac{\mathbf{e}}{(1 - \gamma)d} \sum_x [T_{J,u}(W_n, C(u))(x) - W_n(x)],$$

and $W_n + Hz$ satisfies the bounds on $W(u)$ given in (3.2) and (3.3).

A Nondiscounted Problem. Suppose that there is no discounting, but the states in S are all transient. Then $r(x, y|u(x)) = p(x, y|u(x))$ for $x \in S$. Let us use a single aggregation group. Then $H = \mathbf{e}$ and

$$H'(I - R(u))H = d - \sum_{x,y \in S} r(x, y|u(x)).$$

Also,

$$Hz = \mathbf{e} \frac{\sum_x [T_{J,u}(W_n, C(u))(x) - W_n(x)]}{d - \sum_{x,y \in S} r(x, y|u(x))}.$$

In the context of the system (5.8.16) with $\beta = 0$, the denominator in the last expression can be interpreted as a “mean number” of absorbing states. It is the deficiency between the number of states d and the sum of all the elements of the matrix $R(u)$.

Comment on Computations. The aggregation method as it is presented above lacks a sound theoretical basis. The method works for Markov chain problems of the type in [11] and often does give an improvement for our type of problem as well, particularly in the earlier stages of policy iteration. But the convergence is not always monotonic, and the method is not recommended in the later stages. Generally, we find that the method does give good estimates of the optimal solution in less time than would be required by a direct relaxation method. But it must be used with care until sufficient experience is obtained. Generally, only a few aggregation steps should be taken.

If the state space has a reflecting boundary ∂G_h^+ , then it is important (for the best convergence) to coordinate the aggregation grouping of states in G_h and those in ∂G_h^+ . If a state $x \in \partial G_h^+$ is reflected to a $y \in G_h$ so that $W(x, u) = W(y, u)$, then x and y should be in the same aggregation group.

6.7 Coarse Grid-Fine Grid Solutions

The scalar approximation parameter h will be reintroduced in this and in the next section. A major problem in getting a “quick” solution to (1.1) or (1.2) via any type of iterative procedure concerns the initial condition which

is used to start the iteration. To get a good initial condition, it is useful to first obtain an approximate solution on a coarser state space, because the calculations on a smaller state space require less time. Let $h_2 > h_1 = h$, with the associated state spaces satisfying $S_{h_2} \subset S_{h_1} = S_h$. A useful procedure for the solution of (1.2) is to get an approximation $(\tilde{V}^{h_2}, \tilde{u}^{h_2})$ to the optimal value function and associated control on the coarse state space S_{h_2} , and then start the iteration on the finer state space S_h with an initial condition and control $(\tilde{V}^h, \tilde{u}^h)$ which are suitable interpolations of $(\tilde{V}^{h_2}, \tilde{u}^{h_2})$. For example, one might use the interpolation: $\tilde{V}^h(x) = \tilde{V}^{h_2}(x)$, for $x \in S_{h_2}$, and for $x \in S_h - S_{h_2}$, use a linear interpolation of the values in the smallest set of points in S_{h_2} in whose convex hull x lies, with an analogous interpolation for the control. It seems preferable to do several “smoothings” via a Gauss-Seidel (accelerated or not) relaxation before the first update of the control on the finer grid, whether iteration in policy space or in value space is used.

Such a procedure can be nested. Set $h = h_1 < \dots < h_k$, such that the associated state spaces satisfy

$$S_{h_k} \subset S_{h_{k-1}} \subset \dots \subset S_{h_1} = S_h. \quad (7.1)$$

Start the solution procedure by getting an acceptable approximation $(\tilde{V}^{h_k}, \tilde{u}^{h_k})$ to the optimal value function and control on the coarsest state space S_{h_k} . Then interpolate to get an initial condition and initial control for the iterative procedure for the solution on $S_{h_{k-1}}$, etc. This method is perhaps the crudest of the variable grid methods, but it does save considerable time over what is needed to get a solution of comparable accuracy by working on S_h directly. When applied to the optimization problem (1.2), the solution on a coarser state space might lack some of the local detail of the structure of the optimal control which one would get from a solution to the problem on a finer state space, but it quite commonly shows the salient features of the optimal control, and is itself often good enough for a preliminary study of the optimal control problem.

The state space refinement method discussed here, together with the approximation in policy space and the use of accelerated iterative methods, is relatively easy to program and is a good first choice unless the additional speed which the multigrid method allows is worth the additional programming effort. The use of such a state space refinement method still leaves open the choice of the actual method used for solving the problems on the chosen sequence of state spaces. For this one can use any of the procedures discussed so far, as well as the multigrid method of the next section.

6.8 A Multigrid Method

6.8.1 THE SMOOTHING PROPERTIES OF THE GAUSS-SEIDEL ITERATION

In this section, the state spaces will be referred to as grids. In the previous section, the problems (1.1) and (1.2) were approximately solved on a coarse grid. Then the approximate solution and associated control, suitably interpolated, were used as initial conditions for the solution of (1.1) or (1.2) on the original fine grid. As pointed out, one can nest the procedure and get an initial solution for the coarse grid by starting with an even coarser grid. In this section, we discuss another procedure, the so-called *multigrid method*, of exploiting the relative efficiency of different types of computations at different levels of grid refinement. The multigrid method is a powerful collection of ideas of wide applicability, and only a very brief introduction will be given here. A fuller discussion can be found in [14,97]. The multigrid method was introduced for use in solving optimal stochastic control problems by Akian and Quadrat [2,1], and a full discussion of data and convergence, under appropriate regularity conditions, can be found in [2]. Because variable grid sizes will be used, the basic scale factor h will be used.

A key factor in the success of the method is the “smoothing” property of the Gauss-Seidel relaxation. In order to understand this and to get a better idea of which computations are best done at the “coarser level,” it is useful to compare the rate of convergence of the Gauss-Seidel procedure when acting on smooth and on more oscillatory initial conditions. We will do this for a discretization of the simplest problem, where the system is $x(t) = x + w(t)$, where $w(\cdot)$ is standard Wiener process, and the process is absorbed on hitting the endpoints $\{0, 1\}$ of the interval of concern $[0, 1]$. For the discretization, let $h = 1/N$, where N is an integer. The absorbing points will be deleted, since they do not influence the rate of convergence. Then, the state space is $S_h = \{h, \dots, 1 - h\}$, a set with $N - 1$ points. Let $R^h = \{r^h(x, y), x, y \in S_h\}$ denote the transition matrix of the locally consistent approximating Markov chain which is absorbed on the boundary and is defined by

$$R^h = \begin{pmatrix} 0 & 1/2 & 0 \dots & 0 \\ 1/2 & 0 & 1/2 \dots & 0 \\ 0 & 1/2 & 0 \dots & 0 \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Since only the role of the initial condition is of concern, set $C^h(x) = 0$.

Consider the Gauss-Seidel relaxation $W_n^h \rightarrow W_{n+1}^h$, defined by

$$W_{n+1}^h(x) = \sum_{y < x} r^h(x, y) W_{n+1}^h(y) + \sum_{y \geq x} r^h(x, y) W_n^h(y). \quad (8.1)$$

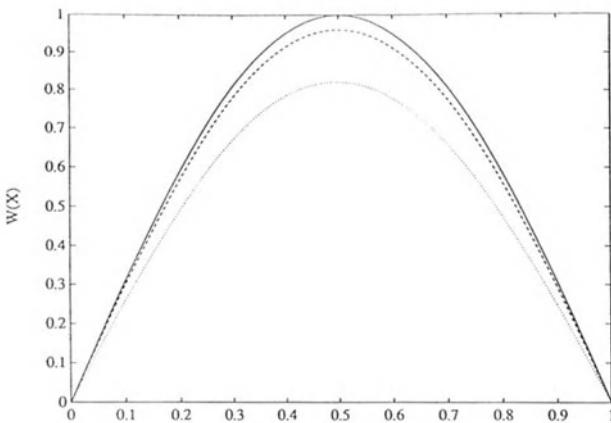


Figure 6.6. A Gauss-Seidel procedure. Smooth initial condition.

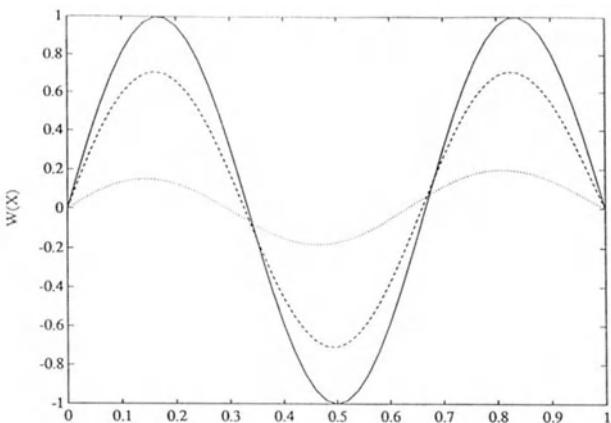


Figure 6.7. A Gauss-Seidel procedure. Oscillatory initial condition.

Refer to Figure 6.6, where $h = 1/N = 1/50$ and the initial condition is the solid line. The two dotted lines show the values after 10 and 50 iterations. Now refer to Figure 6.7, where the initial condition for the same problem is the oscillatory solid line, and the dotted lines have the same meaning as above. Note that the algorithm converges much faster here than in Figure 6.6 for the smoother initial condition. The situation just described is actually the general case for Markov chain approximations to the Wiener process with drift. Loosely speaking, the more “oscillatory” the initial condition [i.e., the energy in the initial condition being essentially in the higher eigenvalues of R^h] for the iteration (8.1), the faster the convergence. When the initial condition in (8.1) is “smooth,” it is reasonable to suppose that we can get a quite good approximation to the solution of (1.1) by working with a coarser grid, because the errors due to the “projection” onto

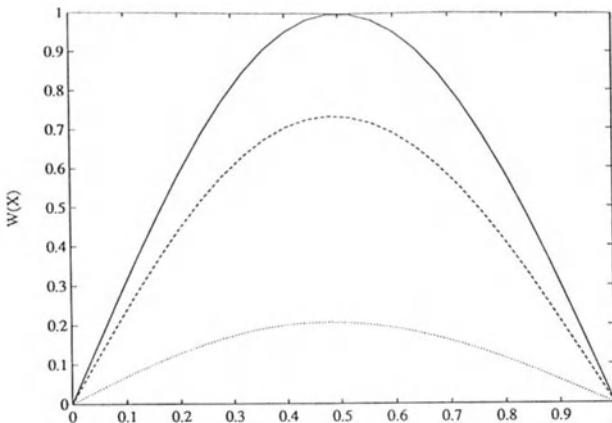


Figure 6.8. Smooth initial condition, coarser grid.

the coarser grid or the “interpolation” of the resulting value back onto the finer grid would not be “large.”

See Figure 6.8, where the problem with the smooth initial condition is treated on a grid with the coarser spacing $2h$. The times required for the computation for the cases of Figure 6.8 are approximately the same as those for the cases in Figure 6.6, although the result is much better in Figure 6.8. The smoother the initial condition, the better the quality of the approximation on a coarser grid. The above comments provide an intuitive picture of the role of the computations on the coarser grid as well as of the smoothing property of the Gauss-Seidel relaxation. This “smoothing” property provides one important basis for the multigrid method. Fuller and very enlightening discussions appear in the introductory books [14,97].

The cited “smoothing” property of the Gauss-Seidel relaxation is not quite shared by the Jacobi relaxation: $W_{n+1}^h = R^h W_n^h + C^h$. Consider the initial condition W_0^h defined by $W_0^h(x) = \sin(k\pi x)$, $1 \leq k < N$. Then the iteration does converge faster as k increases, up to about the “midpoint” $k = N/2$, after which the convergence slows down again. But the weighted Jacobi method can be used to get the faster convergence for the more oscillatory initial conditions, for weights ω less than unity (recall that we used weights greater than unity in Section 6.4, but there the weighting served a different purpose). See [14, Chapter 2] for a fuller discussion.

The so-called smoothing properties of the Gauss-Seidel and weighted Jacobi iteration are shared by systems other than the “random walk” discretization of the Wiener process given above. A fuller discussion is beyond our scope here. But we note that multigrid methods of the type discussed in the next subsection seem to work well on all of the problem classes to which they were applied, including the ergodic, heavy traffic and singular control problems of Chapters 7 and 8, even when their use could not be rigorously

justified. It should be kept in mind that the entire discussion concerning smoothness is heuristic, since not much is known about the smoothness of the solutions of the equations of interest, or about the smoothing properties of the Gauss-Seidel relaxations which arise from Markov chain approximations to “degenerate” control problems.

The previous discussion suggests a procedure of the following type for the solution of the problem (1.1). First, do several Gauss-Seidel relaxations (accelerated or not) on the finer grid. The error between the trial solution at this stage and the true solution should be “smoother” than the error at the start. Then estimate the error via a computation on a coarser grid. With the estimate in hand, correct the last trial solution on the finer grid and continue. This rough outline will now be formalized. The reader should keep in mind that there are numerous variations of the general idea, and we present only the simplest concepts. Our interest is in providing motivation and a rough outline. More detail and much practical information is in [96]. Detail on the stochastic problem is in [2].

6.8.2 A MULTIGRID METHOD

We will be concerned with the solution of the equation (1.1) on the state space S_h , and we rewrite (1.1) here for convenience

$$W^h(u) = R^h(u)W^h(u) + C^h(u). \quad (8.2)$$

Let the operator $T_u^h(X^h, C^h(u))$ denote either the Gauss-Seidel or accelerated Gauss-Seidel relaxation (2.10), (4.2), or (4.3) applied to (8.2):

$$X^h \rightarrow T_u^h(X^h, C^h(u)).$$

For the Gauss-Seidel relaxation,

$$T_u^h(X^h, C^h(u)) = \tilde{R}^h(u)X^h + \tilde{C}^h(u),$$

where the quantities $\tilde{R}^h(u)$ and $\tilde{C}^h(u)$ are defined by (2.13) and (2.14).

These relaxations will be used below with various values of h and replacements for $C^h(u)$. because a variable grid method will be used and it will be necessary to work with S_h as well as “coarser” state spaces, a simplifying assumption will be made on the state space. Once the underlying ideas are understood, it should be apparent that the approach can be applied to other “shapes,” provided that the appropriate projection and interpolation operators (see below) can be defined. The state space S_h will be a “regular rectangular grid,” defined as follows. Let G be a hyperrectangle in \mathbb{R}^r which is centered about the origin. Let $S_h = G \cap \mathbb{R}_h^r$. For some integer $k > 1$, let each side of the hyperrectangle be an integral multiple of $2^{k-1}h$. This last requirement is introduced so that each coarser grid can be defined as a subset of the finer grids, without any further notational

difficulties. k is called the *level* of the procedure. Define the sequence of grid spacings $h_i = h2^{i-1}$, $i = 1, \dots, k$, and the associated state spaces $S_{h_i} = \mathbb{R}_{h_i}^r \cap G = \mathbb{R}_{h_i}^r \cap S_h$. Thus, $h_1 = h$ and $R^{h_i}(u)$ is the $R^h(u)$ in (8.2) with h replaced by h_i .

Next, a specific method of splitting the computation between the different levels will be motivated. That will be followed by a somewhat loose outline of the realization for a two-level procedure, and then a general k -level procedure will be defined. Now consider the following procedure. Let W_0^h denote an initial guess of the solution to (8.2). Then, starting with the initial condition $X^h = W_0^h$, do several (say, n_1) Gauss-Seidel or accelerated Gauss-Seidel relaxations (8.3)

$$X^h \rightarrow T_u^h(X^h, C^h(u)), \quad (8.3)$$

ending up with the final value which will be denoted by W_1^h . Next, define the *residual* ρ_0^h via a Jacobi (or Gauss-Seidel) relaxation

$$\rho_0^h = T_{J,u}^h(W_1^h, C^h(u)) - W_1^h, \quad (8.4)$$

and define the error $\delta W_1^h = W^h(u) - W_1^h$. Then we can write

$$\delta W_1^h = R^h \delta W_1^h + \rho_0^h. \quad (8.5)$$

If (8.5) can be solved, then the solution to (8.2) is given by

$$W^h(u) = W_1^h + \delta W_1^h. \quad (8.6)$$

Loosely speaking, the discussion in the last subsection concerning the “smoothing” properties of the Gauss-Seidel relaxation suggests that the residual ρ_0^h , which is defined on the points of S_h will be “smooth.” This suggests that a good approximation to the solution to (8.5) can be obtained by solving a “projection” of (8.5) onto a coarser grid (say, S_{2h}), and then “interpolating the approximation back” to the original grid S_h . This discussion will now be formalized as an algorithm.

A Simple Two Level Multigrid Procedure. In order to carry out the program outlined above, we need to be more precise concerning the notions of projection from a finer grid to a coarser grid and interpolation from a coarser grid to a finer grid. Consider the following two level method. Define the *projection* or *restriction operator* I_h^{2h} from S_h to S_{2h} by

$$(I_h^{2h} \rho)(x) = \rho(x), \text{ for } x \in S_{2h}.$$

Thus, I_h^{2h} acting on the function $\rho(\cdot)$ simply picks out the values for the points which are common to both S_h and S_{2h} . This is perhaps the crudest projection. See further comments at the end of the section.

(a) Given the initial guess W_0^h of the solution to (8.2), do n_1 accelerated Gauss-Seidel iterations, getting an updated estimate of $W^h(u)$ which will be called W_1^h . Solve for the residual ρ_0^h by (8.4).

(b) Then, starting with initial condition $\delta X^{2h} = 0$, get an approximate solution δW_1^{2h} (say, via n_2 accelerated Gauss-Seidel iterations) to the equation

$$\delta X^{2h} = R^{2h}(u)\delta X^{2h} + I_h^{2h}\rho_0^h \quad (8.7)$$

on S_{2h} .

(c) Next, given δW_1^{2h} , we need to interpolate it “back to the finer” grid S_h . To do this, define an *interpolation* operator I_{2h}^h which takes functions defined on S_{2h} into functions defined on S_h . One natural choice is the following: Let ρ be defined on the points of S_{2h} . For $x \in S_{2h}$, set $(I_{2h}^h\rho)(x) = \rho(x)$. For $x \in S_h - S_{2h}$, define $(I_{2h}^h\rho)(x)$ to be the linear interpolation of the smallest set of points in S_{2h} in whose convex hull x lies. This is perhaps the simplest type of interpolation operator. The new trial solution to (8.2) is now defined by the updated W_1^h

$$W_1^h \rightarrow W_1^h + I_{2h}^h\delta W_1^{2h}. \quad (8.8)$$

(d) The procedure is now repeated. Starting with the initial value W_1^h , get a new trial solution W_2^h to (8.2) [say, by taking n_1 accelerated Gauss-Seidel iterations (8.3)], calculate the residual $\rho_1^h = T_{J,u}^h(W_2^h, C^h(u)) - W_2^h$, project the residual onto S_{2h} , get an approximate solution δW_2^{2h} to (8.7), where ρ_1^h replaces ρ_0^h , interpolate δW_2^{2h} back to S_h by use of I_{2h}^h , and compute the new trial solution to (8.2) by updating as

$$W_2^h \rightarrow W_2^h + I_{2h}^h\delta W_2^{2h}, \quad (8.9)$$

and continue until the given stopping criterion is satisfied.

A General k -Level Method. For $i = 1, \dots, k-1$, define the projection operators $I_{h_i}^{h_{i+1}}$ and interpolation operators $I_{h_{i+1}}^{h_i}$ analogously to the definitions used for $i = 1$ in the two-level case above, and let n_1, \dots, n_k be given integers. Set $n = 0$, and let W_0^h be an initial guess of the solution to (8.2). Consider the following procedure.

1. Starting with the initial value $X^h = W_n^h$, do n_1 iterations of

$$X^h \rightarrow T_u^h(X^h, C^h(u)),$$

and denote the final value by W_{n+1}^h .

2. Calculate the residual ρ_n^h by

$$\rho_n^h = T_{J,u}^h(W_{n+1}^h, C^h(u)) - W_{n+1}^h.$$

3. Define $f_n^{2h} = I_h^{2h}\rho_n^h$, and starting with the initial value $\delta X^{2h} = 0$, do n_2 iterations of

$$\delta X^{2h} \rightarrow T_u^{2h}(\delta X^{2h}, f_n^{2h}),$$

ending up with the final value which we denote by δW_{n+1}^{2h} .

4. Calculate the residual ρ_n^{2h} by

$$\rho_n^{2h} = T_{J,u}^{2h}(\delta W_{n+1}^{2h}, f_n^{2h}) - \delta W_{n+1}^{2h}.$$

5. Define $f_n^{4h} = I_{2h}^{4h}\rho_n^{2h}$, and starting with the initial value $\delta X^{4h} = 0$, do n_3 iterations of

$$\delta X^{4h} \rightarrow T_u^{4h}(\delta X^{4h}, f_n^{4h}),$$

ending up with the final value which we denote by δW_{n+1}^{4h} .

6. Continue until level k is reached and $\delta W_{n+1}^{h_k}$ is obtained.

7. Given $\delta W_{n+1}^{h_k}$, reset the value $\delta W_{n+1}^{h_{k-1}}$ at the $(k-1)^{st}$ -level as

$$\delta W_{n+1}^{h_{k-1}} \rightarrow \delta W_{n+1}^{h_{k-1}} + I_{h_k}^{h_{k-1}}\delta W_{n+1}^{h_k}.$$

With this new initial condition, $X^{h_{k-1}} = \delta W_{n+1}^{h_{k-1}}$, do n_{k-1} iterations of

$$\delta X^{h_{k-1}} \rightarrow T_u^{h_{k-1}}(\delta X^{h_{k-1}}, f_n^{h_{k-1}}),$$

ending up with the final value which is also denoted by $\delta W_{n+1}^{h_{k-1}}$.

⋮

8. Reset the value of δW_{n+1}^h as $W_{n+1}^h \rightarrow W_{n+1}^h + I_{h_2}^h \delta W_{n+1}^{2h}$, and with this new initial condition, do n_2 iterations of

$$X^h \rightarrow T_u^h(X^h, C^h(u)),$$

ending up with the final value which is also denoted by W_{n+1}^h .

9. If the stopping criterion is satisfied, then stop. Otherwise, increase n by 1 and return to step 1 or 2.

The Gauss-Seidel relaxation can also be used for calculating the residual. The procedure just described is referred to as a V -cycle, since one descends from the highest (first) level to the lowest (k^{th}) and then back to the highest. One can vary the procedure in many ways; for example, the so-called W -method goes from the highest level to the lowest, then back to an intermediate level, then down to the lowest, and finally back up to the highest. Such methods are often preferred for the solution of the linear equations which arise from discretizations of many types of partial differential equations. See e.g., the proceedings in [94].

Comments on Computation. The method seems to be robust. It has been used successfully on all of the classes of problems dealt with in this book. It tends to be as good as the other methods, and is generally superior. It is best used as part of a variable grid scheme such as in Section 6.7, where it might be used to solve the appropriate optimization problem at each level.

There is no convergence proof available for the general problem of interest here, and it has been observed on some problems that the values eventually oscillate slightly. But, even then the asymptotic errors have been quite small for the problems that have been tried by the authors. In any case, much more experimentation is needed on the great variety of stochastic control problems.

Due to the crudity of the projection operator, as well as to the poorly understood structure of the general problems of interest, we generally used several accelerated Gauss-Seidel relaxations between each level of projection/interpolation. Thus these relaxations actually contribute to the convergence via their contraction as well as by their smoothing property. In much multigrid practice, acceleration is not used, and one uses fewer relaxations, as well as more sophisticated schemes such as the W -cycle. It was pointed out in [53] that the smoothing property of the accelerated Gauss-Seidel procedure deteriorates as the acceleration parameter increases to its optimal value.

The projection operator $I_{h_k}^{h_{k-1}}$ described above is the simplest possible. It is not a “smooth” operator, and the success of its use is increased with added smoothing relaxations. A common alternative is to define the projection $I_{h_k}^{h_{k-1}} \rho^{h_k}(x)$ to be the arithmetic mean of the $\rho^{h_k}(y)$ for points y which can be reached in one step (including diagonals) from x . Alternatively, define the interpolation operator first and then define the projection operator to be its transpose. The R^{h_k} on the coarser grids can also be chosen in a more sophisticated manner. See the references for more detail. We also direct the readers attention to the algebraic multigrid method [96], where the interpolation and projection operators (as well as the coarser grids) are tailored to the actual values of the transition probabilities.

6.9 Linear Programming Formulations and Constraints

6.9.1 LINEAR PROGRAMMING

The solution to (1.4) can also be obtained via linear programming (LP). At this time, it does not appear that a direct use of the simplex method of linear programming is better than any of the other methods discussed previously in this chapter. But an LP formulation does allow the introduction of “mean value” constraints on the paths or control effort, and the LP codes are being improved continuously. First, the formulation of the general linear programming problem will be reviewed. Then the problem of computing the minimum values $V(x)$ and minimizing controls will be put into an LP form, and both the primal and dual equations exhibited. It will be seen that the dual equations are just (1.4). The basic approach follows

the development in [84]. Assumption (A1.2) will be used throughout.

An optimal stopping problem with a constraint on the mean time to stopping was discussed in [76]. Problems with reflected diffusion models and constraints were dealt with in [85] and [123], and will not be discussed here. Other references which have dealt with LP formulations of Markov chain optimization problems are [32] and [67]. Only a few basic facts concerning linear programming will be stated, and the reader is referred to any of the many standard references (e.g., [8,50]) for more detail.

The Basic LP. For a column vector b , row vector c , and a matrix A , all of compatible dimensions, the basic LP form is: Choose $X = \{X_1, \dots\}$ satisfying

$$AX = b, X \geq 0 \quad (9.1)$$

and which minimizes

$$z = cX.$$

All inequalities between vectors are assumed to hold component by component. Thus $X \geq 0$ means that all components of X are non-negative. Equation (9.1) is called the *primal problem*. The *dual problem* is defined to be

$$\max b'Y : Y'A \leq c. \quad (9.2)$$

Let A_i denote the i^{th} column of A . Define the row vector $D = c - Y'A$. In terms of the components, $D_i = c_i - Y'A_i$. The so-called *complementary slackness* condition is

$$X_i D_i = 0, \text{ for all } i. \quad (9.3)$$

We say that a vector $X \geq 0$ (resp., Y) is primal (resp., dual) feasible if $AX = b$ (resp., $Y'A \leq c$). A well known fact in LP is that a vector X is optimal if and only if it is primal feasible, and there is a dual feasible Y such that complementary slackness holds.

The Simplex Procedure. Only a rough outline will be given. Let m denote the number of rows of A and q the number of columns (the dimension of the vector X) and suppose that A has more columns than rows. Let A be of full rank (this is the case in the Markov chain optimization problem). Suppose that there is a primal feasible solution $X = (X_1, \dots, X_q)$ which can be partitioned as follows: There are $I(X) = \{i_1, \dots, i_m\}$, such that for $j \notin I(X)$, we have $X_j = 0$, and the matrix $\{A_{i_1}, \dots, A_{i_m}\} \equiv B$ is nonsingular. Then X is said to be a *basic feasible solution*, the components $\{X_j, j \in I(X)\}$ are said to be the *basic variables* or basis, and B is said to be the *basis matrix*. If there is an optimal solution, then there is one in the class of basic feasible solutions. The simplex procedure computes the optimal by getting a minimizing sequence in the class of basic feasible solutions. Suppose that we have a basic feasible solution X . Let us reorder the components so that we can write X as $X = (X_B, X_{NB})$, where X_B is the vector of basic

variables and X_{NB} is the vector of nonbasic variables. Similarly partition the matrix A as $A = [B, (NB)]$, and write $c = (c_B, c_{NB})$.

With this new notation, we can write the cost function and constraints (9.1) as

$$\begin{aligned} c_B X_B + c_{NB} X_{NB} &= z, \\ BX_B + (NB)X_{NB} &= b. \end{aligned} \quad (9.4)$$

We wish to replace the current basis by one with a smaller cost, if possible. To help us see what needs to be done, the equation (9.4) will be transformed into a form which allows us to observe directly the derivatives of the cost with respect to each of the basic and nonbasic variables. Let us subtract a linear combination of the rows of the second equation of (9.4) from the first equation such that the term $c_B X_B$ is cancelled. The multipliers in the linear combination are easily seen to be $Y' = c_B B^{-1}$. Then rewrite (9.4) as

$$(0) X_B + [c_{NB} - Y'(NB)]X_{NB} + Y'b = z, \quad (9.5)$$

$$X_B + B^{-1}(NB)X_{NB} = B^{-1}b. \quad (9.6)$$

Because $X_{NB} = 0$, (9.5) implies that $z = c_B B^{-1}b$, the cost under the current basis. The multiplier vector Y is not necessarily dual feasible.

At the current basic solution X , equation (9.5) implies that $D_i = c_i - Y'A_i$ is the derivative of the cost with respect to the variable X_i . Since $D_i = 0$ if X_i is a basic variable, the derivative of the cost (at the current vector X) with respect to a basic variable is zero. Note that if $D_i \geq 0$ for all i , then Y is dual feasible, and complementary slackness holds. In the actual simplex procedure, B^{-1} is not actually explicitly computed anew at each update of the basis. It is calculated in a relatively simple manner using the previous basis inverse.

If $D_i \geq 0$ for all i , then the current solution is optimal, and the procedure can be stopped. Otherwise, it is possible to select an improved basic solution. To do this, define an index i_0 by

$$\min_j D_j = D_{i_0}$$

and introduce the nonbasic variable X_{i_0} into the basis at the maximal level consistent with the preservation of feasibility. This will involve the elimination of one of the current basic variables from the basis. Note that

$$X_B + (B^{-1}A_{i_0})X_{i_0} = B^{-1}b, \quad (9.7)$$

It can be seen from (9.7) that as X_{i_0} increases from the value zero, at least some of the components of the vector X_B might have to change value to assure feasibility. There are three possibilities. Case (a): X_{i_0} can increase without bound, with the cost going to $-\infty$. This would occur if none of the components of $B^{-1}A_{i_0}$ were positive. In this case, stop. Case (b): It might

be impossible to raise X_{i_0} above zero, without driving some other basic variable X_{i_1} (which would have to have been at the zero level) negative. In this case, simply replace X_{i_1} by X_{i_0} in the basis. Case (c): X_{i_0} can be increased to a finite but nonzero level before some other basic variable, say X_{i_1} , becomes zero. Then replace X_{i_1} by X_{i_0} in the basis and set it at the maximum level. The procedure is then repeated. The exact details of this so-called “pivoting” procedure are not important for the discussion here. The method is not quite a gradient procedure, since it attempts to get a decrease in cost by increasing the value of only a single nonbasic variable at a time, with the other variables changing their values only to the extent necessary to maintain feasibility.

6.9.2 THE LP FORMULATION OF THE MARKOV CHAIN CONTROL PROBLEM

The problem of computing the minimum cost and minimizing control for the Markov chain control problem [i.e., solving (1.4)] will now be set up as a LP problem. Assume (A1.2). If, for a given control $u(\cdot)$, $R(u)$ in (1.4) is a contraction, then $\sum_y r(x, y|u(x)) < 1$ for at least one state x . But $R(u)$ is still a transition probability matrix for some “killed” Markov chain, which we denote by $\{\xi_n, n < \infty\}$. In this subsection, the variables i, j will be used to denote the states of the chain, and not the x, y of the previous sections.

The development is simpler if the control action space \mathcal{U} contains only a finite number of points. Many cases of practical interest reduce to this. For example, let $\mathcal{U} = [-1, 1]$, and suppose that the dynamics of the original system are linear in the control and the cost function either does not depend explicitly on the control or else depends linearly on its absolute value. Then the control will take one of the three values $\{0, -1, 1\}$, so an a priori restriction to this finite set can be made. If the finiteness assumption does not hold, then the so-called “generalized programming” [88] could be used and would lead to the same conclusions. Let m denote the number of points in S , the state space of the chain, and write $\mathcal{U} = \{\alpha_1, \dots, \alpha_L\}$. The LP formulation is in terms of randomized Markov or (equivalently) randomized feedback controls. Thus, the control $u(\cdot)$ is to be defined by the conditional probabilities

$$\gamma_{ik} = P\{u_n = \alpha_k | \xi_q, u_q, q < n; \xi_n = i\} = P\{u_n = \alpha_k | \xi_n = i\}.$$

Suppose that the initial state ξ_0 is a random variable with the probability distribution

$$P\{\xi_0 = i\} = \rho_i > 0, \text{ for all } i \in S.$$

It will turn out that the values of ρ_i do not affect the optimal control, unless there are added constraints [84]. Let E_ρ denote the expectation of functionals of the Markov chain which is associated with whatever $R(u)$ is used and the distribution ρ of the initial condition ξ_0 . (The $u(\cdot)$ will be

clear from the context.) For $i \in S$ and $k \leq L$, let M_{ik} denote the joint (state, control) mean occurrence times defined by

$$M_{ik} = \sum_{n=0}^{\infty} E_{\rho} I_{\{\xi_n=i, u_n=\alpha_k\}}.$$

Thus, M_{ik} is the mean number of times that the state is i and the action is α_k simultaneously. We emphasize that these mean values are for the chain with the degenerate transition matrix $R(u)$. The probability that action α_k is used when the state is i is

$$\gamma_{ik} = \frac{M_{ik}}{M_i}, \quad (9.8)$$

where $M_i \equiv \sum_k M_{ik}$. The mean occupancy times satisfy the equation

$$\begin{aligned} M_i &= \rho_i + \sum_{j,k} r(j, i | \alpha_k) M_{jk}, \\ M_{ik} &\geq 0. \end{aligned} \quad (9.9)$$

In terms of the control variables, (9.9) equals

$$M_i = \rho_i + \sum_{j,k} r(j, i | \alpha_k) \gamma_{jk} M_j. \quad (9.9')$$

Equation (9.9) is the primal constraint equation, the analogue of the constraint (9.1) for the Markov chain problem. Let $u(\cdot)$ denote the randomized control defined by (9.8). Let us define the total weighted cost $W_{\rho}(u) = \sum_i W(i, u) \rho_i$ in terms of the M_{ik} as

$$W_{\rho}(u) = \sum_{i,k} C(i, \alpha_k) M_{ik} \equiv cM, \quad (9.10)$$

which is a linear function of the M_{ik} . The constraints (9.9) and cost (9.10) constitute the LP formulation of the optimization problem. Additional constraints of the form

$$\sum_{i,k} d_{ikj} M_{ik} \leq q_j$$

can be added.

Comments. Any basic feasible solution to (9.9) will have at most m nonzero elements. By definition, the basis contains exactly m of the variables M_{ik} . Because we have assumed that $\rho_i > 0$ for all i , for each i we must have $M_{ik} > 0$ for at least one value of k . In addition, the number of constraints in the first line of (9.9) equals the number m of states in S . These considerations imply that, for each i , there is one and only one

value of k [to be called $k(i)$] for which $M_{ik} > 0$. Otherwise, more than m of the variables M_{ik} would be positive. Thus, any basic feasible solution yields a pure Markov control: There is no randomization. If additional linear constraints were added to (9.9), then the basic solution might contain randomized controls for some of the states.

Suppose that $u(\cdot)$ is a control which is determined by the probability law $\{\gamma_{ik}\}$. That is,

$$P\{\xi_{n+1} = j | \xi_n = i, u(\cdot) \text{ used}\} = \sum_k r(i, j | \alpha_k) \gamma_{jk} \equiv r(i, j | u).$$

Let $R(u) = \{r(i, j | u), i, j \in S\}$ be a contraction. Then (9.9') has a unique solution, and this solution must be the vector of mean occupancy times. The solutions to (9.9) and (9.9') are identical. Hence, the solution to (9.9) is also the vector of mean occupancy times.

Let $u_0(\cdot)$ be a pure Markov (nonrandomized) control such that $R(u_0)$ is a contraction. Such a control exists by assumption (A1.2). Let us start the simplex iteration with this control. That is, the variables in the starting basis are $\{M_{iu_0(i)}, i \in S\}$. Let $\{M_{ik}^0\}$ denote the associated initial feasible solution to (9.9). Let $M^n = \{M_{ik}^n, i, k\}$ denote the n^{th} feasible solution generated by the simplex procedure, and define $\gamma_{ik}^n = M_{ik}^n / M_i^n$. Let $\{u_n(\cdot)\}$ denote the associated control law. The sequence of costs cM^n is nonincreasing in n . Hence, by (A1.2) and an argument like that used in the last paragraph, the $R(u_n)$ are all contractions, the solution to (9.9) and (9.9') are equal and unique, and $\{M_{ik}^n, i, k\}$ are the mean occupancy times for all n . See [84] for proofs and further discussion.

The Dual LP. Let Y_i denote the dual variables for the LP (9.9), (9.10). Then, by analogy to the relation between (9.1) and (9.2), we see that the dual equations to (9.9) can be written as

$$Y_i \leq \sum_j r(i, j | \alpha_k) Y_j + C(i, \alpha_k), \quad \text{for all } i, k. \quad (9.11)$$

The dual cost is

$$\sum_i \rho_i Y_i. \quad (9.12)$$

The complementary slackness conditions (9.3) can be expressed in the following form:

- (a) If (9.11) is a strict inequality for i, k , then $M_{ik} = 0$.
- (b) If $M_{ik} > 0$, then there is strict equality in (9.11) for i, k .

Recall that $M_{ik} > 0$ for one and only one value of k for each i . Thus the complementary slackness conditions reduce to

$$Y_i = \min_k \left[\sum_j r(i, j | \alpha_k) Y_j + C(i, \alpha_k) \right] \quad (9.13)$$

which is just the dynamic programming equation (1.4). Thus, the optimal dual variables are the minimum cost values: $Y_i = V(i)$.

The Simplex Procedure and Approximation in Policy Space. Let us recapitulate the simplex procedure for the Markov chain optimization problem. The appropriate form of the “reduced cost” variables $D_\beta = c_\beta - Y' A_\beta$ are

$$D_{ik} = C(i, \alpha_k) + \sum_j r(i, j | \alpha_k) Y_j - Y_i. \quad (9.14)$$

Let $M_{ik(i)}$, $i \in S$, denote the current set of basic variables. Then $D_{ik(i)} = 0$ for all $i \in S$. By examining the set of linear equations

$$0 = C(i, \alpha_{k(i)}) + \sum_j r(i, j | \alpha_{k(i)}) Y_j - Y_i, \quad (9.15)$$

we see that they are the same as (1.3) for the current control $u(i) = k(i)$. This implies that $Y_i = W(i, u)$. If $D_{ik} \geq 0$ for all i and k , then stop, since optimality has been achieved. Otherwise, define (i_0, k_0) by

$$D_{i_0 k_0} = \min_{i, k} D_{ik}.$$

Then the new control for state i_0 will be α_{k_0} . The controls for the other states $i \neq i_0$ remain as before. The pivoting procedure calculates the cost for the new control. Thus, approximation in policy space is equivalent to the simplex procedure if the control for only one state is changed on each policy update (as opposed to the possibility of changing the controls for all of the states simultaneously).

The Ergodic Cost Problem: Formulation and Algorithms

In this chapter, we reformulate some of the concepts in the last chapter so that they can be used on the ergodic cost problem. Before doing that it is useful to discuss the appropriate dynamic programming equations and some additional background material. The natural state spaces for control problems that are of interest over a long time interval are often unbounded, and they must be truncated for numerical purposes. One standard way of doing this involves a reflecting boundary, and this is the case dealt with in this chapter. Thus, there are no absorbing states. The basic process is the controlled diffusion (5.3.1) or jump diffusion (5.6.1). The approximating Markov chains $\{\xi_n^h, n < \infty\}$ will be locally consistent with these processes. As in the previous chapters, S_h denotes the state space, and $\partial G_h^+ \subset S_h$ the set of reflecting states. Recall that the reflection is instantaneous and that we use $\Delta t^h(x, \alpha) = 0$ for $x \in \partial G_h^+$. For a feedback control $u(\cdot)$, the cost function of interest for the original process is

$$\gamma(x, u) = \limsup_T \frac{1}{T} \int_0^T E_x^u k(x(s), u(x(s))) ds,$$

where the function $k(\cdot)$ is continuous. If the limit does not depend on the initial state x , then we omit it from the notation.

In the first few sections, we leave the approximation problem aside and work with general controlled Markov chain models. In Section 7.1, the appropriate form of the dynamical equation for the value function for the ergodic cost problem for a fixed control is given, as well as the dynamic programming equation for the optimal control and cost. Sections 7.2 and 7.3 concern the appropriate forms of the approximation in value space and approximation in policy space algorithms for the case where there is a single ergodic class under each feedback control. The matrices $P(u)$ which appear in these algorithms are not contractions. To formulate the numerical algorithms, a centered form of the transition matrix is introduced in Section 7.3. This centered form enjoys the contraction property. The appropriate adaptations of the numerical algorithms of Chapter 5 are discussed in Section 7.4.

In Section 7.5, we return to the ergodic cost problem for the approximating chain. The appropriate form of the cost function for the approximating

Markov chain is given together with a heuristic discussion of why it is the correct one. The dynamic programming equation for the optimal value function is also stated. Owing to the possibility of nonconstant interpolation intervals, and to our desire that the limits of the cost functions for the chains approximate that for the original diffusion or jump diffusion, the formulation is slightly different than the one used for the Markov chain model of Sections 7.1-7.4. In Section 7.6, it is shown that the dynamic programming equation given in Section 7.5 is also the correct one for the analogous ergodic cost problem for the continuous parameter Markov chain interpolation $\psi^h(\cdot)$ introduced in Chapter 4.

The main difficulty in directly applying the computational methods of Chapter 5 to the approximating chain and the ergodic cost problem stems from the fact that the interpolation interval might depend on either the state or the control. With this dependence, the dynamic programming equation for the ergodic cost problem cannot usually be easily solved by a recursive procedure. In Section 7.7, we review the procedure for getting chains with constant interpolation intervals from general approximating chains and discuss the computational consequences.

In Sections 7.6 and 7.7, we suppose that there is neither a cost nor a control acting on the boundary (the boundary must be reflecting and not absorbing). Section 7.8 gives the few changes that are needed for a more general case.

7.1 The Control Problem for the Markov Chain: Formulation

Let $\{\xi_n, n < \infty\}$ be a controlled Markov chain on a finite state space S . The control at each state x takes values in the set \mathcal{U} . For an admissible feedback law $u(\cdot)$, let $P(u) = \{p(x, y|u(x)), x, y \in S\}$ denote the corresponding transition matrix. Unless otherwise mentioned, we use the following assumptions.

A1.1. *For each control, the state space consists of transient states plus a single communicating aperiodic class.*

A1.2. *$C(x, \cdot)$ and $p(x, y|\cdot)$ are continuous functions of the control parameter.*

Other conditions as well as replacements for (A1.1) will be stated as needed. For a feedback control $u(\cdot)$, let the cost function be

$$\gamma(x, u) = \limsup_N \frac{1}{N} E_x^u \sum_{n=0}^{N-1} C(\xi_n, u(\xi_n)). \quad (1.1)$$

If the \limsup is not a function of the initial condition, then it will be written as $\gamma(u)$. Assumption (A1.1) does cover many applications, and it simplifies the development considerably. It does not hold in all cases of interest for the approximating Markov chains. For example, in the singular control problem of Chapter 8, one can easily construct controls for which the assumption is violated. Nevertheless, the algorithms seem to perform well for the “usual” problems on which we have worked. Weaker conditions are used in Theorem 1.1, under a finiteness condition on \mathcal{U} .

The Functional Equation for the Value Function for a Given Control. Let us recapitulate some of the statements in Subsection 2.1.3. By (A1.1), for each feedback control $u(\cdot)$, there is a unique invariant measure which will be denoted by the row vector $\pi(u) = \{\pi(x, u), x \in S\}$, and (1.1) does not depend on x . As noted in Subsection 2.1.3, there is a vector valued function $W(u)$ with values $\{W(x, u), x \in S\}$, such that $(W(x, u), \gamma(u))$ satisfy

$$W(x, u) = \sum_y p(x, y|u(x))W(y, u) + C(x, u(x)) - \gamma(u). \quad (1.2)$$

In fact, the function defined by

$$W(x, u) = \sum_{n=0}^{\infty} E_x^u[C(\xi_n, u(\xi_n)) - \gamma(u)] \quad (1.3)$$

satisfies (1.2).

The solution to (1.2) is not unique, since if $(W(u), \gamma(u))$ is a solution and k a constant, then the function defined by $W(x, u) + k$ together with the same $\gamma(u)$ is also a solution. However, it is true that if we normalize the set of possible solutions by selecting some state $x_0 \in S$ and insisting that $W(x_0, u) = K$, for some given constant K , then the solution $(W(u), \gamma(u))$ is unique. Generally, K is chosen to equal either zero or $\gamma(u)$.

Let $(\tilde{W}, \tilde{\gamma})$ satisfy

$$\tilde{W}(x) = \sum_y p(x, y|u(x))\tilde{W}(y) + C(x, u(x)) - \tilde{\gamma}. \quad (1.4)$$

Then $\tilde{\gamma} = \gamma(u)$. To see this, multiply each side of (1.4) on the left by $\pi(x, u)$, sum over x , and use the fact that $\pi(u)P(u) = \pi(u)$.

The Dynamic Programming Equation. Define $\bar{\gamma} = \inf_u \gamma(u)$, where the infimum is over all feedback controls $u(\cdot)$. Then there is an auxiliary function $V(\cdot)$ such that the dynamic programming equation is

$$V(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_y p(x, y|\alpha)V(y) + C(x, \alpha) - \bar{\gamma} \right]. \quad (1.5)$$

Theorem 1.1. Assume either condition (i), (ii), or (iii) below.

(i): \mathcal{U} has only a finite number of points. For each pair of states x, y there is an integer k and a feedback control $u(\cdot)$ such that

$$P\{\xi_k = y | \xi_0 = x, u(\cdot) \text{ used}\} > 0.$$

(ii): (A1.2) holds and S is a single recurrent and aperiodic class for each feedback control.

(iii): (A1.1) and (A1.2) hold.

Then, there exists a unique solution $\bar{\gamma}$ to (1.5).

The proof under (i) is [10, Proposition 4, Chapter 7]. The proof under (iii) is Theorem 8.12 in [100]. The proof under the stronger condition (ii) is in Theorem 3.1 below.

Theorem 1.2. Assume (A1.2) and suppose that there is a solution $(V, \tilde{\gamma})$ to (1.6).

$$V(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_y p(x, y | \alpha) V(y) + C(x, \alpha) - \tilde{\gamma} \right] \quad (1.6)$$

Then there is a feedback control $\bar{u}(\cdot)$ such that $\tilde{\gamma} = \gamma(\bar{u})$. Let $u = \{u_0(\cdot), u_1(\cdot), \dots\}$ be any sequence of feedback controls. Then

$$\tilde{\gamma} \leq \liminf_n \frac{1}{n} E_x^u \sum_{i=0}^{n-1} C(\xi_i, u_i(\xi_i)).$$

The result remains true for any admissible control sequence. Thus $\tilde{\gamma} = \bar{\gamma}$.

Proof. Suppose that there are V and $\tilde{\gamma}$ satisfying (1.6). Let the minimum be taken on at $\bar{u}(x)$ and let $u(\cdot)$ be another feedback control. Then

$$\begin{aligned} V(x) &= \sum_y p(x, y | \bar{u}(x)) V(y) + C(x, \bar{u}(x)) - \tilde{\gamma} \\ &\leq \sum_y p(x, y | u(x)) V(y) + C(x, u(x)) - \tilde{\gamma}. \end{aligned} \quad (1.6')$$

Define the $|S|$ -dimensional column vector $\mathbf{e} = (1, \dots, 1)$. By iterating, we get

$$V = P^n(\bar{u})V + \sum_{i=0}^{n-1} P^i(\bar{u})[C(\bar{u}) - \mathbf{e}\tilde{\gamma}].$$

Divide all terms by n and let $n \rightarrow \infty$ to get

$$\tilde{\gamma} = \lim_n \frac{1}{n} E_x^{\bar{u}} \sum_{i=0}^{n-1} C(\xi_i, \bar{u}(\xi_i)) = \gamma(\bar{u}).$$

Now, let $u = \{u_0(\cdot), u_1(\cdot), \dots\}$ be any sequence of feedback controls. Then, iterating (1.6') yields

$$V \leq P(u_0)P(u_1)\dots P(u_n)V + \sum_{i=0}^n P(u_0)\dots P(u_{i-1})[C(u_i) - \mathbf{e}\gamma].$$

Divide both sides by n , and let $n \rightarrow \infty$ to get

$$\tilde{\gamma} \leq \liminf_n \frac{1}{n} \sum_{i=0}^{n-1} E_x^u C(\xi_i, u_i(\xi_i)),$$

which proves the optimality with respect to all sequences of feedback controls. The additional details needed to get the result for an arbitrary admissible control sequence will not be given. See [74, Section 6.6], or [103].

■

A Centered Form of the Dynamical Equations. There are alternative “centered” forms of (1.2) and (1.6) which will be useful in the next few sections. Suppose that, for a feedback control $u(\cdot)$, we normalize the function $W(u)$ in (1.2) by choosing a particular state x_0 and setting $W(x_0, u) = \gamma(u)$. Define the $|S|$ -dimensional column vector $C(u) = \{C(x, u(x)), x \in S\}$. Then we can write (1.2) as

$$\begin{aligned} W(u) &= P(u)w(u) + C(u), \\ w(u) &= W(u) - \mathbf{e}W(x_0, u). \end{aligned} \tag{1.7}$$

Because $P(u)\mathbf{e} = \mathbf{e}$, we see that we must have $\gamma(u) = W(x_0, u)$. Analogously, the dynamic programming equation (1.6) can be written as

$$\begin{aligned} V &= \min_{u(x) \in \mathcal{U}} [P(u)v + C(u)], \\ v &= V - \mathbf{e}V(x_0). \end{aligned} \tag{1.8}$$

In this form, $\bar{\gamma} = V(x_0)$. In (1.8), the minimum is taken line by line, as usual.

7.2 A Jacobi Type Iteration

One way of looking at the average cost per unit time problem is as a control problem over a very large but finite time interval. Consider a control problem over the time interval $[0, N]$ with cost function

$$W(x, n, u) = E_x^u \sum_{i=0}^n C(\xi_i, u(\xi_i)). \tag{2.1}$$

The associated dynamic programming equation is (see Section 2.5)

$$V(x, n+1) = \min_{\alpha \in \mathcal{U}} \left[\sum_y p(x, y|\alpha) V(y, n) + C(x, \alpha) \right], \quad (2.2)$$

where $V(x, 0) = 0$. For large values of N , the cost function (2.1) differs from (1.1) mainly in the normalization by N . Thus it is intuitively reasonable to expect that the optimal control for the ergodic problem will be well approximated by that for the finite time problem if N is large, and that the normalized cost function $V(x, N)/N$ will converge to $\bar{\gamma}$ as $N \rightarrow \infty$. Under appropriate conditions this does happen, and we now give a more precise statement of the result.

Let $\{u_n(\cdot), n < \infty\}$ be a sequence of admissible feedback controls, where $u_i(\cdot)$ is to be used at time i . Define the $(n+1)$ -step probability transition matrix

$$\{p^n(x, y|u_0, \dots, u_n), x, y \in S\} = P(u_0) \cdots P(u_n).$$

Assume

A2.1. *There is a state x_0 , an $n_0 < \infty$, and an $\epsilon_0 > 0$ such that for all $u_0(\cdot), \dots, u_{n_0}(\cdot)$ and all $x \in S$, we have*

$$P\{\xi_{n_0} = x_0 | \xi_0 = x, u_0, \dots, u_{n_0-1} \text{ used}\} > \epsilon_0.$$

Define vectors V_n and v_n recursively by $V_0(x) = v_0(x) = 0$, and

$$\begin{aligned} V_{n+1} &= \min_{u(x) \in \mathcal{U}} [P(u)v_n + C(u)], \\ v_n &= V_n - \mathbf{e}V_n(x_0). \end{aligned} \quad (2.3)$$

Note the similarity between (2.2) and (2.3). Equation (2.3) is essentially (2.2) with a centering or normalization at each time step in order to keep the cost from blowing up.

Theorem 2.1. *Assume (A1.2) and either of (i)-(ii) below.*

- (i): (A2.1) holds.
- (ii): (A1.1) holds for every optimal feedback control. For some optimal feedback control, the chain is also aperiodic.

Then $V_n(x_0)$ converges to $\bar{\gamma}$ and the dynamic programming equation (1.6) holds.

Remarks. The result under condition (i) is due to White ([10],[74, Section 6.5], [100, Theorem 8.18], [124]) and that under (ii) is in [43]. See the survey [100] for additional information and references. Note that the algorithm defined by (2.3) is a Jacobi type iteration. There does not seem to be a Gauss-Seidel version known at this time, although there is for the centered form introduced below. Because the approximation in policy space type methods seem to be preferable for most problems, (2.3) is not widely used.

7.3 Approximation in Policy Space

The basic approximation in policy space algorithm is the following. Suppose that a feedback control $u_n(\cdot)$ is given, and let $(W(u_n), \gamma(u_n))$ denote the corresponding solution to (1.2) or, equivalently, (1.7). The algorithm defines the sequence of feedback controls $u_{n+1}(\cdot)$ by (6.2.1); that is, by

$$u_{n+1}(x) = \arg \min_{\alpha \in \mathcal{U}} \left[\sum_y p(x, y | \alpha) W(y, u_n) + C(x, \alpha) \right] \quad (3.1)$$

or, alternatively, by

$$u_{n+1}(x) = \arg \min_{\alpha \in \mathcal{U}} \left[\sum_y p(x, y | \alpha) w(y, u_n) + C(x, \alpha) \right]. \quad (3.2)$$

We can now state the following theorem. Other results are in the references.

Theorem 3.1. *Assume (A1.2), and that S is a single recurrent and aperiodic class under each feedback control. Then there is $\tilde{\gamma}$ such that $\gamma(u_n) \downarrow \tilde{\gamma}$, and there is a V such that $(V, \tilde{\gamma})$ satisfy (1.6), or equivalently, (1.8).*

Proof. We show that (1.6) holds. By the minimization in (3.1) and an argument like that below (1.6), it follows that $\gamma(u_{n+1}) \leq \gamma(u_n)$. Hence there is some $\tilde{\gamma}$ such that $\gamma(u_n) \downarrow \tilde{\gamma}$.

We next show that the associated sequence $\{W(u_n), n < \infty\}$ can be assumed to be bounded. Because any solution $W(u)$ will have the form (1.3) modulo a vector with identical components, we can suppose that the $W(u)$ are given by (1.3). Note that for each feedback $u(\cdot)$, the absolute value of the subdominant eigenvalue of $P(u)$ is bounded away from unity. This assertion follows from the following argument: First, note that the hypotheses imply that there is only one eigenvalue whose norm equals one (which is the eigenvalue of value one). Suppose that for some sequence $\{\hat{u}_n(\cdot)\}$ the absolute values of the sequence of subdominant eigenvalues of $P(\hat{u}_n)$ converges to one. We can suppose (by choosing a subsequence if necessary) that $\hat{u}_n(\cdot)$ converges to a control $\hat{u}(\cdot)$. Then $P(\hat{u})$ has at least two eigenvalues whose norms equal one, a contradiction to the condition that there is a single recurrent class under $\hat{u}(\cdot)$. It follows from the assertion that the $W(u)$ defined by (1.3) are bounded uniformly in $u(\cdot)$.

Let $n_k \rightarrow \infty$ be a sequence such that $u_{n_k}(\cdot), u_{n_k+1}(\cdot), W(u_{n_k})$, respectively, converge to limits $\tilde{u}(\cdot), \hat{u}(\cdot), \tilde{W}$, respectively. We have

$$\mathbf{e}\gamma(u_{n_k}) = (P(u_{n_k}) - I)W(u_{n_k}) + C(u_{n_k})$$

and

$$\mathbf{e}\tilde{\gamma} = (P(\tilde{u}) - I)\tilde{W} + C(\tilde{u}),$$

which implies that $\tilde{\gamma} = \gamma(\tilde{u})$. On the other hand, by (3.1)

$$\begin{aligned}\mathbf{e}\gamma(u_{n_k}) &\geq \min_{u(x) \in \mathcal{U}} [(P(u) - I)W(u_{n_k}) + C(u)] \\ &= [(P(u_{n_k+1}) - I)W(u_{n_k}) + C(u_{n_k+1})].\end{aligned}$$

Thus

$$\begin{aligned}\mathbf{e}\tilde{\gamma} &\geq \min_{u(x) \in \mathcal{U}} [(P(u) - I)\tilde{W} + C(u)] \\ &= [(P(\hat{u}) - I)\tilde{W} + C(\hat{u})].\end{aligned}. \quad (*)$$

Because there are no transient states by hypothesis, $\pi(x, \hat{u}) > 0$ for all $x \in S$. Now multiply the left and right sides by $\pi(\hat{u})$ to get $\tilde{\gamma} = \gamma(\tilde{u}) \geq \gamma(\hat{u})$, where the inequality is strict unless $(*)$ is an equality. But $\gamma(\hat{u}) = \gamma(\tilde{u}) = \lim_n \gamma(u_n)$, which implies that $(*)$ is an equality. ■

Remarks. Although it is not part of the theorem statement, it is a common experience that the sequence of controls $u_n(\cdot)$ also converges. One generally uses approximations to the solutions of (1.2) or (1.7) for each control $u_n(\cdot)$. Recall that the matrix $P(u)$ has an eigenvalue of value unity. Thus it is not a contraction, and the usual Jacobi or Gauss-Seidel iterative methods cannot be used directly. The method of Theorem 2.1 involved a centering or renormalization on each step, and that centering was crucial to the convergence. There is a modified form of the centering in (1.7) which will allow the use of the Jacobi, Gauss-Seidel, and similar methods, and which will be described next. The centered equations will be the basis of the numerical algorithms to be discussed in Section 7.4.

A “Centered” Form of (1.7). Fix $x_0 \in S$ and let $u(\cdot)$ be a feedback control. Let $P(x_0, u)$ denote the row vector which is the x_0^{th} row of $P(u)$. As in (1.7), we center $W(u)$ by setting $\gamma(u) = W(x_0, u)$. Define the centered transition matrix $P_e(u)$ and cost vector $C_e(u)$ by

$$P_e(u) = P(u) - \mathbf{e}P(x_0, u) = \{p_e(x, y|u(x)), x, y \in S\},$$

$$C_e(u) = C(u) - \mathbf{e}C(x_0, u(x_0)).$$

Equation (1.7) implies that

$$\begin{aligned}w(u) + \mathbf{e}W(x_0, u) &= P(u)w(u) + C(u), \\ W(x_0, u) &= P(x_0, u)w(u) + C(x_0, u(x_0)).\end{aligned} \quad (3.3)$$

Using these expressions, we can rewrite (1.7) as

$$w(u) = P_e(u)w(u) + C_e(u). \quad (3.4)$$

Note that the x_0^{th} rows of $P_e(u)$ and $C_e(u)$ are zero. Given the value of $w(u)$, both $W(u)$ and $\gamma(u)$ can be calculated. The key property of the

matrix $P_e(u)$ is the fact that its spectral radius is less than unity, so that the iterative algorithms of Chapter 6 can be used for the solution of (3.4). This centering idea seems to have originated in [11] in their work on aggregation algorithms for the ergodic problem.

Lemma 3.2. *If the chain under the feedback control $u(\cdot)$ contains only transient states and a single aperiodic recurrent class, then the spectral radius of $P_e(u)$ is less than unity.*

Proof. Note that $P(x_0, u)\mathbf{e} = 1$ and $P(u)\mathbf{e} = \mathbf{e}$. Thus,

$$P_e^2(u) = P_e(u)P(u) - [P(u) - \mathbf{e}P(x_0, u)]\mathbf{e}P(x_0, u) = P_e(u)P(u).$$

Thus, for $n > 1$, $P_e^n(u) = P_e(u)P^{n-1}(u)$. Let $\{\lambda_i, v_i\}$ denote the eigenvalues and associated eigenvectors of $P(u)$, where we order them such that $\lambda_1 = 1$ and $v_1 = \mathbf{e}$. For a given vector x , define the coefficients c_i by $x = \sum_i c_i v_i$. Then

$$P_e^n(u)x = P_e(u)P^{n-1}(u)x = \sum_i P_e(u)c_i \lambda_i^{n-1} v_i.$$

Because $v_1 = \mathbf{e}$ and $P_e(u)\mathbf{e} = 0$, we see that the dominant eigenvalue of $P_e(u)$ is the subdominant eigenvalue of $P(u)$. ■

7.4 Numerical Methods for the Solution of (3.4)

In this section, we comment briefly on the use of the Jacobi (6.2.7), Gauss-Seidel (6.2.10), semi-accelerated Gauss-Seidel (6.4.2), the accelerated Gauss-Seidel (6.4.3), and the other methods of Chapter 6. Assumptions (A1.1) and (A1.2) will continue to be used.

Given $w_0(u)$, the Jacobi iteration for solving (3.4) is

$$w_{n+1}(x, u) = \sum_y p_e(x, y|u)w_n(y, u) + C_e(x, u). \quad (4.1)$$

The Gauss-Seidel iteration for solving (3.4) is

$$w_{n+1}(x, u) = \sum_{y < x} p_e(x, y|u)w_{n+1}(y, u) + \sum_{y \geq x} p_e(x, y|u)w_n(y, u) + C_e(x, u), \quad (4.2)$$

where it is supposed that the states are ordered in some way. Recall that $\gamma(u) = W(x_0, u)$. Even if the transition probabilities for the approximating Markov chain are selected such that $p(x, x|u(x)) = 0$ for all x (see the discussion concerning normalization in Subsection 5.2.2), due to the centering we will not have $p_e(x, x|u) = 0$ for any state x for which $p(x_0, x|u(x)) > 0$.

From a numerical point of view, it is preferable to “normalize” the equations (4.1) and (4.2) by using the forms

$$w_{n+1}(x, u) = \left[\sum_{y \neq x} p_e(x, y|u) w_n(y, u) + C_e(x, u) \right] / (1 - p_e(x, x|u)), \quad (4.3)$$

$$\begin{aligned} w_{n+1}(x, u) &= \left[\sum_{y < x} p_e(x, y|u) w_{n+1}(y, u) + \sum_{y > x} p_e(x, y|u) w_n(y, u) \right. \\ &\quad \left. + C_e(x, u) \right] / (1 - p_e(x, x|u)). \end{aligned} \quad (4.4)$$

Similar remarks hold for the various “accelerated” forms of these algorithms.

Because the matrix $P_e(u)$ depends on the chosen centering state x_0 , it is conceivable that the eigenvalues will depend on x_0 . Lemma 3.2 implies that this is not the case for the Jacobi method. But the eigenvalues of the matrices which define the Gauss-Seidel type iterations do depend on the chosen state, and this dependence can be quite important. The type of dependence which seems to be typical is presented in Tables 4.1 and 4.2 below for one particular case, which is an approximating chain for a diffusion process. Define the one dimensional system

$$dx = cxdt + dw, \quad c > 0$$

on the interval $[0, 1]$, with instantaneously reflecting boundaries. The state space for the Markov chain approximation is $\{h, 2h, \dots, 1-h\}$, where $1/h$ is assumed to be an integer. Let $ch \leq 1$, and define $x_0 = n_0 h$ for some integer n_0 . Define the transition probabilities $p^h(x, y)$ by the “central finite difference formula” (see Section 5.1)

$$p(x, x+h) = \frac{1 + chx}{2}, \quad x = h, \dots, 1-2h,$$

$$p(x, x-h) = \frac{1 - chx}{2}, \quad x = 2h, \dots, 1-h,$$

$$p(h, h) = 1 - p(h, 2h), \quad p(1-h, 1-h) = 1 - p(1-h, 1-2h).$$

In the tables, $h = 1/20$. The first three lines of Table 4.1 compare the spectral radius of the iteration matrix with an acceleration parameter $\omega = 1.2$ and x_0 equalling h (the left hand state), $10h$ (the center state) and $19h$ (the right hand state). The advantages of the AGS over the SAGS over the GS over the Jacobi are clear. Note the improvement due to the use of the normalization (although the effects of the normalization will be less in higher dimensions). The use of the left hand states for the centering state is preferable. This seems to be due to the fact that the Gauss-Seidel method

starts iterating from the left. A comparison of the first three and the second three lines in the table and with the unaccelerated GS column show the advantages of acceleration. Of course, if the acceleration parameter is too large, the iteration becomes unstable. The choice of the centering state is important in general. It should be a state which has a “high connectivity” from other states.

The multigrid method is used exactly as in Chapter 6. The aggregation method is also used exactly as in Chapter 6, with the sole exception that the logic of the method requires that the state x_0 must be an “aggregation group” by itself. See the discussion in [11] for additional information on this point.

n_0	ω	J	GS	GS(norm)	SAGS(norm)	AGS(norm)
1	1.2	.987	.975	.963	.956	.920
10	1.2	.987	.977	.971	.966	.965
19	1.2	.987	.978	.974	.969	.971
1	1.4	.987	.975	.963	.949	.909
10	1.4	.987	.977	.971	.960	.953
19	1.4	.987	.978	.974	.964	.963

Table 4.1. Spectral radius, $c = 1$.

n_0	ω	J	GS	GS(norm)	SAGS(norm)	AGS(norm)
1	1.2	.986	.973	.965	.958	.939
10	1.2	.986	.976	.970	.964	.964
19	1.2	.986	.977	.973	.967	.971

Table 4.2. Spectral radius, $c = 0$.

7.5 The Control Problem for the Approximating Markov Chain

We will adapt the discussion concerning the problem formulation and numerical algorithms in Sections 7.1 to 7.4 to the approximating Markov chain. First, the dynamical equation for the value function for a fixed control, and the dynamic programming equation for the optimal value will be given. Assumptions (A1.1) and (A1.2) will be used for the approximating chains for the values of h of interest. In this section, we let the interpolation interval $\Delta t^h(x, \alpha)$ depend on the state and control in order to get a better understanding of the relationship between the approximating chain and the original diffusion process. It will be seen later that the approximation in policy space algorithm requires that the intervals not depend on either the

state or control. In Section 7.7, we will modify the transition probabilities so that they are not state dependent.

We suppose the setup for a reflecting jump diffusion given in Subsection 5.8.3. The process $x(\cdot)$ is constrained to stay in a compact set G . G_h denotes the set of states for the approximating chain in G , and ∂G_h^+ denotes the reflecting boundary for the chain, which is disjoint from G_h . Define the state space $S_h = G_h \cup \partial G_h^+$. Recall that $\Delta t^h(x, \alpha) = 0$, for $x \in \partial G_h^+$. By (A1.1), there cannot be an absorbing boundary unless it is a single point. Recall the notation: $u_n^h = u(\xi_n^h)$, $\Delta t_n^h = \Delta t^h(\xi_n^h, u_n^h)$, and $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$.

It will be seen below that the appropriate cost function for the Markov chain, under an admissible control sequence $u = \{u_0(\cdot), u_1(\cdot), \dots\}$, is

$$\limsup_n \frac{E_x^u \sum_0^n k(\xi_i^h, u_i^h) \Delta t_i^h}{E_x^u \sum_0^n \Delta t_i^h}. \quad (5.1)$$

Boundary costs can be added, and an example will be given in Chapter 8; also, some remarks appear in Section 7.8. Here, we wish to keep the formulation simple. If (5.1) does not depend on the initial condition, then we write it as $\gamma^h(u)$. Suppose that the feedback control $u(\cdot)$ is used at each time step and let $\pi^h(u) = \{\pi^h(x, u), x \in S_h\}$ denote the associated invariant measure. The ergodic theorem for Markov chains [20] yields

$$\gamma^h(u) = \lim_n \frac{E_x^u \sum_0^n k(\xi_i^h, u_i^h) \Delta t_i^h}{E_x^u \sum_0^n \Delta t_i^h} = \frac{\sum_x k(x, u(x)) \Delta t^h(x, u(x)) \pi^h(x, u)}{\sum_x \Delta t^h(x, u(x)) \pi^h(x, u)}. \quad (5.2)$$

There is a vector valued function $W^h(u)$ with values $\{W^h(x, u), x \in S_h\}$, such that $(W^h(x, u), \gamma^h(u))$ satisfy

$$W^h(x, u) = \sum_y p^h(x, y|u(x)) W^h(y, u) + [k(x, u(x)) - \gamma^h(u)] \Delta t^h(x, u(x)). \quad (5.3)$$

For $x \in \partial G_h^+$, (5.3) reduces to

$$W^h(x, u) = \sum_y p^h(x, y|u(x)) W^h(y, u). \quad (5.4)$$

In fact, analogous to the case in Section 7.1, the function defined by

$$W^h(x, u) = \sum_{n=0}^{\infty} E_x^u [k(\xi_n^h, u_n^h) - \gamma^h(u)] \Delta t_n^h \quad (5.5)$$

satisfies (5.3). As was the case in Section 7.1, the solution to (5.3) is unique up to an additive constant on the function $W^h(u)$. Let $(\tilde{W}^h, \tilde{\gamma}^h)$ satisfy

$$\tilde{W}^h(x) = \sum_y p^h(x, y|u(x)) \tilde{W}^h(y) + [k(x, u(x)) - \tilde{\gamma}^h] \Delta t^h(x, u(x)). \quad (5.6)$$

Then, following the argument used below (1.4), $\tilde{\gamma}^h = \gamma^h(u)$.

We next show why the cost function (5.1) or (5.2) is a natural choice for the approximating Markov chain model. Fix the feedback control $u(\cdot)$. So far, we have been working with the invariant measure of the approximating chain. The amount of time that the chain spends at any one state once it arrives there might depend on that state (and perhaps on the control used in that state as well). Thus, the invariant measure of the chain needs to be adjusted to account for this (possibly) state and control dependent sojourn time, if it is to be used to get the mean value of the cost per unit time for the continuous parameter interpolated process $\psi^h(\cdot)$. That is, because $\Delta t^h(x, u(x))$ depends (or might depend) on the state x , and on the control used at that state, the invariant measure for the chain needs to be “weighted” in order for it to account for the time that the interpolated process (which is the process of primary interest) spends at each state x . Define the measure $\mu^h(u) = \{\mu^h(x, u), x \in S_h\}$ by

$$\mu^h(x, u) = \frac{\Delta t^h(x, u(x))\pi^h(x, u)}{\sum_y \Delta t^h(y, u(y))\pi^h(y, u)}. \quad (5.7)$$

Note that the value is zero for the “instantaneous” reflecting states. Now, (5.2) can be written in the simpler form

$$\gamma^h(u) = \sum_x k(x, u(x))\mu^h(x, u). \quad (5.8)$$

Recall the definition of the interpolation $\xi^h(\cdot)$ from Chapter 4, and define the interpolation $u^h(\cdot)$ here by: $u^h(t) = u_n^h$ for $t \in [t_n^h, t_{n+1}^h]$. The ergodic theorem for Markov chains also implies the following identities (all with probability one)

$$\begin{aligned} \gamma^h(u) &= \lim_n \sum_0^{n-1} k(\xi_i^h, u_i^h) \Delta t_i^h / t_n^h = \lim_n \frac{1}{t_n^h} \int_0^{t_n^h} k(\xi^h(s), u(\xi^h(s))) ds \\ &= \lim_t \frac{1}{t} \int_0^t k(\xi^h(s), u(\xi^h(s))) ds \\ &= \lim_t \frac{1}{t} \int_0^t E_x^u k(\xi^h(s), u(\xi^h(s))) ds. \end{aligned} \quad (5.9)$$

A similar expression can be obtained for $\psi^h(\cdot)$.

The expressions (5.9) make it clear that the cost function (5.1) or (5.2) is the correct one for the approximating Markov chain model. The scaling by the mean interpolation interval [in the denominators of (5.1) and (5.2)] accounts for the (possibly) state and control dependent time that the process spends at each state until the next transition. It also transforms [as in (5.7)] the invariant measure for the chain into one which measures the relative time that the interpolated process spends at each state of the

interpolated process. It will be seen in Section 7.6 that $\mu^h(u)$ is just the invariant measure for the continuous parameter Markov chain interpolation $\psi^h(\cdot)$, which was defined in Chapter 4. Let us extend the measure $\mu^h(u)$ to a measure on the Borel subsets of the “continuous” state space G of the original reflected process $x(\cdot)$ by defining $\mu^h(A, u) = \sum_{x \in A} \mu^h(x, u)$. In fact, it will be shown in Chapter 11 that this extended $\mu^h(u)$ is an approximation to an invariant measure of $x(\cdot)$.

Generally, it is much harder to calculate the invariant measure $\mu^h(u)$ than it is to calculate the values of explicit functionals such as

$$\sum_x k(x, u(x)) \mu^h(x, u).$$

The numerical methods for the latter calculation converge much faster than do the numerical methods which might be used for the calculation of the invariant measure itself. Because of this, even if one is interested in the invariant measure of $x(\cdot)$, we might content ourselves with the values of approximations to “stationary expectations” of a small number of functions.

An interesting alternative for numerically approximating the invariant measure for certain classes of “heavy traffic” problems is in [30,54].

The Dynamic Programming Equation. Define $\bar{\gamma}^h = \inf_u \gamma^h(u)$, where the infimum is over all feedback controls $u(\cdot)$. Then the dynamic programming equation is

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_y p^h(x, y | \alpha) V^h(y) + [k(x, \alpha) - \bar{\gamma}^h] \Delta t^h(x, \alpha) \right]. \quad (5.10)$$

The fact that the infimum $\bar{\gamma}^h$ of the costs (together with an auxiliary function V^h) satisfies (5.10) as well as the convergence of the approximation in policy space algorithm can be shown by the methods used in Theorems 1.2 and 3.1. Here we will show only that (5.10) is the correct dynamic programming equation. Suppose that there are V^h and $\tilde{\gamma}^h$ satisfying

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \left[\sum_y p^h(x, y | \alpha) V^h(y) + [k(x, \alpha) - \tilde{\gamma}^h] \Delta t^h(x, \alpha) \right] \quad (5.11)$$

for all $x \in S_h$. Let the minimum be taken on at $\bar{u}(x)$ and let $u(\cdot)$ be another feedback control. Then we claim that $\tilde{\gamma}^h = \gamma^h(\bar{u})$ which equals the minimum cost. Following the method used to deal with (1.6),

$$V^h(x) = \left[\sum_y p^h(x, y | \bar{u}(x)) V^h(y) + [k(x, \bar{u}(x)) - \tilde{\gamma}^h] \Delta t^h(x, \bar{u}(x)) \right]$$

$$\leq \left[\sum_y p^h(x, y|u(x)) V^h(y) + [k(x, u(x)) - \tilde{\gamma}^h] \Delta t^h(x, u(x)) \right].$$

The first equality implies that $\tilde{\gamma}^h = \gamma^h(\bar{u})$, as shown in connection with (1.4). Next, multiply the left and right sides of the above inequality by $\pi^h(x, u)$, sum over x , and use the fact that $\pi^h(u)$ is invariant for the transition matrix $P^h(u)$ to get that

$$\sum_x \pi^h(x, u) k(x, u(x)) \Delta t^h(x, u(x)) \geq \tilde{\gamma}^h \sum_x \pi^h(x, u) \Delta t^h(x, u(x)).$$

But this last inequality implies that $\gamma^h(u) \geq \tilde{\gamma}^h = \gamma^h(\bar{u})$, which proves the claim.

Computation of $\gamma^h(u)$. Let $u(\cdot)$ be a feedback control and assume (A1.2). Choose a centering state x_0 , and write (analogously to (3.4)) for $x \in G_h \cup \partial G_h^+$

$$\begin{aligned} C^h(x, u(x)) &= k(x, u(x)) \Delta t^h(x, u(x)), \\ P_e^h(u) &= P^h(u) - \mathbf{e} P^h(x_0, u), \\ C_e^h(u) &= C^h(u) - \mathbf{e} C^h(x_0, u(x_0)), \\ w^h(u) &= P_e^h(u) w^h(u) + C_e^h(u). \end{aligned} \quad (5.12)$$

Then

$$\begin{aligned} W^h(x_0, u) &= P^h(x_0, u) w^h(u) + C^h(x_0, u(x_0)) \\ &= \sum_x k(x, u(x)) \Delta t^h(x, u(x)) \pi^h(x, u). \end{aligned} \quad (5.14)$$

Now repeat (5.12) for $C^h(x, u(x)) = \Delta t^h(x, u(x))$ yielding the new value

$$W^h(x_0, u) = \sum_x \Delta t^h(x, u(x)) \pi^h(x, u). \quad (5.15)$$

Divide (5.14) by (5.15) to get $\gamma^h(u)$. So we can calculate the cost under each control $u(\cdot)$. Nevertheless, to use the approximation in policy space method, we need to use a constant interpolation interval, for $x \in G_h$. See Section 7.7.

7.6 The Continuous Parameter Markov Chain Interpolation

Recall the continuous parameter approximating Markov chain $\psi^h(\cdot)$ introduced in Section 4.3. Let $u(\cdot)$ be a feedback control (as usual, it does not depend on time). By the law of large numbers for continuous parameter

Markov chains [20], one can show that the $\mu^h(u)$ defined by (5.7) is an invariant measure for $\psi^h(\cdot)$, if $\pi^h(u)$ is an invariant measure for $\{\xi_n^h, n < \infty\}$. In order to prove this assertion, first recall the definitions in Section 4.3 of τ_n^h , the times of change of $\psi^h(\cdot)$, and of the differences $\Delta\tau_n^h = \tau_{n+1}^h - \tau_n^h$. Recall that the mean holding time in state x for $\psi^h(\cdot)$ is $\Delta t^h(x, u(x))$. The *mean amount of time* that $\psi^h(\cdot)$ takes the value x is (where the limits are with probability one)

$$\lim_n \frac{\sum_{i=0}^{n-1} \Delta\tau_i^h I_{\{\xi_i^h=x\}}}{\tau_n^h} = \frac{\Delta t^h(x, u(x))\pi^h(x, u)}{\sum_y \Delta t^h(y, u(y))\pi^h(y, u)} \quad (6.1)$$

which is just $\mu^h(x, u)$. Hence, $\mu^h(\cdot)$ is an invariant measure for $\psi^h(\cdot)$ under control $u(\cdot)$. It is unique if there is a unique invariant measure for $\{\xi_n^h, n < \infty\}$ under $u(\cdot)$.

The mean cost per unit time for $\psi^h(\cdot)$ under $u(\cdot)$ can be written as

$$\lim_t E_x^u \frac{1}{t} \int_0^t k(\psi^h(s), u(\psi^h(s)))ds = \int k(x, u(x))\mu^h(dx, u) = \gamma^h(u). \quad (6.2)$$

We can also write

$$\gamma^h(u) = E_{\mu^h(u)}^u \int_0^1 k(\psi^h(s), u(\psi^h(s)))ds, \quad (6.3)$$

where $E_{\mu^h(u)}^u$ is the expectation for the stationary process, under the control $u(\cdot)$. Again, we note that boundary costs can be added. See Section 7.8 and Chapter 8. Let us note the following without proof, because it partially justifies considering the process $\psi^h(\cdot)$: The equivalence of (6.2) and (6.3) and a weak convergence argument can be used to show that the weak limits of the invariant measures $\mu^h(u)$ are invariant measures for the original process $x(\cdot)$. See Chapter 11, where convergence results for the ergodic cost problem will be given.

The interpolation $\psi^h(\cdot)$ is useful for the mathematical convergence analysis. But the computational problem is the same as that for the discrete parameter chain. The recursive equations for the average cost per unit time under any given feedback control $u(\cdot)$ are the same for the discrete and the continuous parameter chains. By comparing (6.2) and (6.3) with (5.1) and (5.2), it is evident that the minimum of (5.1) for the discrete parameter chain equals the minimum of

$$\limsup_t E_x^u \frac{1}{t} \int_0^t k(\psi^h(s), u(s))ds$$

over the admissible controls.

7.7 Computations for the Approximating Markov Chain

There are two minor issues that need to be dealt with in order to get a centered equation (with the desired contraction properties) for the cost, analogous to (3.4). The first problem concerns the possible state and control dependence of the interpolation intervals. The second issue concerns the fact that the $\gamma^h(u)$ in (5.3) is multiplied by the interpolation interval. With these problems taken care of, we will get a centered equation of the form (3.4) and the numerical methods of Chapter 6 can be used.

7.7.1 CONSTANT INTERPOLATION INTERVALS

The Necessity of a Constant Interpolation Interval. A representation of the mean cost per unit time by a formula analogous to (3.4) as well as the algorithms in Section 7.4 cannot be used if the interpolation intervals are dependent on the state or control, because then the e would need to be replaced with a vector with components $\Delta t^h(x, u(x))$.

Furthermore, if the interpolation intervals are not independent of the state and the control for each value of h of interest, then the solution $\gamma^h(u)$ to (5.3) is the ratio of expectations given by (5.2), which suggests that it would be difficult to solve (5.3) by a “dynamical procedure” say, akin to the Gauss-Seidel method. Indeed, it was shown at the end of Section 7.5 that the mean cost could be calculated, but by solving two cost equations. In order to solve for the optimal ergodic cost for the approximating chain by the computational methods of Chapter 6, it seems to be necessary at this time to have a constant interpolation interval $\Delta t^h(x, \alpha) \equiv \Delta t^h$, for $x \in G_h$.

For purposes of completeness, we will next discuss an example of the construction of a locally consistent approximating chain whose interpolation intervals do not depend on either the control or the state. The example to follow concerns the approximation in the set G , because the reflection states are instantaneous and their interpolation intervals are zero, they will be (easily) dealt with separately in Section 7.7.2 below.

Construction of an Approximating Chain with a Constant Interpolation Interval. Example. It was seen in Section 5.2 how to construct approximating chains with interpolation intervals which do not depend on the state and control. In order to refresh our memory and to firmly associate the procedure with the computational methods for the ergodic problem, we will repeat the details for one simple example here. For concreteness, we work with the system

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= (a_1 x_1 + a_2 x_2)dt + u dt + \sigma dw, \end{aligned} \tag{7.1}$$

where $\sigma^2 \geq 1, |u(t)| \leq 1$. Let $a_i \leq 0$, so that the system (7.1) is stable. For this illustrative example, the state space for $x(\cdot)$ will be the closed “square” $G = [-B, B] \times [-B, B]$, for some $B > 0$, and with a reflecting boundary. It is sufficient for our purposes to let G_h be the regular h -grid ($\mathbb{R}_h^r \cap G$) on G . Let ∂G_h^+ , the reflecting boundary for the chain, be the points in $\mathbb{R}_h^r - G$ which are at most a distance h from G in any coordinate direction. The procedure is as follows: First, we obtain any locally consistent chain, using any of the methods of Chapter 5. If the interpolation intervals are not constant, we then use the ideas of Section 5.2 to eliminate the state and control dependence.

For the first step for this case, and in order to illustrate one concrete example, let us follow the “operator splitting” procedure outlined in Sub-section 5.2.3. Write $x = (x_1, x_2)$. The differential operator of the process defined by (7.1) is

$$\frac{\sigma^2}{2} \frac{\partial^2}{\partial x_2^2} + x_2 \frac{\partial}{\partial x_1} + u(x) \frac{\partial}{\partial x_2} + (a_1 x_1 + a_2 x_2) \frac{\partial}{\partial x_2}. \quad (7.2)$$

Define

$$Q^h(x) = \sigma^2 + h|x_2| + h|a_1 x_1 + a_2 x_2|.$$

For use below, define the maximum of $Q^h(x, \alpha)$

$$Q^h = \max_{x \in G, \alpha \in U} Q^h(x, \alpha) = \sigma^2 + hB + hB|a_1 + a_2|.$$

The transition probabilities will be defined by the finite difference method. A central difference approximation (5.3.4) will be used for the second derivative term and also for the first partial derivative involving the control term $u(x)$ in (7.2), as in (5.1.18). A one sided difference approximation (5.3.5) is used for the other two first derivative terms in (7.2). This procedure yields the transition probabilities

$$p^h(x, x \pm e_1 h | \alpha) = \frac{n^h(x, x \pm e_1 h | \alpha)}{Q^h(x)},$$

where

$$n^h(x, x \pm e_1 h | \alpha) = h x_2^\pm,$$

and

$$p^h(x, x \pm e_2 h | \alpha) = \frac{n^h(x, x \pm e_2 h | \alpha)}{Q^h(x)},$$

where

$$n^h(x, x \pm e_2 h | \alpha) = \sigma^2/2 + h\alpha/2 + h(a_1 x_1 + a_2 x_2)^\pm,$$

$$\Delta t^h(x) = \frac{h^2}{Q^h(x)}.$$

The transition probabilities from x to all nonlisted y are zero. The above constructed transition probabilities are locally consistent with the diffusion model (7.1).

Next, let us construct transition probabilities where the interpolation intervals are constant. This is easily done as follows, using the $p^h(x, y|\alpha)$ given above. Define

$$\begin{aligned}\bar{p}^h(x, x \pm e_1 h|\alpha) &= \frac{n^h(x, x \pm e_1 h|\alpha)}{Q^h}, \\ \bar{p}^h(x, x \pm e_2 h|\alpha) &= \frac{n^h(x, x \pm e_2 h|\alpha)}{Q^h}, \\ \bar{p}^h(x, x|\alpha) &= 1 - \frac{Q^h(x)}{Q^h}, \\ \Delta\bar{t}^h(x) &= \Delta t^h = \frac{h^2}{Q^h}.\end{aligned}\tag{7.3}$$

The chain with these transition probabilities and interpolation interval is also locally consistent with the diffusion (7.1). The transition probabilities (7.3) are used for $x \in G_h$. The transition probabilities for the reflecting states ∂G_h^+ are assumed to be locally consistent with the reflection directions as in Section 5.7. Recall that the interpolation interval for states on ∂G_h^+ is zero.

7.7.2 THE EQUATION FOR THE COST (5.3) IN CENTERED FORM

Suppose that we are given a locally consistent chain with interpolation intervals constant for $x \in G_h$. This might be obtained as above from a locally consistent chain with a nonconstant interpolation interval for $x \in G_h$. Because $\Delta t^h(x) = 0$ for $x \in \partial G_h^+$, the interpolation intervals are not actually constant over all states. Because of this, we need to eliminate the reflection states. This is easy to do and the algebraic details will be given below. Once these reflecting states are eliminated, the analog of the centered form (3.4) for (5.3) can be used and solved with any of the numerical methods alluded to in Section 7.4. Let $\bar{p}^h(x, y|\alpha)$, $x \in G_h$, denote transition probabilities of a locally consistent Markov chain with constant interpolation intervals Δt^h for $x \in G_h$. Let $p^h(x, y)$ denote the transition probabilities for $x \in \partial G_h^+$. Then (7.8) below is the appropriate centered form of (5.3), which is analogous to (3.4).

Since the interpolation interval is zero for the reflecting states, we need to eliminate them in order to get an analog of the centered form (3.4) for the $\bar{p}^h(x, y|\alpha)$. Fix the centering state x_0 to be an ergodic state not on the boundary ∂G_h^+ . Suppose, without loss of generality, that $\bar{p}^h(x_0, y|\alpha) = 0$ for all states $y \in \partial G_h^+$ and control actions α . Also, to slightly simplify the

development, suppose that the states in ∂G_h^+ communicate only to states in G_h .

In order to eliminate the instantaneous reflecting boundary states, we define the transition probability for $x, y \in G_h$ as

$$\tilde{p}^h(x, y|\alpha) = \bar{p}^h(x, y|\alpha) + \sum_{z \in \partial G_h^+} \bar{p}^h(x, z|\alpha) p^h(z, y).$$

Now, for a feedback control $u(\cdot)$, define the transition matrix for the “reduced” chain on G_h by

$$\tilde{P}^h(u) = \{\tilde{p}^h(x, y|u(x)), x, y \in G_h\}.$$

Then (5.3) can be rewritten as

$$W^h(u) = \tilde{P}^h(u)W^h(u) + C^h(u) - \mathbf{e}\Delta t^h\gamma^h(u),$$

where

$$C^h(x, u(x)) = k(x, u(x))\Delta t^h, x \in G_h,$$

$$W^h(u) = \{W^h(x, u), x \in G_h\},$$

since we supposed that there is no cost on the boundary. Now follow the procedure which led to (3.4). Choose the centering value $W^h(x_0, u)$ to satisfy

$$\gamma^h(u) = \frac{W^h(x_0, u)}{\Delta t^h}. \quad (7.4)$$

Define the centered values

$$w^h(u) = W^h(u) - \mathbf{e}W^h(x_0, u).$$

Then

$$w^h(u) + \mathbf{e}W^h(x_0, u) = \tilde{P}^h(u)w^h(u) + C^h(u), \quad (7.5a)$$

$$W^h(x_0, u) = \tilde{P}^h(x_0, u)w^h(u) + k^h(x_0, u(x_0))\Delta t^h. \quad (7.5b)$$

Using the fact that $\tilde{p}^h(x_0, y|u(x_0)) = \bar{p}^h(x_0, y|u(x_0))$, we thus obtain the following form of (5.3)

$$\begin{aligned} w^h(x, u) &= \sum_{y \in G_h} [\bar{p}^h(x, y|u(x)) - \bar{p}^h(x_0, y|u(x_0))]w^h(y, u) \\ &\quad + [k(x, u(x)) - k(x_0, u(x_0))]\Delta t^h, \end{aligned} \quad (7.6)$$

for $x \in G_h$.

It is not necessary to calculate the $\tilde{p}^h(x, y|\alpha)$ in order to evaluate the right hand side. The calculation can be done in terms of the original $\bar{p}^h(x, y|\alpha)$. First, rewrite the sum in (7.6) as

$$\begin{aligned} &\sum_{y \in G_h} [\bar{p}^h(x, y|u(x)) - \bar{p}^h(x_0, y|u(x_0))]w^h(y, u) \\ &+ \sum_{y \in G_h} \sum_{z \in \partial G_h^+} \bar{p}^h(x, z|u(x))p^h(z, y)w^h(y, u). \end{aligned}$$

For $x \in \partial G_h^+$, define $w^h(x, u)$ by

$$w^h(x, u) = \sum_y p^h(x, y) w^h(y, u). \quad (7.7)$$

Then, for $x \in G_h$, we can write (7.6) as

$$\begin{aligned} w^h(x, u) &= \sum_y [\bar{p}^h(x, y|u(x)) - \bar{p}^h(x_0, y|u(x_0))] w^h(y, u) \\ &\quad + [k(x, u(x)) - k(x_0, u(x_0))] \Delta t^h. \end{aligned} \quad (7.8)$$

Equation (3.4) is thus replaced by (7.4), (7.7), (7.8).

The Algorithm in Terms of the Original Data. Before proceeding, let us backtrack a little. Often the $\bar{p}^h(x, y|\alpha)$, $x \in G_h$, are obtained by starting with a chain with nonconstant interpolation intervals, and then doing a transformation, as in the last subsection. Since it is useful to write the actual computational algorithm in terms of the original data, let us suppose next that we start with locally consistent transition probabilities $p^h(x, y|\alpha)$ for which the interpolation intervals (for $x \in G_h$) might *not* be constant, and then obtain locally consistent transition probabilities $\bar{p}^h(x, y|\alpha)$ with constant interpolation intervals from them as outlined in the above subsection. Below, it will be useful to represent the original $p^h(x, y|\alpha)$ in the common form

$$p^h(x, y|\alpha) = \frac{n^h(x, y|\alpha)}{Q^h(x, \alpha)}, \quad x \in G_h.$$

We will write the algorithm in terms of the $p^h(x, y|\alpha)$, since that is the starting data. The final result will be (7.11) and (7.12) below. Now, (7.8) will be rewritten in terms of $p^h(x, y|\alpha)$ to get (7.11) and (7.12).

In terms of the original transition probabilities $p^h(x, y|\alpha)$, we can rewrite (7.8) as

$$w^h(x, u) = \sum_{y \neq x} [n^h(x, y|u(x)) - n^h(x_0, y|u(x_0))] \frac{w^h(y, u)}{Q^h} + \quad (7.9)$$

$$\left[1 - \frac{Q^h(x, \alpha)}{Q^h} - \frac{n^h(x_0, x|u(x_0))}{Q^h} \right] w^h(x, u) + [k(x, u(x)) - k(x_0, u(x_0))] \frac{h^2}{Q^h},$$

where $Q^h = \sup_{x, \alpha} Q^h(x, \alpha)$. As will be seen below, it is not necessary to calculate Q^h .

The Normalized Equation. Recall the definition of the normalized equation in Section 5.2. The normalized form of (3.4) is

$$w(x, u) = \frac{\sum_{y \neq x} p_e(x, y|u) w(y, u) + C_e(x, u)}{1 - p_e(x, x|u)}. \quad (7.10)$$

It is generally preferred to use the normalized form in the Gauss-Seidel or related relaxations. Let $\bar{p}^h(x, y|\alpha)$ be the locally consistent transition probability with constant interpolation interval for $x \in G_h$ used above. Suppose that a procedure such as in the last subsection is used to get them from a locally consistent set $p^h(x, y|\alpha)$ for which the interpolation intervals are not constant in G_h . Then $\bar{p}^h(x, x|\alpha) > 0$ for some states $x \in G_h$. The normalized form of (7.7), (7.8) for this case will now be written in terms of the original $p^h(x, y|\alpha)$. Because one generally tries to construct the transition probabilities such that $p^h(x, x|\alpha) = 0$, let us assume this here. Using the normalization analogous to (7.10) and noting that

$$1 - \bar{p}_e^h(x, x|u) = 1 - [Q^h - Q^h(x, u(x)) - n^h(x_0, x|u(x_0))] / Q^h,$$

yields for $x \in G_h$,

$$\begin{aligned} w^h(x, u) &= \sum_{y \neq x} \frac{n^h(x, y|u(x)) - n^h(x_0, y|u(x_0))}{Q^h(x, u(x)) + n^h(x_0, x|u(x_0))} w^h(y, u) \\ &\quad + [k(x, u(x)) - k(x_0, u(x_0))] \frac{h^2}{Q^h(x, u(x)) + n^h(x_0, x|u(x_0))}. \end{aligned} \tag{7.11}$$

For the reflecting states,

$$w^h(x, u) = \sum_y p^h(x, y) w^h(y, u). \tag{7.12}$$

Recall that the actual average cost is given by (7.4).

Approximation in Policy Space. For the approximation in policy space method, given a current control $u(\cdot)$ and an associated approximation $w^h(u)$, the next control $u_1(\cdot)$ are the minimizers in

$$\min_{\alpha \in \mathcal{U}} \left[\sum_y \bar{p}^h(x, y|\alpha) w^h(y, u) + k(x, \alpha) \Delta t^h \right], \quad x \in G_h. \tag{7.13}$$

7.8 Boundary Costs and Controls

In the last two sections, the boundary cost was identically zero and no control is allowed on the boundary. In many examples, one needs either or both boundary cost and control (Chapter 8, [72,85,86,123]). Here, we note only that for many routing problems in telecommunication, control is on the “boundary” because that corresponds to additional resources (say, a circuit) being unavailable. Only a few “algorithm oriented” comments will be made here. Let the running cost $k(x, \alpha)$ be zero on ∂G_h^+ , and define the boundary cost rate $k_0(x, \alpha)$, a continuous real valued function which is zero for $x \in G_h$. The typical cost that is accrued on the boundary is of

the form $k_0(x, \alpha)h$ and we will use this. For examples, see the above cited references.

Now (5.1) is replaced by

$$\limsup_n \frac{E_x^u \sum_0^n k(\xi_i^h, u_i^h) \Delta t_i^h + E_x^u \sum_0^n k_0(\xi_i^h, u_i^h)h}{E_x^u \sum_0^n \Delta t_i^h}. \quad (8.1)$$

Equation (5.3) continues to hold for $x \in G_h$. Let $p^h(x, y|\alpha), x \in \partial G_h^+$, denote the controlled transition probabilities on the reflecting states. For $x \in \partial G_h^+$, (5.4) is replaced by

$$W^h(x, u) = \sum_y p^h(x, y|u(x))W^h(y, u) + k_0(x, u(x))h. \quad (8.2)$$

The function defined by (5.5) continues to satisfy (5.3) and (8.2) if the cost $k(\xi_n^h, u_n^h) \Delta t_n^h$ is replaced by

$$k(\xi_n^h, u_n^h) \Delta t_n^h + k_0(\xi_n^h, u_n^h)h.$$

Now let us use the terminology of Section 7.7 above. Equations (7.8), (7.9) and (7.11) continue to hold for $x \in G_h$. Recall that x_0 does not communicate with states on ∂G_h^+ by assumption. For $x \in \partial G_h^+$, (7.12) is replaced by

$$w^h(x, u) = \sum_y p^h(x, y|u(x))w^h(y, u) + k_0(x, u(x))h. \quad (8.3)$$

In fact, suppose that (7.8) and (8.3) hold. Recall that by (7.5b), and using the fact that x_0 does not communicate to ∂G_h^+ ,

$$W^h(x_0, u) = \sum_y \bar{p}^h(x_0, y|u(x_0))w^h(y, u) + k(x_0, u(x_0))\Delta t^h.$$

Let $\pi^h(u) = \{\pi^h(x, u), x \in G_h \cup \partial G_h^+\}$ be the invariant measure for the chain $\{\xi_n^h, n < \infty\}$ on the extended state space. Multiplying (7.8) and (8.3) by $\pi^h(x, u)$ appropriately, adding, and using the invariance of $\pi^h(u)$ as in Section 7.5 yields

$$W^h(x_0, u) = \frac{\sum_{x \in G_h} \pi^h(x, u)k(x, u(x))\Delta t^h + \sum_{x \in \partial G_h^+} \pi^h(x, u)k_0(x, u(x))h}{\sum_{x \in G_h} \pi^h(x, u)} \quad (8.5)$$

Because $\sum_{x \in G_h} \pi^h(x, u)\Delta t^h$ equals the mean interpolation interval, the cost (8.1) equals $W^h(x_0, u)/\Delta t^h$, as previously.

Approximation in Policy Space. The approximation in policy update uses (7.13) for $x \in G_h$ and the minimizing controls in

$$\min_{\alpha \in \mathcal{U}} \left[\sum_y p^h(x, y|\alpha)w^h(y, u) + k_0(x, \alpha)h \right], \quad x \in \partial G_h^+. \quad (8.6)$$

Similar considerations hold for other cost functionals, when there is boundary cost and control and a reflecting process.

8

Heavy Traffic and Singular Control Problems: Examples and Markov Chain Approximations

Many of the process models which are used for purposes of analysis or control are approximations to the true physical model. Perhaps the dimension of the actual physical model is very high, or it might be difficult to define a manageable controlled dynamical (Markov) system model which describes well the quantities of basic interest. Sometimes the sheer size of the problem and the nature of the interactions of the component effects allows a good approximation to be made, in the sense that some form of the central limit theorem might be used to "summarize" or "aggregate" many of the random influences and provide a good description of the quantities of interest. Because these simpler or aggregate models will be used in an optimization problem, we need to be sure that optimal or nearly optimal controls (and the minimum value function, resp.) for the aggregated problem will also be nearly optimal for the actual physical problem (and a good approximation to the associated minimum value function, resp.).

This chapter is concerned with two classes of problems where this averaging effect can be used to simplify the model. The first class, the so-called class of *heavy traffic* problems, originated in the study of uncontrolled queueing systems [60, 90] and has applications to a broad class of such systems which include certain communication and computer networks and manufacturing systems. For these systems, "traffic" is heavy in the sense that at some processors there is little idle time. The distributions of the service and interarrival times might be dependent on the system state. The dimension of the physical problem is usually enormous. With an appropriate scaling, a functional central limit theorem argument can be used to show that the basic elements of the system can be well approximated by the solution to a certain reflected diffusion process. The approximation also allows us to compute good approximations to optimal controls and optimal value functions for the original physical system.

The second class of problems to which this chapter is devoted (and which includes part of the first class) is the class of so-called *singular control* problems. Perhaps, they are best understood as being limits of either a sequence

of discrete time problems, or as approximations to controlled continuous time systems, where the control is well approximated by a sequence of (possibly) small impulses but the control “rate” might not be representable as a bounded function. Loosely speaking, the cumulative control effort can be represented as a nondecreasing process, but not necessarily as an integral of a “control rate.”

The interest in these classes of stochastic control problems has been increasing rapidly, because they can be used to model many systems which occur in manufacturing, communication, modelling of financial transactions, and elsewhere. Numerical procedures are needed. The models have only been studied in recent years and limited analytical results are available, but the Markov chain approximation method is easily adapted and provides useful procedures. Due to the relative newness of the models, and to the fact that they often cannot be well understood without an understanding of the underlying physical processes, in Section 8.1 we will give several motivating examples. Because our interest is only in providing motivation for these controlled system models, only a brief and formal discussion will be given. The examples in Section 8.1 deal mainly with problems where the actual model which is used for numerical purposes is obtained as an “aggregation” of or other approximation to a complex physical problem.

In Section 8.2, Markov chain approximations for the “heavy traffic” case will be discussed. These are actually special cases of those in Section 5.7 for the reflected controlled diffusion, but it is useful to show how they specialize to the cases at hand. Section 8.3 deals with Markov chain approximations for the singular control problem.

8.1 Motivating Examples

8.1.1 EXAMPLE 1. A SIMPLE QUEUEING PROBLEM

In order to illustrate the size of the physical state space of the problem which is of interest, let us consider the most classical of the models of queueing theory, a single queue where the interarrival and service times are all mutually independent and exponentially distributed, and only one customer can arrive or be served at a time. In this first example, there is no control [the so-called M/M/1 queue]. Let the buffer size be N ; i.e., if there are N customers in or waiting for service and a customer arrives, then that customer is rejected from the system. Let λ_a and λ_d denote the arrival rate and service rate, resp., and let $\pi_i(t)$ denote the probability that there are i customers in the system at time t . Then it is well known that the $\pi_i(\cdot)$

satisfy the differential equations [70]

$$\begin{aligned}\dot{\pi}_0 &= \lambda_d\pi_1 - \lambda_a\pi_0, \\ \dot{\pi}_i &= \lambda_a\pi_{i-1} + \lambda_d\pi_{i+1} - (\lambda_a + \lambda_d)\pi_i, \quad i \neq 0, N, \\ \dot{\pi}_N &= \lambda_a\pi_{N-1} - \lambda_d\pi_N.\end{aligned}\quad (1.1)$$

The system (1.1) is one of the few queueing system equations that can be solved. Nevertheless, it is still quite difficult to calculate the distributions at finite times for large N . The situation is considerably worse if we allow the distributions of the service or interarrival times to be other than exponential (even if we are only concerned with the stationary distributions), or if services or arrivals can occur in batches. If the arrival or service rates depend on the state of the system (e.g., faster service or slower arrivals for longer queues), then the analytic solution of the counterpart of (1.1) can be obtained at present only in a few special cases, and even numerical solutions are generally difficult to get if N is not small. The required computation rapidly gets out of bounds if the system is a network of interconnected queues (except for the stationary solutions to the so-called Jackson cases). If a control is added to (1.1) (e.g., a control on the service rate), then the resulting control problem has a very high dimensional state space, even under the “exponential” distribution. One is strongly tempted to use some sort of approximation method to simplify the model for the queueing process. Such approximations normally require that “certain parameters” be either small or large. It was recognized in the late 1960’s [60, 90] that if the service and arrival rates are close to each other, then by a suitable scaling, such a simplifying approximation can be obtained, at least for simple systems, and the approximating process was a reflected Wiener process with drift. Subsequently, the same type of result was shown to be true in a fairly general setting [102] and for controlled problems as well [86, 95]. A further motivation for seeking aggregative or simplified models is that some sort of stochastic evolution equation and a Markov model are very helpful if the control problem is to be treated. Much of control theory uses stochastic dynamical models, and the dynamical model for the system described by (1.1) is a continuous time Markov jump process on a state space with $N + 1$ points. If N is large, then an “aggregative” model might be “numerically” approximated by a simpler process.

8.1.2 EXAMPLE 2. A HEURISTIC LIMIT FOR EXAMPLE 1

We now continue with the classical example defined above (1.1), and formally discuss how we might get a simple approximating diffusion process. Let us work in discrete time and suppose that only one arrival or departure event can occur at each discrete time. Since we are concerned with an approximation result, we will consider a family of queueing problems parametrized by ϵ , with the probability of an arrival at any time and the probability of completion of any service, conditioned on all past data, being

$\lambda + b_a\sqrt{\epsilon}$ and $\lambda + b_d\sqrt{\epsilon}$, resp., where the b_a and b_d can be either positive or negative. With these values, the arrival and service rates are within $O(\sqrt{\epsilon})$ of one another. Marginal differences of the order of $\sqrt{\epsilon}$ in the rates can make a considerable difference in the queue occupancy statistics. As the traffic intensity increases (i.e., as $\epsilon \rightarrow 0$), the mean occupancies increase, and it makes sense to scale the buffer size also. We let the buffer size be scaled as *Buffer Size*= $B/\sqrt{\epsilon}$, for some $B > 0$.

A simplifying convention concerning departures. Let Q_n^ϵ denote the number of customers waiting for or in service at discrete time n . There is a convention which is used in writing the evolution equations which simplifies the analysis considerably. In our current example, it is clear that the arrival process is independent of the queue size and departure process, but the departure process does depend on the state of the queue. In particular, if the queue is empty then there can be no departure. The convention to be used in writing the evolution equation is that even if the queue is empty, the processor will keep working and sending out outputs at the usual rate. But to keep the equations correct, a correction or “reflection” term (the ΔY terms below) will be subtracted from the departure process whenever such a “fictitious” output occurs. This clever device was used by [57, 60, 102] and others and it leads to a considerable simplification of the derivation of the approximations. If there is an arrival in the midst of such a “fictitious” interval, it is supposed that the service time of this arrival is just the residual service time for the current service interval. It can be shown that this convention does not affect the form of the limit [60, 86]

Input-Output Equation for the Queue; A Stochastic Evolution Equation. We can now write the mass balance or input-output equation as

$$Q_n^\epsilon = Q_0^\epsilon + \sum_{m=0}^{n-1} \Delta A_m^\epsilon - \sum_{m=0}^{n-1} \Delta D_m^\epsilon + \sum_{m=0}^{n-1} I_m^{d,\epsilon} - \sum_{m=0}^{n-1} I_m^{a,\epsilon}, \quad (1.2)$$

where the ΔA_m^ϵ and ΔD_m^ϵ , resp., are the number (zero or one) of arrivals or departures, resp., at time m . Keeping in mind the above convention concerning departures and fictitious departures, the ΔD_m^ϵ are the indicators of departure events assuming that the queue is never empty. The $I_m^{d,\epsilon}$ corrects for a “fictitious” output at time m if the queue is empty but a “fictitious” output occurs at that time. The term $I_m^{a,\epsilon}$ subtracts any input which arrives at time m if the buffer is full at that time.

An Approximating Process. Equation (1.2) is an evolution equation for the physical system. We would like to approximate it by simpler process. Equation (1.2) will next be rearranged in a way that will suggest the sort of approximating limit process that can be expected. Write ΔA_m^ϵ and ΔD_m^ϵ as a sum of a mean value and a random difference

$$\begin{aligned}\Delta A_m^\epsilon &= (\lambda + b_a \sqrt{\epsilon}) + \gamma_m^{a,\epsilon}, \\ \Delta D_m^\epsilon &= (\lambda + b_d \sqrt{\epsilon}) + \gamma_m^{d,\epsilon}.\end{aligned}\tag{1.3}$$

The random variables $\{\gamma_m^{a,\epsilon}, \gamma_m^{d,\epsilon}, m, a, d\}$ are mutually independent with mean zero. For use below, note that they can also be defined by

$$\gamma_m^{a,\epsilon} = \Delta A_m^\epsilon - E[\Delta A_m^\epsilon | \Delta A_i^\epsilon, \Delta D_i^\epsilon, i \leq m-1] \tag{1.4}$$

and similarly for $\gamma_m^{d,\epsilon}$. Thus, the partial sums of the $\{\gamma_m^{a,\epsilon}, m < \infty\}$ form a martingale sequence for α equal to a or d . Note that

$$\begin{aligned}E|\gamma_m^{a,\epsilon}|^2 &= \lambda(1-\lambda) + O(\sqrt{\epsilon}), \\ E|\gamma_m^{d,\epsilon}|^2 &= \lambda(1-\lambda) + O(\sqrt{\epsilon}).\end{aligned}$$

Let $[t/\epsilon]$ denote the integer part of t/ϵ . Define the continuous parameter scaled queue length process $X^\epsilon(\cdot)$ by

$$X^\epsilon(t) = \sqrt{\epsilon} Q_{[t/\epsilon]}^\epsilon.$$

Define $\Delta Y_m^\epsilon = \sqrt{\epsilon} I_m^{d,\epsilon}$ and $\Delta U_m^\epsilon = \sqrt{\epsilon} I_m^{a,\epsilon}$. Then, letting the ratio t/ϵ henceforth denote the integer part only, we can write

$$\begin{aligned}X^\epsilon(t) &= X^\epsilon(0) + (b_a - b_d)t + \sqrt{\epsilon} \sum_{m=0}^{t/\epsilon-1} \gamma_m^{a,\epsilon} - \sqrt{\epsilon} \sum_{m=0}^{t/\epsilon-1} \gamma_m^{d,\epsilon} \\ &\quad + \sum_{m=0}^{t/\epsilon-1} \Delta Y_m^\epsilon - \sum_{m=0}^{t/\epsilon-1} \Delta U_m^\epsilon + \text{"small error" } .\end{aligned}\tag{1.5}$$

For motivational purposes, note first that the first two sums in (1.5) tend to normally distributed random variables with mean zero and variance $\lambda(1-\lambda)t$ as $\epsilon \rightarrow 0$. More generally, when considered as functions of t , they converge weakly to mutually independent Wiener processes $w^a(\cdot), w^d(\cdot)$, each with variance $\lambda(1-\lambda)t$. If $X^\epsilon(0)$ converges weakly (i.e., in distribution; see Chapter 9 for the definitions) to a random variable $X(0)$ as $\epsilon \rightarrow 0$, then the sequence of processes defined in (1.5) converges weakly to limits that satisfy

$$X(t) = X(0) + (b_a - b_d)t + w^a(t) - w^d(t) + Y(t) - U(t). \tag{1.6}$$

The $Y(\cdot)$ is nondecreasing and can increase only when $X(t) = 0$. It is this reflection term which keeps the queue from going negative. The term $U(\cdot)$ is nondecreasing and can increase only when $X(t) = B$, and it keeps the buffer from exceeding the normalized limit B . Thus for small ϵ , the queueing process can be well represented by a rather simple reflected Brownian motion process with drift $b_a - b_d$.

Note that (1.6) and the other models of this section are special cases of the Skorokhod problem formulation of the reflected diffusion.

For small ϵ , many functionals of interest for the physical process $X^\epsilon(\cdot)$ can be well approximated by the same functionals of the simpler limit defined by (1.6). For example, the distribution of $X^\epsilon(t)$ and $U^\epsilon(t)$ can be well approximated by those of $X(t)$ and $U(t)$, resp. Many functionals of the entire paths can also be approximated. For example, the distribution of the first hitting time of some level $B_1 \leq B$ by $X^\epsilon(\cdot)$ can be approximated by the distribution of the first hitting time of that level by $X(\cdot)$. For another example, $\sqrt{\epsilon}$ times the number of customers lost by time t/ϵ is approximated by $U(t)$.

A Controlled Version of Example 2. The model (1.6) can also be used for control purposes. Consider one simple case, where the probabilities of arrivals or service completions are controlled and we use the “controlled” rates

$$\lambda + b_a \sqrt{\epsilon} + c_a u_a \sqrt{\epsilon}, \quad \lambda + b_d \sqrt{\epsilon} + c_d u_d \sqrt{\epsilon},$$

where u_a and u_d are the control variables. The actual value of the control at each time n is to be selected at that time, based only on information available then, and we assume that the control values u_a and u_d are bounded in absolute value by some given constants \bar{u}_a, \bar{u}_d .

Suppose that the cost for the physical system is defined by

$$W^\epsilon(x, u) = E_x^u \sum_{n=0}^{\infty} e^{-\beta n \epsilon} k(X^\epsilon(n\epsilon), u_n^\epsilon) \epsilon + E_x^u \sum_{n=0}^{\infty} e^{-\beta n \epsilon} \Delta U_n^\epsilon, \quad (1.7)$$

where $\beta > 0$. This cost weighs the loss of a customer heavily relative to the cost of control or the cost of the waiting time. Define $V^\epsilon(x) = \inf_u W^\epsilon(x, u)$ where the infimum is over all admissible controls. The appropriate controlled form of (1.6) and the associated cost are

$$\begin{aligned} X(t) &= X(0) + (b_a - b_d)t + \int_0^t (c_a u_a(s) - c_d u_d(s)) ds \\ &\quad + w^a(t) - w^d(t) + Y(t) - U(t). \end{aligned} \quad (1.8)$$

$$W(x, u) = E_x^u \int_0^{\infty} e^{-\beta t} k(X(t), u(t)) dt + E_x^u \int_0^{\infty} e^{-\beta t} dU(t). \quad (1.9)$$

Let $V(x) = \inf_u W(x, u)$, where the infimum is over all “admissible” controls for (1.8). Loosely speaking, what we mean by admissible (see Chapter 1 or 9 for a fuller discussion) is that the controls satisfy $|u_\alpha(t)| \leq \bar{u}_\alpha$ for α equal to a or d , and the $u_\alpha(t)$ are independent of the future of the Wiener processes in the sense that $u_\alpha(t)$ is independent of $\{w_a(t+s) - w_a(t), w_d(t+s) - w_d(t), s \geq 0\}$. It can be shown that $V^\epsilon(x) \rightarrow V(x)$, as $\epsilon \rightarrow 0$, and that

continuous “nearly optimal” controls for the limit problem are also “nearly optimal” when used on the physical problem. See similar results for related but more complex problems are in [80, 72, 86, 95].

These representations and approximations (1.8) and (1.9) hold under quite broad conditions, as can be seen from the references. The interarrival or service intervals need not be exponentially or geometrically distributed, and only the first and second moments of their distributions appear in the limit equation. “Batch” arrivals and services can also be introduced. Our aim here is to motivate the limit system equations only, and the reader is referred to the literature for further details. Such approximations (1.6) and (1.8) have been called *Brownian flow* systems by [55]. The numerical problem consists of computing an approximation to the optimal cost $V(x)$ and associated control (and then applying the obtained control to the physical problem). Note that the cost structure is different from that used for the problems in Chapter 3, because the process $U(\cdot)$, which is included in the cost, is not differentiable.

8.1.3 EXAMPLE 3. CONTROL OF ADMISSION, A SINGULAR CONTROL PROBLEM

In many problems in queueing and production systems, one seeks to reduce the cost associated with the waiting time by controlling the number of customers entering the system, while suitably penalizing the number denied entry. If a customer is denied entry, then that customer is assumed to disappear from the system. An alternative model might allow it to reappear later with a given probability. Generally, a higher cost is assigned to a denial of entry than to “waiting.” Example 2 will now be revised to accommodate this type of control. Let $\Delta\hat{F}_m^\epsilon$ denote the indicator of the event that an arriving customer or job has been denied entry into the queue. By convention, we suppose that a customer which arrives when the buffer is full is denied entry. Then the state equations (1.2) are

$$Q_n^\epsilon = Q_0^\epsilon + \sum_{m=0}^{n-1} \Delta A_m^\epsilon - \sum_{m=0}^{n-1} \Delta D_m^\epsilon + \sum_{m=0}^{n-1} I_m^{d,\epsilon} - \sum_{m=0}^{n-1} \Delta\hat{F}_m^\epsilon.$$

Define $\Delta F_m^\epsilon = \sqrt{\epsilon}\Delta\hat{F}_m^\epsilon$. Then in scaled form, we can write (modulo a negligible error term) the controlled analogue of (1.5)

$$\begin{aligned} X^\epsilon(t) &= X^\epsilon(0) + (b_a - b_d)t + \sqrt{\epsilon} \sum_{m=0}^{t/\epsilon-1} \gamma_m^{a,\epsilon} - \sqrt{\epsilon} \sum_{m=0}^{t/\epsilon-1} \gamma_m^{d,\epsilon} \\ &\quad + \sum_{m=0}^{t/\epsilon-1} \Delta Y_m^\epsilon - \sum_{m=0}^{t/\epsilon-1} \Delta F_m^\epsilon. \end{aligned} \tag{1.10}$$

Let the cost be (1.7) with F replacing U . The function

$$F^\epsilon(t) = \sum_{m=0}^{t/\epsilon-1} \Delta F_m^\epsilon$$

represents the scaled cumulative control action. It is implied in [95] that (with F replacing U), the sequences defined in (1.10) converge in an appropriate sense to those in (1.6) and the cost (1.7) converges to (1.9), where the $F(\cdot)$, $X(\cdot)$, $Y(\cdot)$ are non anticipative with respect to the Wiener process.

The control term $F(t)$ does not necessarily have a representation as $F(t) = \int_0^t f(s)ds$, for some bounded control rate $f(\cdot)$. The paths of $F(\cdot)$ are simply non-negative and nondecreasing right continuous functions. Such controls are called “singular” or, more accurately, singular with respect to Lebesgue measure, because they are not necessarily representable as integrals with respect to Lebesgue measure. In recent years, such controls have appeared more frequently as models of many types of systems [55, 68, 89, 19, 111, 116]. Such a control rate might not be “physical” since it might be unbounded, but as we just saw, they arise very naturally as approximations to physical systems.

Under broad conditions, one can show [95] that the optimal value functions for the physical system [(1.7), (1.10)] and for the limit system [(1.6), (1.9), with F replacing U] are arbitrarily close for small ϵ . The interarrival intervals need not be exponentially distributed. Also, nearly optimal controls for the limit system are nearly optimal for the physical system for small ϵ . In particular, the optimal control for the limit system is often of the form: There is a $B_0 \leq B$ such that $F(\cdot)$ keeps the process in the interval $[0, B_0]$ [e.g., when $k(\cdot)$ is increasing with x and is continuous]. In this case, the cost under the control which admits an arriving customer only if $X^\epsilon(t) \in [0, B_0]$ is arbitrarily close to the optimal cost for the physical system for small ϵ . This might not be surprising for this simple case, but it holds under quite weak assumptions on the arrival and service statistics. It is generally easier to compute the value of the threshold B_0 for the limit system (1.6) than it would be to work with the actual physical model.

8.1.4 EXAMPLE 4. A MULTIDIMENSIONAL QUEUEING OR PRODUCTION SYSTEM UNDER HEAVY TRAFFIC: NO CONTROL

We will next describe a multidimensional system of the type of Example 1. For the sake of expository simplicity, we first discuss the problem in the absence of control. The control will be added in the following examples. The general discussion will be loose, because our main purpose is the motivation of the somewhat nonclassical reflected diffusion systems models which arise. Let us work with a general K -dimensional system. There are K service stations or processors, each serving one customer at a time, and

the i^{th} station is referred to as P_i . Each station might have a sequence of external inputs and sends its output either to the exterior of the system or to another processor. Let p_{ij} denote the probability (conditioned on the “past” data) that the completed output of P_i goes to P_j , and with p_{i0} being the probability that the output leaves the system. It is assumed that the spectral radius of the transition matrix $P = \{p_{ij}, i, j = 1, \dots, K\}$ is less than unity. This implies that all customers leave the system eventually. Such systems are called *open systems* [56]. We also work in discrete time for notational convenience.

We use P_0 to denote the exterior, whether it is the source or the destination of a customer. Let $Q_n^{i,\epsilon}$ denote the total number of customers in or waiting for service at time n at $P_i, i = 1, \dots, K$. Without loss of generality, let $p_{ii} = 0$. This can be accomplished by a redefinition of the service intervals. Let $B_i/\sqrt{\epsilon}$ denote the buffer sizes, where B_i is assumed to be an integral multiple of $\sqrt{\epsilon}$.

The Input-Output Equations. We continue to use the convention introduced in Example 2, where the processors “keep processing” even when there are no customers, and the fictitious outputs thus created are compensated for by a cancellation or reflection term, which are the various Y -terms below. Thus the mass balance equation analogous to (1.2) is

$$\begin{aligned} Q_n^{i,\epsilon} &= Q_0^{i,\epsilon} + \sum_{m=0}^{n-1} (\text{arrivals to } P_i \text{ from the exterior at time } m) \\ &\quad + \sum_{j \neq 0} \sum_{m=0}^{n-1} (\text{arrivals to } P_i \text{ from } P_j \text{ at time } m) \\ &\quad - \sum_{j} \sum_{m=0}^{n-1} (\text{departures from } P_i \text{ to } P_j \text{ at time } m) \\ &\quad + \sum_{j} \sum_{m=0}^{n-1} (\text{corrections for fictitious departures to } P_j \text{ from } P_i \text{ at } m) \\ &\quad - \sum_{j \neq 0} \sum_{m=0}^{n-1} (\text{corrections for fictitious departures from } P_j \text{ to } P_i \text{ at } m) \\ &\quad - \sum_{m=0}^{n-1} (\text{corrections for lost inputs due to a full buffer at } P_i \text{ at } m). \end{aligned}$$

Define the scaled occupancies $X_n^{i,\epsilon} = \sqrt{\epsilon} Q_n^{i,\epsilon}$, and let X_n^ϵ denote the vector with components $\{X_n^{i,\epsilon}, i \leq K\}$. In general, for any sequence $\{Z_n^\epsilon, n < \infty\}$ define the continuous time parameter interpolation $Z^\epsilon(\cdot)$ by $Z^\epsilon(t) = Z_n^\epsilon$ on the interval $[n\epsilon, n\epsilon + \epsilon]$. Define $\Delta Y_m^{ij,\epsilon} = \sqrt{\epsilon}$ times the indicator of the event that a fictitious departure occurred at P_i at time m and it was sent to P_j . Let $\Delta A_n^{i,\epsilon}$ and $\Delta D_n^{i,\epsilon}$ be the indicators of the events that there is an external arrival at P_i and a departure from P_i to P_i at time n , resp.

Let $\Delta U_n^{i,\epsilon}$ denote the indicator of the event that there is an arrival to P_i at time n which is rejected due to a full buffer. Rewrite the above equation with the obvious notation, and where t/ϵ is used to denote the integer part

$$\begin{aligned} X^{i,\epsilon}(t) = & X_0^{i,\epsilon} + \sqrt{\epsilon} \sum_{m=0}^{t/\epsilon-1} \Delta A_m^{i,\epsilon} + \sqrt{\epsilon} \sum_{j \neq 0} \sum_{m=0}^{t/\epsilon-1} \Delta D_m^{ji,\epsilon} \\ & - \sqrt{\epsilon} \sum_j \sum_{m=0}^{t/\epsilon-1} \Delta D_m^{ij,\epsilon} + \sum_j \sum_{m=0}^{t/\epsilon-1} \Delta Y_m^{ij,\epsilon} - \sum_{j \neq 0} \sum_{m=0}^{t/\epsilon-1} \Delta Y_m^{ji,\epsilon} \\ & - \sum_{m=0}^{t/\epsilon-1} \Delta U_m^{i,\epsilon}. \end{aligned} \quad (1.11)$$

Suppose that more than one arrival to some processor occurs at some time. If the buffer capacity is not exceeded by these arrivals, then their order is unimportant. Otherwise, order them in some way and reject those that arrive when the buffer is full.

The Heavy Traffic and Other Assumptions. In keeping with the intuitive idea of heavy traffic, the average total input rate to each processor from all sources combined will be close to the service rate of that processor. As $\epsilon \rightarrow 0$, the average input and service rates for each processor converge to each other. The rates will be introduced via their “inverses”, the (conditional) means of the interarrival or service intervals. In particular, we suppose that:

The expectation of the $(n+1)^{st}$ interarrival interval (arrivals from the exterior) and service interval for P_i , each conditioned on the “past” up to the beginning of the interval in question, take the following forms, respectively

$$[g_{ai} + \sqrt{\epsilon} a_i(\text{state of system at start of that interval})]^{-1} \quad (1.12)$$

$$[g_{di} + \sqrt{\epsilon} d_i(\text{state of system at start of that interval})]^{-1},$$

where the $g_{\alpha i}$ are constants and the “marginal rates” $a_i(\cdot)$ and $d_i(\cdot)$ are bounded and continuous functions of their arguments.

Note the analogue to the form used in Example 2. The g_{ai} and g_{di} are the dominant parts of the rates of arrival (for external arrivals) and service, respectively, at P_i . We also suppose that the conditional variances of these random intervals, given the “past data,” are also continuous functions of the current state, modulo an error which goes to zero as $\epsilon \rightarrow 0$. The mathematical proof of convergence also needs the additional assumption that the set of the squares of these intervals is uniformly integrable in n, ϵ .

In order to operate in the heavy traffic environment, it is necessary that the dominant parts of the mean arrival and service rates be equal for each processor. This implies that

$$g_{ai} + \sum_{j \neq 0} p_{ji} g_{dj} = g_{di}. \quad (1.13)$$

If (1.13) does not hold, then the scaled queue length process at some processor will always be either near zero or at the upper limit for small ϵ . Form (1.12) and equation (1.13) are known as the *heavy traffic assumptions*.

Simplification of the Evolution Equation (1.11). Define $Y^{ij,\epsilon}(t) = \sum_{m=0}^{t/\epsilon-1} \Delta Y_m^{ij,\epsilon}$ and

$$Y^{i,\epsilon}(t) = \sum_{j \neq 0} Y^{ij,\epsilon}(t).$$

Then it can be shown [86, 95] that

$$Y^{ij,\epsilon}(t) = p_{ij} Y^{i,\epsilon}(t) + \text{asymptotically negligible error.}$$

In (1.4) of Example 1, we split the indicator of the event that there is an external arrival at time n into the sum of the conditional mean value $\lambda + b_a \sqrt{\epsilon}$ (which was also the unconditional mean value in that case) and the random ‘‘martingale’’ difference $\gamma_m^{a,\epsilon}$, and similarly for the indicators of the service completions. An analogous approach can be taken in the current case. The sums of the analogous ‘‘martingale differences’’ lead to the $M^\epsilon(\cdot)$ term in (1.14) below. Details are in the references. Doing this splitting and using (1.12) and (1.13) yields that the mass balance equation (1.11) can be written in the vector form

$$\begin{aligned} X^\epsilon(t) &= X^\epsilon(0) + H^\epsilon(t) + M^\epsilon(t) + (I - P')Y^\epsilon(t) - U^\epsilon(t) \\ &\quad + \text{‘‘small error terms’’.} \end{aligned} \quad (1.14)$$

The i^{th} component of $H^\epsilon(t)$ has the form

$$H^{i,\epsilon}(t) = \int_0^t b_i(X^\epsilon(s)) ds,$$

where $b_i(\cdot) = a_i(\cdot) - d_i(\cdot) + \sum_j p_{ji} d_j(\cdot)$. The $M^\epsilon(\cdot)$ is a martingale with a quadratic variation of the form $\int_0^t \Sigma(X^\epsilon(s)) ds$, where the bounded and continuous matrix valued function $\Sigma(\cdot)$ can be calculated from the conditional covariances of the external interarrival and service intervals and the p_{ij} , as shown in the references. The i^{th} component of the reflection terms $Y^\epsilon(\cdot)$ and $U^\epsilon(\cdot)$ are nonnegative, nondecreasing, and can increase only when the state $X^{i,\epsilon}(t)$ takes the value zero or B_i , resp.

Suppose that $X^\epsilon(0)$ converges to a random variable $X(0)$. Then, under suitable conditions on the $b_i(\cdot)$ and $\Sigma(\cdot)$, the sequence of processes defined in (1.14) converges to a limit which satisfies

$$X(t) = X(0) + H(t) + M(t) + (I - P')Y(t) - U(t), \quad (1.15)$$

where

$$H^i(t) = \int_0^t b_i(X(s))ds,$$

$M(\cdot)$ is a stochastic integral with respect to some Wiener process $w(\cdot)$, and can be written in the form

$$M(t) = \int_0^t \sqrt{\Sigma(X(s))}dw(s),$$

and the reflection terms $Y(\cdot)$ and $U(\cdot)$ have the properties described above in connection with (1.14). Also $X(\cdot)$, $Y(\cdot)$ and $U(\cdot)$ are nonanticipative with respect to $w(\cdot)$. Equation (1.15) describes the limit process which we wished to motivate. Note that it is in the form of the Skorokhod problem. Numerical methods for such systems and their controlled forms are of increasing interest. The Markov chain approximation method is easily adapted for use on such processes.

Remarks on (1.15). See Figure 8.1 for an example of a two dimensional problem. The directions of reflection are as indicated. At the origin, the set of allowed reflection directions is the convex cone formed by the directions at the neighboring points. See the discussion of the Skorokhod problem in Chapter 1 and the discussion of the reflecting problem in Section 5.7. The directions of reflection on each side of the box can be obtained by noting that there can be an “increment” in $Y^i(\cdot)$ [resp., in $U^i(\cdot)$] at time t only if $X^i(t) = 0$ [resp., $X^i(t) = B_i$].

Adding a Control to (1.11). Suppose that a control of the type used in (1.7) is added to (1.11). Then the limit dynamical system (1.15) is replaced by

$$X(t) = X(0) + H(t) + M(t) + (I - P')Y(t) - U(t), \quad (1.16)$$

where

$$H^i(t) = \int_0^t b_i(X(s), u(s))ds$$

for appropriate continuous functions $b_i(\cdot)$ and a control $u(\cdot)$.

Cost Functions for (1.11) and (1.16). Let G denote the state space $[0, B_1] \times [0, B_2]$ for (1.16). In order to illustrate the scaling which is needed for the cost function for the physical system (1.11), we introduce one particular but interesting case. We now restrict attention to a two dimensional

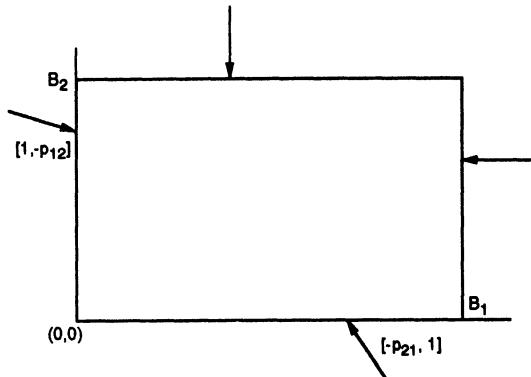


Figure 8.1. A two dimensional problem

example for notational simplicity. Let $j \neq i$, and let $k(\cdot)$ be a continuous function and k_i positive numbers. Define

$$\Delta L_m^{i,\epsilon} = \Delta A_m^{i,\epsilon} I_{\{X_m^{i,\epsilon} = B_i\}} + \Delta D_m^{ji,\epsilon} I_{\{X_m^{i,\epsilon} = B_i, X_m^{j,\epsilon} \neq 0\}}, \quad (1.17)$$

which represents an “overflow” of processor i . For an admissible control sequence $u = \{u_n^\epsilon, n < \infty\}$, let the cost for the physical system (1.11) be of the discounted form

$$W^\epsilon(x, u) = E_x^u \sum_{m=0}^{\infty} e^{-m\beta\epsilon} [\epsilon k(X_m^\epsilon, u_m^\epsilon) + \sqrt{\epsilon} k_1 \Delta L_m^{1,\epsilon} + \sqrt{\epsilon} k_2 \Delta L_m^{2,\epsilon}]. \quad (1.18)$$

The $\Delta L_m^{1,\epsilon}$ term in (1.18) penalizes the loss of a customer to processor 1 when the buffer of that processor is full. That lost customer can come from either an output of processor 2 or from the exterior of the system. Similarly for $\Delta L_m^{2,\epsilon}$.

From (1.18), we can see that the cost assigned to waiting or to the control itself (the $\epsilon k(\cdot)$ term) is small relative to the cost assigned to lost customers due to full buffers (the $\sqrt{\epsilon} \Delta L$ terms). The difference in scale is ϵ vs. $\sqrt{\epsilon}$. The reason is simply that under other scalings either one or both components of the cost become either negligible or unbounded in the limit as $\epsilon \rightarrow 0$. In addition, one might wish to heavily penalize lost customers. We note that the form (1.18) is only one possibility.

It can be shown that the process

$$\sqrt{\epsilon} \sum_{m=0}^{t/\epsilon} \Delta D_m^{ji,\epsilon} I_{\{X_m^{i,\epsilon} = B_i\}} I_{\{X_m^{j,\epsilon} = 0\}}$$

converges weakly to the zero process. Hence, the limit form of (1.18), which

is the cost for (1.16), is

$$W(x, u) = E_x^u \int_0^\infty e^{-\beta t} [k(X(t), u(t))dt + k_1 dU^1(t) + k_2 dU^2(t)]. \quad (1.19)$$

8.1.5 EXAMPLE 5. A PRODUCTION SYSTEM IN HEAVY TRAFFIC WITH IMPULSIVE CONTROL

There are many ways in which controls can be introduced into the model discussed in Example 4. For example, the marginal arrival or service rates $a_i(\cdot)$ and $d_i(\cdot)$ might be controlled. Marginal differences in the service or arrival rates can make a substantial difference in the statistics of the queues, when the traffic is heavy. We will actually discuss a different type of model, where the limit reflected diffusion model is impulsively controlled. The problem is interesting partly because the nature of the impulsive control is somewhat nonstandard and is a good example of a new class of problems to which the numerical methods can be applied. The discussion will be descriptive only and the reader is referred to [86] for more detail as well as for a discussion of a numerical algorithm. In this problem there might be “simultaneous impulses,” but the “order within the simultaneity” is crucial. By simultaneous impulses, we mean that the control action at some time t might be a sequence of impulses in different directions taken in a particular order, with no time gap between them. The possibility of multiple simultaneous impulses arises due the scaling which is used as we go to the limit in the physical model. Events which are separated in time in the physical model can occur at the same time in the limit due to the way that time is “squeezed” to get the limit model. But the order in which they occur in the physical model must be preserved in the order in which we need to take the “simultaneous” impulses in the limit model. This phenomenon affects the numerical procedure, which must keep track of the correct order. But that is not a serious problem. The physical model is generally quite hard to work with, so that the impulsively controlled reflected diffusion approximation can be quite useful.

The impulsive nature of the control for the limit model is a consequence of the way the effects of the actual control actions for the physical model accumulate as the traffic increases. The actual allowable impulses for the limit model can only be understood in the context of the physical system. This is the situation for many impulsive and singular control problems, where the actual model which is used for the calculation or analysis makes sense only as a limit of a sequence of physical processes as some scaling parameter goes to its limit.

In the system of concern, the only controls which are allowed are the actual shutting down of a processor, and the opening or closing of the links connecting the processors to each other or which connect the external sources to the processors. To each action, there is an associated immediate

cost as well as a cost due to the actual lost inputs and production. If a link P_{ij} connecting P_i to P_j is shut down but processor P_i continues to operate, then the outputs from P_i are assumed to be sent to the outside. They are lost to the processing system, but there might be a salvage value associated with them. There will be a cost for these lost customers or for the customers who cannot enter the system due to full buffers or to processors being shut down.

Due to the effects of the control actions, the input-output or the mass balance equation (1.11) needs to be modified by adding or subtracting the gains or losses which might occur due to the control actions. The mass balance equation is modified by accounting for the terms

decreases due to lost inputs to P_i when P_{0i} or some $P_{ji}, i \neq 0$ is shut off,

increases due to P_i being shut down, but some input is not shut off.

These terms give rise to the impulsive control terms of the limit model. The directions of the segments of the impulsive controls in the limit model depend on which combination of links or processors are turned off. See [86] for more detail.

8.1.6 EXAMPLE 6. A TWO DIMENSIONAL ROUTING CONTROL PROBLEM

More interesting versions of the admission control problem of Example 3 can be given when there is more than one processor, so that the internal routing as well as admissions can be controlled. See [95], where the routing only is controlled. Refer to Figure 8.2. Processor P_0 is the routing controller. It routes the inputs instantaneously to either of the other processors. There might be external inputs coming directly into processors P_1 or P_2 also. In the examples below, it is supposed that some prior preferred routing is associated with each of the arrivals to P_0 , but that this can be changed by P_0 with an associated profit or loss. For motivation, consider the following two particular cases.

Case 1. There are two classes of customers which arrive at random at P_0 (with some prior probability q_i that any new arrival will be in class i). But $P_i, i = 1, 2$, is more efficient for the class i . A prior assignment of class i to P_i is made, but P_0 can reassign to the other less efficient processor if the sizes of the waiting lines warrant it. The cost of rerouting might be the relative cost of the less efficient processor.

Case 2. Continue with Case 1, but let there be three classes of customers arriving at P_0 at random. Classes 1 and 2 must be assigned to P_1 and P_2 , resp., but there is discretion with Class 3. One of the processors is more efficient for Class 3, and a prior assignment is made to that processor. Suppose, for example, that the processors contain data bases with some overlap in their data set. Class 3 needs only the overlapping data, but

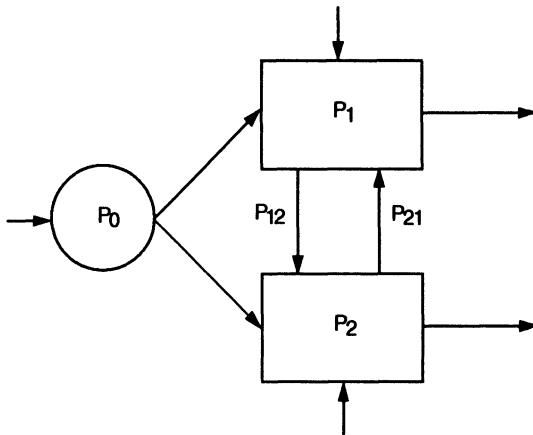


Figure 8.2. A routing control problem.

one of the processors is faster, and the prior assignment is made to that one. The prior assignment can be altered at the discretion of P_0 , with an associated cost.

The Dynamical System and Control Problem. As in Examples 4 and 5, let the buffer size for P_i be $B_i/\sqrt{\epsilon}$. Let $\Delta \hat{F}_m^{ij,\epsilon}$ denote the indicator of the event that there is an arrival at P_0 at discrete time m and which has a prior assignment to P_i , but is reassigned to P_j , where $j \neq i$. Define

$$F_n^{ij,\epsilon} = \sqrt{\epsilon} \sum_{m=0}^{n-1} \Delta \hat{F}_m^{ij,\epsilon},$$

$$F_n^{i,\epsilon} = F_n^{ji,\epsilon} - F_n^{ij,\epsilon}.$$

Again, let $U_n^{i,\epsilon}$ denote $\sqrt{\epsilon}$ times the total number of customers lost to P_i by time n due to a full buffer. The $F_n^{ij,\epsilon}(\cdot)$ represent the control. Let $A_n^{0i,\epsilon}$ denote the number of customers which would be routed to P_i from P_0 by time n according to the *a priori assignment*, and $A_n^{i,\epsilon}$ the number of customers which come directly to P_i from the exterior by time n . Let $D_n^{i,\epsilon}$ denote the number of (true or fictitious) customers departing P_i by time n . Recall that, for any sequence B_n , we define $B(t) = B_{t/\epsilon}$, where t/ϵ denotes the integer part. Then, using the other notation of Example 4, the evolution or mass balance equation which is analogous to (1.14) can be shown to be

$$\begin{aligned} X^{i,\epsilon}(t) &= X_0^{i,\epsilon} + \sqrt{\epsilon} A^{i,\epsilon}(t) + \sqrt{\epsilon} A^{0i,\epsilon}(t) - \sqrt{\epsilon} D^{i,\epsilon}(t) + F^{ji,\epsilon}(t) \\ &\quad - F^{ij,\epsilon}(t) + Y^{i,\epsilon}(t) - p_{ji} Y^{j,\epsilon}(t) - U^{i,\epsilon}(t) + \text{"small error". } \end{aligned} \tag{1.20}$$

Let $q_i > 0, k_i \geq 0$, and let $k(\cdot)$ be continuous. A cost functional which is analogous to those used in the previous examples is

$$\begin{aligned} W^\epsilon(x, F^\epsilon) = & E_x^{F^\epsilon} \int_0^\infty e^{-\beta t} k(X^\epsilon(t)) dt \\ & + E_x^{F^\epsilon} \int_0^\infty e^{-\beta t} [q_1 dF^{12,\epsilon}(t) + q_2 dF^{21,\epsilon}(t)] \\ & + E_x^{F^\epsilon} \int_0^\infty e^{-\beta t} [k_1 dU^{1,\epsilon}(t) + k_2 dU^{2,\epsilon}(t)]. \end{aligned} \quad (1.21)$$

Under the heavy traffic assumptions of the type used in Example 4, (1.20) can be approximated [95] by the “limit system”

$$X^i(t) = X_0^i + H^i(t) + M^i(t) + F^{ji}(t) - F^{ij}(t) + Y^i(t) - p_{ji}Y^j(t) - U^i(t) \quad (1.22)$$

where $i \neq j$. The limit cost functional is

$$\begin{aligned} W(x, F) = & E_x^F \int_0^\infty e^{-\beta t} k(X(t)) dt \\ & + E_x^F \int_0^\infty e^{-\beta t} [q_1 dF^{12}(t) + q_2 dF^{21}(t)] \\ & + E_x^F \int_0^\infty e^{-\beta t} [k_1 dU^1(t) + k_2 dU^2(t)]. \end{aligned} \quad (1.23)$$

Here the $H(\cdot)$ and $M(\cdot)$ terms take the forms of Example 4, with the appropriate $b(\cdot)$ and $\Sigma(\cdot)$ functions. The term $b_i(\cdot)$ is the (scaled) marginal difference between the input and service rates, and $\Sigma(\cdot)$ depends on the “randomness” of the arrival and service processes. The term

$$F^i(t) = F^{ji}(t) - F^{ij}(t)$$

is the control term for the limit system. The $F^{ij}(\cdot)$ are processes whose paths are non-negative, nondecreasing, and right continuous. They are “singular” controls, and represent the limits of the reassessments. In this problem, $F^1(\cdot) = -F^2(\cdot)$, but that is not necessarily the case in general. Thus, we have a singular control problem defined in a state space which is the hyperrectangle $B = \{x : 0 \leq x^i \leq B_i\}$, and with the type of boundary reflection directions of the “heavy traffic” type as in Example 4. The reader is referred to the references for further detail.

An Extension. For use in Section 8.3, let us write the following extension of the model (1.22) and (1.23)

$$X(t) = X(0) + \int_0^t b(X(s)) ds + \int_0^t \sigma(X(s)) dw(s) + F(t) + (I - P')Y(t) - U(t), \quad (1.22')$$

where for some integer q , $F(t)$ has the representation

$$F(t) = \sum_{i=1}^q v_i F^i(t),$$

and the $F^i(\cdot)$ are non-negative, nondecreasing, and nonanticipative processes which are right continuous. The cost function is

$$\begin{aligned} W(x, F) &= E_x^F \int_0^\infty e^{-\beta t} k(X(t)) dt \\ &\quad + E_x^F \int_0^\infty e^{-\beta t} \left[\sum_i q_i dF^i(t) + \sum_i k_i dU^i(t) \right]. \end{aligned} \quad (1.23'')$$

A Formal Dynamic Programming Equation for the Minimum Value Function for (1.22'), (1.23'). We will next give a formal development of the dynamic programming equation for the problem (1.22'), (1.23'). Let $V(x)$ denote the infimum of the value function $W(x, F)$ over all admissible controls. Let $\delta > 0$ be small, and let E_x^0 denote the expectation under zero control and initial condition x . Then, by a formal use of the principle of optimality, for $x \in G^0$, we can write

$$V(x) = \min [e^{-\beta\delta} E_x^0 V(X(\delta)) + k(X(\delta))\delta, \min_i (V(x + v_i\delta) + q_i\delta)]. \quad (1.24)$$

See Figure 8.1 for a description of the boundary conditions: On the north $V_{x_1}(x) = 0$, on the west $V_{x_2} = 0$, on the south $-p_{21}V_{x_1} + V_{x_2} = 0$, and on the east, $V_{x_1} - p_{12}V_{x_2} = 0$. In (1.24), we suppose that the “approximating” choices are either to not control over the time interval $[0, \delta]$ or else to instantaneously add an increment δ to some F^i ; i.e., either no control is used, which leads to the first term inside the outer minimum, or else there is an increment of size δ in some F^i , which leads to the term in the inner minimum. The impulsive control term in (1.24) is not discounted because the control is supposed to act instantaneously. Let \mathcal{L}^0 be the differential operator of the uncontrolled and unreflected diffusion process part of (1.22'). Next subtracting $V(x)$ from each side of (1.24) and formally expanding the terms yields that

$$\begin{aligned} \mathcal{L}^0 V(x) + k(x) - \beta V(x) &\geq 0 \\ V'_x(x)v_i + q_i &\geq 0, \quad i = 1, \dots, k \end{aligned}$$

and at each point x , at least one of the $k+1$ terms equals zero. Thus, we formally have

$$\min [\mathcal{L}^0 V(x) + k(x) - \beta V(x), \min_i (V'_x(x)v_i + q_i)] = 0. \quad (1.25)$$

The reflecting boundary conditions need to be added to (1.25).

Equation (1.25), together with its boundary conditions, is known as a *variational inequality*. For the singular control problem, it is the replacement for the PDE's obtained in Chapter 3. See [111] for a more mathematical derivation for some related problems (without reflection). The numerical method based on the Markov chain approximation allows us to avoid dealing with (1.25), because we approximate the original control problem (1.22'), (1.23'), rather than the equation (1.25). In Section 8.3, the dynamic programming equation for the approximating Markov chain will be given and its formal similarity to (1.25) noted.

8.1.7 EXAMPLE 7

An interesting problem in portfolio selection which involves a combination of singular and ordinary control is in [31]. Let $x = (x_0, x_1)$, where $x_0 \geq 0$ is the bank account balance and $x_1 \geq 0$ the amount invested in stocks. Let $U(t)$ [resp., $L(t)$] denote the total value of stock sales (resp., purchases) by time t , and let $c(\cdot)$ be the “consumption rate.” There are transactions costs for sales (resp., purchases): one pays a fraction μ [resp., λ] of the transactions amount. “Infinitesimal” transactions are allowed and the model is

$$\begin{aligned} dx_0 &= (r_0 x_0 - c)dt - (1 + \lambda)dL + (1 - \mu)dU, \\ dx_1 &= r_1 x_1 dt + \sigma x_1 dw + dL - dU, \end{aligned} \quad (1.26)$$

where r_0 and r_1 are the bank interest rate and the mean rate of increase of the value of the stocks. The controls are $u = (c, L, U)$.

For a suitable utility function, one wishes to maximize the profit

$$W(x, u) = E_x^u \int_0^\infty e^{-\beta t} k(c(t)) dt. \quad (1.27)$$

In [31], the form $k(c) = c^\lambda$ for $\lambda \in (0, 1)$ was used, and this allowed the authors to get an (essentially) explicit analytic solution in terms of the ratio x^0/x^1 . For other types of utility functions or processes, a numerical method might be needed.

Since the state space in (1.26) is unbounded, we might have to bound it for numerical purposes. This can be done by putting upper limits \bar{x}_0 and \bar{x}_1 on the bank deposit and stock holding, resp. The upper bounds should be large enough so that they do not seriously affect the numerical results for typical values of the state variables.

The model (1.26), (1.27) in [31] follows a common usage in financial modeling in that a strictly concave utility function is used. This causes a difficulty with unbounded consumption “rates” of the type that would result when upper bounds on the state variable are imposed, because it would give them zero value. This problem is easy to avoid by defining an appropriate value for the forced consumption at the upper bound.

Comments. Only a small sample of the many types of heavy traffic and singular control problems have been considered. [111] treat a singular control problem where there are infinitely many directions of control. The numerical method for the problem given in [85] is a straightforward extension of that given here. [56] and [123] treat a problem which arises in heavy traffic modelling, where there are several classes of customers, a “throughput” constraint, and an ergodic cost criterion. A numerical method is developed in [85], which gives a general approach to the Markov chain approximation method for the ergodic cost problem with a singular control. Forms of the reflecting diffusions which arise as heavy traffic limits of the “trunk line” problems in telephone routing can also be dealt with. Ergodic costs can be handled for all of the problems of this chapter

8.2 The Heavy Traffic Problem: A Markov Chain Approximation

8.2.1 THE BASIC MODEL

Consider the model (1.16) and cost function (1.19), where $X^i(t) \in [0, B_i]$, $i = 1, 2$. Under appropriate conditions, they are limits of the system and cost function (1.11) and (1.18), resp. With these limits in hand, we would like to obtain numerical approximations to the optimal cost function $V(x)$ and the optimal control. In practice, one would use an appropriate adaptation of a “nearly” optimal control for (1.16) and (1.19) on the actual physical system. It is helpful to know that the resulting controls will be nearly optimal for the physical system under broad conditions if ϵ is small enough. See [86] and [95] for further information on some specific classes of problems. Thus the main remaining problem concerns the numerical solution of the optimization problem for (1.16) and (1.19). The construction of a locally consistent Markov chain is actually a slight extension of the method used in Section 5.7 for the reflected problem and we will review the details for our special case.

The models (1.15) and (1.16) can be put into the form of the Skorokhod Problem of Chapter 1. The possible reflection directions are dictated by the form of the reflection terms $(I - P')Y(t)$ and $-U(t)$. We next show how to read the correct reflection directions from (1.15) or (1.16) (they are the same for both systems). Refer to Figure 8.3a.

The Reflection Directions for (1.15) and (1.16). Write $x = (x^1, x^2)$. G is the inside box in the figure. The reflection direction are constant on each of the four open sides of the boundary of G and are as follows:

- (a) For $x^2 = B_2$, $x^1 \in (0, B_1)$, the reflection direction is $r(x) = -e_2 = (0, -1)$; i.e., it points downward.

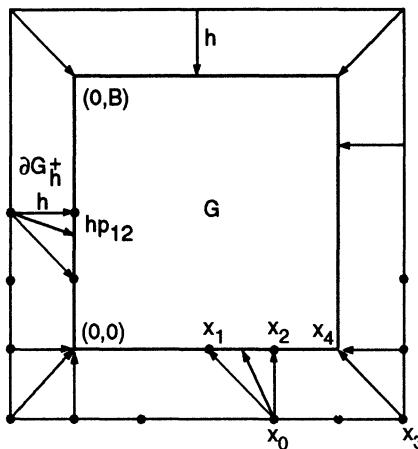


Figure 8.3a. The boundary transitions.

(b) For $x^1 = B_1$, $x^2 \in (0, B_2)$, the reflection direction is $r(x) = -e_1 = (-1, 0)$.

(c) For $x^2 = 0$, $x^1 \in (0, B_1)$, the reflection direction is $r(x) = (-p_{21}, 1)$.

(d) For $x^1 = 0$, $x^2 \in (0, B_2)$, we have $r(x) = (1, -p_{12})$.

The set of allowed reflection directions at a corner is the convex hull of those associated with the adjoining sides. Recall that the spectral radius of the “connection” probability matrix P is assumed to be less than unity, in order to guarantee that each customer spends only a finite average time in the system. We have also normalized the system such that we can assume $p_{ii} = 0$. Thus, $p_{12} + p_{21} < 2$. Let us now examine (1.16) a little more closely (but still heuristically) to see how these reflection directions are actually obtained.

More Details of the Calculation of the Reflection Directions for (1.16). Write (1.16) in the form

$$X(s) = R(s) + (I - P')Y(s) - U(s), \quad R(s) = (R^1(s), R^2(s)).$$

Fix t , and suppose that $X^2(t) = 0$, $X^1(t) \in (0, B_1)$. Loosely speaking, suppose that $R^2(\cdot)$ tries to “pull” $X^2(\cdot)$ negative at time t . Then this pull needs to be compensated for by an increase in $Y^2(\cdot)$. In particular, let $X^2(s) = 0$, $X^1(s) \in (0, B_1)$ on the time interval $[t, t + \delta]$ with $R^i(t + \delta) - R^i(t) = \Delta R^i < 0$ for $i = 2$. Write $Y^i(t + \delta) - Y^i(t) = \Delta Y^i$. Then we must have

$$\begin{pmatrix} 1 & -p_{21} \\ -p_{12} & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \Delta Y^2 \end{pmatrix} = - \begin{pmatrix} 0 \\ \Delta R^2 \end{pmatrix}.$$

Because $\Delta Y^2 > 0$, this is equivalent to the reflection direction being $(-p_{21}, 1)$ for the boundary points in question.

Next, let us repeat this procedure for the corner point $(0, 0)$. Let $X^2(s) = X^1(s) = 0$ on the interval $[t, t + \delta)$, and suppose that $\Delta R^i < 0, i = 1, 2$. Then using the terminology introduced above,

$$\begin{pmatrix} 1 & -p_{21} \\ -p_{12} & 1 \end{pmatrix} \begin{pmatrix} \Delta Y^1 \\ \Delta Y^2 \end{pmatrix} = - \begin{pmatrix} \Delta R^1 \\ \Delta R^2 \end{pmatrix}.$$

This implies that the set of reflection directions at the corner is the convex hull of those at the adjoining sides.

8.2.2 THE NUMERICAL METHOD

The numerical problem is quite similar to that for the variation of the Example 2 in Section 5.7 drawn in Figure 5.8. Let G_h denote the restriction to G of the state space of a controlled Markov chain which is locally consistent with the unreflected form of (1.16), and with interpolation interval $\Delta t^h(x, \alpha)$. In particular, for this illustrative example we will suppose that the state space is a regular grid with spacing h in each direction and that the B_i are integral multiples of h . To complete the description of the approximating chain for the reflected diffusion (1.16), we need only describe its behavior on the reflecting boundary. Refer to Figure 8.3a, where an acceptable reflecting boundary ∂G_h^+ is shown as the grid points on the “outer” boundary. The ∂G_h^+ is disjoint from points on G . We could use points on G as a “numerical boundary,” but it is often more convenient from the programming point of view to create an “external” reflecting boundary. Indeed, the external boundary is often closer to the physics of the problem, where the reflection actually arises from a constraint.

A Locally Consistent Transition Probability on the Boundary ∂G_h^+ . Recall that the transition probabilities for the reflecting states are not controlled in our examples here (see Section 7.8). Consider point x_0 . If p_{21} equals one (resp., zero) then the state goes from x_0 to x_1 (resp., x_2). Suppose that $p_{21} \in (0, 1)$. To realize this *desired mean reflection direction*, use the randomization method of Section 5.7 and set

$$p^h(x_0, x_1) = p_{21} = 1 - p^h(x_0, x_2).$$

Similarly for all points x such that $x^2 = -h$ and $x^1 \in (0, B_1]$. The procedure is the same for the left hand boundary. The assignment can be completed by taking all other points on ∂G_h^+ to the nearest point on G_h . This set of rules gives us a natural and locally consistent approximation.

Further Examination of the Lower Right Hand Corner of G . The rules just suggested for the boundary behavior do yield a locally consistent approximating chain, and they seem to be the simplest rules for this problem. For possible use on other classes of problems, it is useful to note

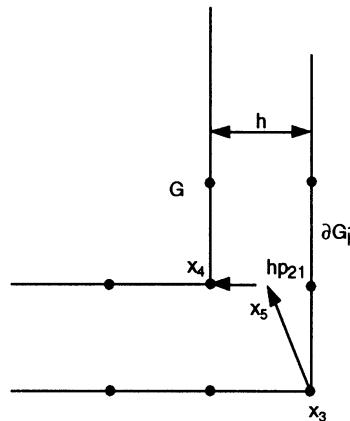


Figure 8.3b. Transitions at the corner.

that there are other possibilities, and one will be illustrated for the lower right hand corner. Before doing that let us “dissect” in Case 1 below the transition from x_3 to $x_4 = (B_1, 0)$ which is implied by the $p^h(x, y)$ given by the above stated rules. The results of this discussion will be needed in order to get the correct cost functional for the approximating Markov chain.

Case 1. Refer to Figure 8.3b. Let $\xi_n^h = x_3$. By the above stated rule, we have $p^h(x_3, x_4) = 1$ so that $\xi_{n+1}^h = x_4$. At the point x_3 two constraints are violated, since $x_3^2 < 0$ and $x_3^1 > B_1$. It is instructive to break the movement from x_3 to x_4 into two segments, correcting for each constraint violation in turn. Let us first correct for the second component of x_3 which is negative, and then for the first component which is larger than B_1 . To do this, we move first from x_3 along the reflection direction $(-p_{21}, 1)$ to the point x_5 in Figure 8.3b. Then move the remaining distance $(1-p_{21})h$ from x_5 to x_4 . We can consider the point x_3 to correspond physically to the event that there is a lost customer at queue 1, due to a full buffer and also a simultaneous fictitious output from queue 2. If $p_{21} = 1$, then the fictitious output of queue 2 was the actual input to the buffer of queue 1. In this case, $x_5 = x_4$. If $p_{21} < 1$, then the fictitious output of queue 2 equals the overflow of the buffer of queue 1 only with probability p_{21} . Hence with probability $1 - p_{21}$, the lost input was an actual true input (which came from the exterior of the system in this case). We will return to this decomposition below when we discuss the appropriate cost function to use for the controlled chain. The method of case 1 is the most natural and is the one which will be used.

Case 2. A second possibility for a locally consistent transition probability at x_3 can be obtained by reversing the order of the “corrections” of case 1. Refer to Figure 8.3c. Let us correct for the overflow first, by taking x_3 to x_6 . Then correct for the fact that the first component of the state is still negative by moving along the reflection direction $(-p_{21}, 1)$ to the point x_7 .

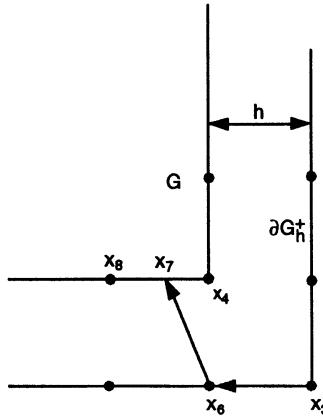


Figure 8.3c. Transitions at the corner.

If \$p_{21} < 1\$, then \$x_7\$ is not a grid point and we need to randomize between \$x_8\$ and \$x_4\$ in the usual way so as to achieve the desired mean value. It can be verified that this choice of the transition probability at \$x_3\$ is also locally consistent. It will yield the same asymptotic results as Case 1.

We note that, in the limit, the contribution of the corner points to the cost is zero, so that the actual form of the cost used at the corners is asymptotically irrelevant.

The Continuous Time Interpolation \$\psi^h(\cdot)\$. Refer to Section 5.7.3 for some of the terminology. Let \$n\$ be a reflection step, and use the terminology of Section 5.7. Thus, \$\xi_n^h = x \in \partial G_h^+\$. Write \$\Delta \xi_n^h = \Delta z_n^h + \Delta \tilde{z}_n^h\$, where \$\Delta z_n^h = E_{x,n}^{h,\alpha} \Delta \xi_n^h\$. In the problem of concern here, the grid spacing is \$O(h)\$ and we have

$$E_{x,n}^{h,\alpha} |\Delta \xi_n^h - \Delta z_n^h|^2 = O(h^2) = O(h) |\Delta z_n^h|.$$

Thus,

$$E \sup_{m \leq n} \left| \sum_{i=0}^m \Delta \tilde{z}_i^h \right|^2 = O(h) E \left| \sum_{i=0}^n \Delta z_i^h \right| \quad (2.1)$$

and (5.7.3) holds.

Owing to the special rectangular shape of \$G\$, and using the Case 1 “decomposition” of the corners, for \$\xi_n^h = x \in \partial G_h^+\$ we can write \$\Delta \xi_n^h\$ in the form

$$\Delta z_n^h = (I - P') \Delta Y_n^h - \Delta U_n^h, \quad (2.2)$$

where \$\Delta Y_n^{h,i}\$ (resp., \$\Delta U_n^{h,i}\$) are non-negative and can increase only if \$\xi_n^{h,i} < 0\$ (resp., \$> B_i\$). Because in our case only the “lower” reflection terms might be randomized, the right side of (2.1) is bounded above by \$O(h) E |Y_{n+1}^h|\$.

Define the interpolations $u^h(t) = u_n^h$ on $[\tau_n^h, \tau_{n+1}^h)$ and

$$Y^h(t) = \sum_{n: \tau_{n+1}^h \leq t} \Delta Y_n^h, \quad U^h(t) = \sum_{n: \tau_{n+1}^h \leq t} \Delta U_n^h$$

and similarly define $\tilde{z}^h(\cdot)$. In Chapter 11, it will be shown that $\tilde{z}^h(\cdot) \rightarrow$ zero process. Now, (5.7.4) takes the form

$$\begin{aligned} \psi^h(t) = x + & \int_0^t b(\psi^h(s), u^h(s)) ds + M^h(t) + J^h(t) \\ & + (I - P')Y^h(t) - U^h(t) + \delta_1^h(t), \end{aligned} \quad (2.3)$$

where $M^h(\cdot)$, $J^h(\cdot)$ and $\delta_1^h(\cdot)$ are defined as in (5.7.4).

The Dynamic Programming Equation for the Markov Chain Approximation. Let $\{\xi_n^h, n < \infty\}$ denote a Markov chain which is locally consistent with the reflecting diffusion (1.16), and use the transition functions described above for the reflecting states. Let $u = \{u_n^h, n < \infty\}$ be an admissible control sequence for the chain. A cost functional which is analogous to (1.19) is, for $k_i \geq 0$,

$$\begin{aligned} W^h(x, u) = & E_x^u \sum_{n=0}^{\infty} e^{-\beta t_n^h} k(\xi_n^h, u_n^h) \Delta t_n^h \\ & + E_x^u \sum_{n=0}^{\infty} e^{-\beta t_n^h} k_1 h [I_{\{\xi_n^{h,1} > B_1, \xi_n^{h,2} \geq O\}} \\ & \quad + (1 - p_{21}) I_{\{\xi_n^{h,1} > B_1, \xi_n^{h,2} < O\}}] \\ & + E_x^u \sum_{n=0}^{\infty} e^{-\beta t_n^h} k_2 h [I_{\{\xi_n^{h,2} > B_2, \xi_n^{h,1} \geq O\}} \\ & \quad + (1 - p_{12}) I_{\{\xi_n^{h,2} > B_2, \xi_n^{h,1} < O\}}] \end{aligned} \quad (2.4)$$

The first sum on the right of (2.4) is obviously an appropriate discretization of the corresponding part of the integral in (1.19). The next sum on the right hand side is an appropriate analogue of the integral in (1.19) involving $U^1(\cdot)$ as we shall now see. Suppose that $\xi_n^{h,1} = B_1 + h$, $\xi_n^{h,2} \geq 0$. Because then $\xi_{n+1}^{h,1} = B_1$, the correct overflow correction is h . Suppose that $\xi_n^{h,1} = B_1 + h$, $\xi_n^{h,2} < 0$. Then recalling the discussion in Case 1 above, we see that the mean overflow “correction” is $h(1 - p_{21})$. The third sum in (2.4) is explained in the same way. Again we note that the terms in (2.4) with $(1 - p_{ij})$ have no effect in the limit and can be dropped.

By the decomposition in Case 1, (2.4) can be written as

$$W^h(x, u) = E_x^u \sum_{n=0}^{\infty} e^{-\beta t_n^h} k(\xi_n^h, u_n^h) \Delta t_n^h + E_x^u \sum_{n=0}^{\infty} e^{-\beta t_n^h} [k_1 \Delta U_n^{h,1} + k_2 \Delta U_n^{h,2}]. \quad (2.5)$$

Let $V^h(x)$ denote the infima of $W^h(x, u)$ over the admissible control sequences u . For $x \in G_h$, the dynamic programming equation is

$$V^h(x) = \min_{\alpha \in \mathcal{U}} \left[e^{-\beta \Delta t^h(x, \alpha)} \sum_y p^h(x, y | \alpha) V^h(y) + k(x, \alpha) \Delta t^h(x, \alpha) \right]. \quad (2.6a)$$

For $x \in \partial G_h^+$, the interpolation interval equals zero, the transition probabilities are not controlled and the dynamic programming equation is

$$\begin{aligned} V^h(x) &= \sum_y p^h(x, y) V^h(y) \\ &\quad + k_1 h [I_{\{x^1 > B_1, x^2 \geq 0\}} + (1 - p_{21}) I_{\{x^1 > B_1, x^2 < 0\}}] \\ &\quad + k_2 h [I_{\{x^2 > B_2, x^1 \geq 0\}} + (1 - p_{12}) I_{\{x^2 > B_2, x^1 < 0\}}]. \end{aligned} \quad (2.6b)$$

8.3 Singular Control: A Markov Chain Approximation

We will work with a two dimensional problem for ease of visualization. It should be apparent that the ideas are of quite general applicability. The system and cost function will be (1.22') and (1.23'), resp., and the sets G , G_h , and ∂G_h^+ are the same as in Section 8.2. Let $q_i > 0$, $k_i \geq 0$. Let $p^h(x, y)$ and $\Delta t^h(x)$ be a transition probability and interpolation interval which are locally consistent with the reflected diffusion (1.22') on the state space $S_h = G_h \cup \partial G_h^+$ when the control term $F(\cdot)$ is dropped. Without loss of generality, let the v_i in (1.22') satisfy:

All the components of the vectors v_i are no greater than unity in absolute value, and at least one component equals unity in absolute value.

The control in (1.22') can be viewed as a sequence of small impulses acting “instantaneously.” With this in mind, we divide the possible behavior of the approximating Markov chain into three classes:

(i) Suppose that $\xi_n^h = x \in \partial G_h^+$. Then we have a “reflection step”, as in Section 5.7 or Section 8.2 and $\Delta t^h(x) = 0$.

Otherwise, we are in the set $G_h \in G$ and there are two choices, only one of which can be exercised at a time:

(ii) Do not exercise control and use $p^h(x, y)$ and $\Delta t^h(x)$ which are locally consistent with the uncontrolled and unreflected diffusion.

(iii) Exercise control and choose the control as described in the next paragraph.

The Control Step. In order to illustrate the procedure in the simplest way, let there be only two distinct “impulsive directions,” namely, v_1 and v_2 . Again, the extensions to the general case should be clear. Suppose that

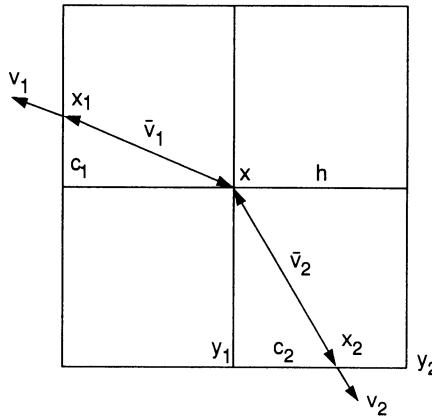


Figure 8.4 The control directions.

$x = \xi_n^h \in G_h$ and we decide to exert control. Define the increment

$$\Delta F_n^h = \sum_{i=1}^2 v_i \Delta F_n^{h,i},$$

where the increments $\Delta F_n^{h,i}$ are non-negative. The impulsive control action is determined by the choice of the direction among the $\{v_i\}$ and the magnitude of the impulse in the chosen direction. An impulse in only one direction will be chosen at each control step (not necessary, but convenient for programming). Let v_i be the chosen direction. For convenience in programming, it is preferable if the states move only “locally.” For this reason, the value of the increment $\Delta F_n^{h,i}$ is chosen to take the state x only to neighboring points.

The procedure is illustrated in Figure 8.4 in a canonical case. In the figure, x_i denotes the point of first intersection of the direction vectors v_i with the neighboring grid lines. Define $\bar{v}_i = x_i - x$. Obviously, the x_i and \bar{v}_i depend on h . We define $\Delta F_n^{h,i}$ by the relationship $v_i \Delta F_n^{h,i} = \bar{v}_i$. If more choices were allowed for the increments, then the asymptotic results for the optimal value function would be the same. If the absolute value of each component of the vector v_i were either unity or zero, then x_i would be a point in G_h . Otherwise, it is not. Analogously to what was done for the reflection problem in Section 5.7 or in Section 8.2 the actual transition is chosen by a randomization which keeps the mean value as \bar{v}_i . Thus, for the example in the figure, at a control step we will have $E_{x,n}^{h,\alpha} \Delta \xi_n^h$ equalling either \bar{v}_1 or \bar{v}_2 , according to the choice of the control direction; that is, the mean increments will be \bar{v}_1 or \bar{v}_2 , resp. Suppose, in particular, that v_2 is the desired control direction. Then $\Delta F_n^{h,1} = 0$, $\Delta F_n^{h,2} = h$. Write $p^h(x, y | \bar{v}_i)$ for the transition probability if the control direction is v_i . The corresponding transition probability is (see the figure for the notation) for

$i = 2$:

$$p^h(x, y_1 | \bar{v}_2) = 1 - c_2 = 1 - p^h(x, y_2 | \bar{v}_2) \quad (3.1)$$

The transition probabilities under the choice of direction v_1 are analogous. In general the components $\Delta F_n^{h,i}$ take values either zero or h . In analogy to (1.23'), if v_i is the chosen direction at time n , then an appropriate cost to assign to this control step is

$$q_i h = q_i \Delta F_n^{h,i}.$$

A Comment on the Limit of the Control Terms. In the limit, as $h \rightarrow 0$, the effects of the randomization disappear. A short calculation which suggests this will now be given. It will be used when convergence is discussed in Chapter 11. If n is a control step, then by definition

$$E_n^h \Delta \xi_n^h = \Delta F_n^h. \quad (3.2)$$

Define the error $\Delta \tilde{F}_n^h$ due to the randomization by

$$\Delta \tilde{F}_n^h = \Delta \xi_n^h - \Delta F_n^h. \quad (3.3)$$

Since the partial sums of the $\Delta \tilde{F}_n^h$ form a martingale sequence, we can write

$$E \sup_{n \leq N} \left| \sum_{j=0}^{n-1} \Delta \tilde{F}_j^h \right| = O(h) E \sum_{j=0}^{N-1} |\Delta F_j^h|. \quad (3.4)$$

Equation (3.4) implies that the “error” goes to zero if the sequence of the costs due to the control are bounded.

The Cost Function and Dynamic Programming for the Approximating Chain. Set $\xi_0^h = x$. Define $\Delta t_n^h = \Delta t^h(\xi_n^h)$ for n a diffusion step, and set $\Delta t_n^h = 0$ otherwise. Define $t_n^h = \sum_0^{n-1} \Delta t_i^h$. For F^h an admissible control sequence, a suitable analogue of (1.23') is

$$W^h(x, F^h) = E_x^{F^h} \sum_{n=0}^{\infty} e^{-\beta t_n^h} \left[k(\xi_n^h) \Delta t_n^h + \sum_i q_i \Delta F_n^{h,i} + \sum_i k_i \Delta U_n^{h,i} \right]. \quad (3.5)$$

Then, for $x \in G_h$, the dynamic programming equation is

$$\begin{aligned} V^h(x) &= \min \left\{ e^{-\beta \Delta t^h(x)} \sum_y p^h(x, y) V^h(y) + k(x) \Delta t^h(x), \right. \\ &\quad \left. \min_i \left[\sum_y p^h(x, y | \bar{v}_i) V^h(y) + q_i h \right] \right\}, \end{aligned} \quad (3.6)$$

and any suitable approximation to the discount factor can be used. For $x \in \partial G_h^+$, the dynamic programming equation is (2.6b).

A Relationship Between (3.6) and the Dynamic Programming Equation for (1.22') and (1.23'). The formal dynamic programming equation for (1.22') and (1.23') is (1.25), to which the reader is referred.

Suppose that the $p^h(x, y)$ in (3.6) has been obtained via a finite difference approximation of the type discussed in Sections 5.1-5.3. Thus, we can represent the sum

$$e^{-\beta \Delta t^h(x)} \sum_y p^h(x, y) V^h(y) - V^h(x) + k(x) \Delta t^h(x)$$

as $\Delta t^h(x)$ times a finite difference approximation to $\mathcal{L}^0 V(x) + k(x) - \beta V(x)$. Let us now rearrange the control term in (3.6) such that it resembles a finite difference approximation to the control term in (1.25). We work with the special two dimensional case of Figure 8.4.

Recall that

$$\begin{aligned} 1 - c_1 &= p^h(x, x - e_1 h | \bar{v}_1), \\ 1 - c_2 &= p^h(x, x - e_2 h | \bar{v}_2). \end{aligned}$$

Subtract $V^h(x)$ from both sides of (3.6). Then the inner minimum divided by h equals (here $i \neq j$)

$$\min_i \left[\frac{V^h(x - e_i h) - V^h(x)}{h} (1 - c_i) + \frac{V^h(x - e_i h + e_j h) - V^h(x)}{h} c_i + q_i \right],$$

which is a finite difference approximation to

$$\min_i \left[-V_{x_i}^h(x) + c_i V_{x_j}^h(x) + q_i \right] = \min_i [V_x^h(x)' v_i + q_i].$$

This last expression is just the inner minimum in (1.25), once the superscript h is dropped.

9

Weak Convergence and the Characterization of Processes

This chapter begins the section of the book devoted to the convergence proofs and related matters. The purpose of the chapter is to introduce the mathematical machinery that is needed in the later chapters. Because particular applications are intended, we do not, in general, give the most elaborate versions of the theorems to be presented.

Our method for proving convergence of numerical schemes is based on the theory of weak convergence of probability measures. The theory of weak convergence of probability measures provides a powerful extension of the notion of convergence in distribution for finite dimensional random variables. For the particular problems of this book, the probability measures are the induced measures defined on the path spaces of controlled processes. This notion of convergence is important for our purposes, since our approximations to value functions always have representations as expectations of functionals of the controlled processes.

The first section of the chapter is concerned with general results and the standard methodology used in weak convergence proofs. Included in this section is a statement of the Skorokhod representation, which allows the replacement of weak convergence of probability measures by convergence with probability one of associated random variables (in an appropriate topology) for the purposes of certain calculations. The usual application of weak convergence requires a compactness result on the sequence of probability measures (to force convergence of subsequences), together with a method of identification of limits. In Section 9.2 we present sufficient conditions for the required compactness. The conditions will turn out to be simple to verify for the problems considered in later chapters. Section 9.3 discusses useful characterizations of the Wiener process and Poisson random measures. These results will be used as part of a direct method of characterizing stochastic processes that will be used often in the sequel. Therefore, after introducing and discussing the method, we expose some of the details of its typical application via an example involving uncontrolled processes in Section 9.4. In Section 9.5, we define what is meant by a “relaxed control.” Relaxed controls provide a very powerful tool in the study of the convergence properties of sequences of optimally (or “nearly optimally”) controlled processes. This is due to the fact that under general

conditions, arbitrary sequences of relaxed controls have compact closure. This is not true of ordinary controls. However, our use of relaxed controls is simply as a device for proving convergence of numerical schemes. The controls computed for the discrete state approximating chains will always be feedback, or Markov, controls.

9.1 Weak Convergence

9.1.1 DEFINITIONS AND MOTIVATION

Let S denote a metric space with metric d and let $C(S)$ denote the set of real valued continuous functions defined on S . Let $C_b(S)$ and $C_0(S)$ denote the subsets of $C(S)$ given by all continuous functions that are bounded and have compact support, respectively.

Suppose we are given S -valued random variables $X_n, n < \infty$ and X , which may possibly be defined on different probability spaces and which take values in S . Let E_n and E denote expectation on the probability spaces on which the X_n and X are defined, respectively. Then we say that the sequence $\{X_n, n < \infty\}$ converges in distribution to X if $E_n f(X_n) \rightarrow Ef(X)$ for all $f \in C_b(S)$. Let $P_n, n < \infty$, and P denote the measures defined on $(S, \mathcal{B}(S))$ that are induced by $X_n, n < \infty$ and X , respectively. Clearly, the property of converging in distribution depends only on these measures: for any $f \in C_b(S)$,

$$E_n f(X_n) \rightarrow Ef(X) \Leftrightarrow \int_S f(s) P_n(ds) \rightarrow \int_S f(s) P(ds).$$

We will refer to this form of convergence of probability measures as *weak convergence* and use the notation $P_n \Rightarrow P$. Often, we abuse terminology and notation and also say that the sequence of random variables X_n that are associated in the manner described above with the measures P_n converges weakly to X and denote this by $X_n \Rightarrow X$. Let $g(\cdot)$ be any continuous function from S into any metric space. A direct consequence of the definition of weak convergence is that $X_n \Rightarrow X$ implies $g(X_n) \Rightarrow g(X)$. A general reference for the theory of weak convergence is [12]. More recent works that emphasize the case when the limit is a Markov process and applications are [42] and [79].

We can jump ahead a bit and indicate the reasons for our particular interest in this notion of convergence. Consider for the moment the special case of $S = D^k[0, \infty)$. An example of a random variable that takes values in the space S is the uncontrolled diffusion process $x(\cdot)$ where

$$dx(t) = b(x(t))dt + \sigma(x(t))dw(t),$$

with $x(0) = x$ given. Of course, $x(\cdot)$ also takes values in the smaller space $C^k[0, \infty)$, and it is not a priori evident why we have chosen to use the

larger space $D^k[0, \infty)$. One important reason is that a basic compactness result that will be needed in the approach described momentarily is easier to prove for processes in $D^k[0, \infty)$.

Suppose that one is interested in calculating a quantity such as

$$W(x) = E_x \left[\int_0^T k(x(s))ds \right],$$

where the function $k(\cdot)$ is continuous and bounded. Then when considered as a function defined on $D^k[0, \infty)$, the mapping $\phi \rightarrow \int_0^T k(\phi(s))ds$ is bounded and continuous. Suppose also that there are “approximations” $\xi^h(\cdot)$ to $x(\cdot)$ available for which the analogous functional

$$W^h(x) = E_x \left[\int_0^T k(\xi^h(s))ds \right]$$

could be readily computed. Recall that some candidate approximations were developed in Chapter 5. A finite state Markov chain $\{\xi_n^h, n < \infty\}$ and an interpolation interval $\Delta t^h(x)$ satisfying the local consistency conditions were constructed, and the process $\xi^h(\cdot)$ was then defined as a piecewise constant interpolation of $\{\xi_n^h, n < \infty\}$. Thus, the processes $\{\xi^h(\cdot), h > 0\}$ take values in $D^k[0, \infty)$. Then, if the sense in which the processes $\xi^h(\cdot)$ approximate $x(\cdot)$ is actually $\xi^h(\cdot) \Rightarrow x(\cdot)$, we may conclude $W^h(x) \rightarrow W(x)$. This simple observation is the basis for the convergence proofs to follow.

In order to make the procedure described above applicable in a broad setting, we will need convenient methods of verifying whether or not any given sequence of processes (equivalently sequence of measures) converges weakly and also for identifying the limit process (equivalently limit measure).

9.1.2 BASIC THEOREMS OF WEAK CONVERGENCE

Let $\mathcal{P}(S)$ be the space of probability measures on $(S, \mathcal{B}(S))$, and suppose that P_1 and P_2 are in $\mathcal{P}(S)$. For a given set $A \in \mathcal{B}(S)$, define $A^\epsilon = \{s' : d(s', s) < \epsilon \text{ for some } s \in A\}$. We define the *Prohorov metric* on $\mathcal{P}(S)$ by

$$\pi(P_1, P_2) = \inf \{\epsilon > 0 : P_1(A^\epsilon) \leq P_2(A) + \epsilon \text{ for all closed } A \in \mathcal{B}(S)\}.$$

We will see below that convergence in the Prohorov metric is equivalent to weak convergence when S is separable. This equivalence makes the result which follows significant for our purposes. A standard reference for most of the material of this section is Billingsley [12], where proofs of the theorems can be found.

Theorem 1.1. *If S is complete and separable, then $\mathcal{P}(S)$ is complete and separable.*

Let $\{P_\gamma, \gamma \in \Gamma\} \subset \mathcal{P}(S)$, where Γ is an arbitrary index set. The collection of probability measures $\{P_\gamma, \gamma \in \Gamma\}$ is called *tight* if for each $\epsilon > 0$ there exists a compact set $K_\epsilon \subset S$ such that

$$\inf_{\gamma \in \Gamma} P_\gamma(K_\epsilon) \geq 1 - \epsilon. \quad (1.1)$$

If the measures P_γ are the induced measures defined by some random variables X_γ , then we will also refer to the collection $\{X_\gamma, \gamma \in \Gamma\}$ as tight. The condition (1.1) then reads (in the special case where all the random variables are defined on the same space)

$$\inf_{\gamma \in \Gamma} P\{X_\gamma \in K_\epsilon\} \geq 1 - \epsilon.$$

Theorem 1.2. (Prohorov's Theorem) *If S is complete and separable, then a set $\{P_\gamma, \gamma \in \Gamma\} \subset \mathcal{P}(S)$ has compact closure in the Prohorov metric if and only if $\{P_\gamma, \gamma \in \Gamma\}$ is tight.*

Assume that S is complete and separable and that a given sequence of probability measures has compact closure with respect to the Prohorov metric. It then follows from Theorem 1.1 that existence of a convergent subsequence is guaranteed. In typical applications we will then show that the limits of all convergent subsequences are the same. Arguing by contradiction, this will establish the convergence of the original sequence. Prohorov's theorem provides an effective method for verifying the compact closure property. The usefulness of this result is in part due to the fact that tightness can be formulated as a property of the random variables associated to the measures P_n . Often these objects have representations (e.g., SDE) which allow a convenient verification of the tightness property.

Remark 1.3. A simple corollary that will be useful for our purposes is the following. Let S_1 and S_2 be complete and separable metric spaces, and consider the space $S = S_1 \times S_2$ with the usual product space topology. For $\{P_\gamma, \gamma \in \Gamma\} \subset \mathcal{P}(S)$, let $\{P_{\gamma,1}, \gamma \in \Gamma\} \subset \mathcal{P}(S_1)$ and $\{P_{\gamma,2}, \gamma \in \Gamma\} \subset \mathcal{P}(S_2)$ be defined by taking $P_{\gamma,i}$ to be the marginal distribution of P_γ on S_i , for $i = 1, 2$. Then $\{P_\gamma, \gamma \in \Gamma\}$ is tight if and only if $\{P_{\gamma,1}, \gamma \in \Gamma\}$ and $\{P_{\gamma,2}, \gamma \in \Gamma\}$ are tight.

We next present several statements which are equivalent to weak convergence. In particular, we note that statements (ii) and (iii) in the theorem below will be used in some of the convergence proofs of Chapters 10 and 11. Let ∂B be the boundary of the set $B \in \mathcal{B}(S)$. A set B is said to be a P -continuity set if $P(\partial B) = 0$.

Theorem 1.4. *Let S be a metric space and let $P_n, n < \infty$, and P be elements of $\mathcal{P}(S)$. Then statements (i) – (iv) below are equivalent and are*

implied by (v). If S is separable, then (i) – (v) are equivalent.

- (i) $P_n \Rightarrow P$
- (ii) $\limsup_n P_n(F) \leq P(F)$ for closed sets F
- (iii) $\liminf_n P_n(O) \geq P(O)$ for open sets O
- (iv) $\lim_n P_n(B) = P(B)$ for P – continuity sets B
- (v) $\pi(P_n, P) \rightarrow 0$.

Part (iv) of Theorem 1.4 suggests the following useful extension.

Theorem 1.5. Let S be a metric space, and let $P_n, n < \infty$, and P be probability measures on $\mathcal{P}(S)$ satisfying $P_n \Rightarrow P$. Let f be a real valued measurable function on S and define D_f to be the measurable set of points at which f is not continuous. Let X_n and X be random variables which induce the measures P_n and P on S , respectively. Then $f(X_n) \Rightarrow f(X)$ whenever $P\{X \in D_f\} = 0$.

Consider once again the example of an uncontrolled diffusion $x(\cdot)$ discussed at the beginning of this section. Suppose that we are now interested in estimating

$$W(x) = E_x \left[\int_0^T k(x(s))ds + g(x(T)) \right],$$

where g is a smooth bounded function. Assume we have approximations $\xi^h(\cdot)$ to $x(\cdot)$ in the sense that $\xi^h(\cdot) \Rightarrow x(\cdot)$ in $D^k[0, \infty)$. Because the function $\phi \rightarrow g(\phi(T))$ is not continuous on $D^k[0, \infty)$, we cannot directly apply the definition of weak convergence to conclude $W^h(x) \rightarrow W(x)$, where

$$W^h(x) = E_x \left[\int_0^T k(\xi^h(s))ds + g(\xi^h(T)) \right].$$

However, Theorem 1.5 implies the convergence still holds, since the limit process $x(\cdot)$ has continuous sample paths (w.p.1) and $\phi(\cdot) \rightarrow g(\phi(T))$ is continuous at all $\phi(\cdot)$ which are continuous.

Remark 1.6. Note that there is an obvious extension of the definition of weak convergence of probability measures, in which the requirement that the measures be probability measures is dropped. For $T < \infty$, let $\mathcal{M}_T(S)$ denote the set of Borel measures $M(\cdot)$ on S satisfying $M(S) = T$. Via the identification $\mathcal{M}_T(S) = \{TP : P \in \mathcal{P}(S)\}$, we have analogues of all the statements above regarding weak convergence of probability measures. In particular, we note the following. If we define tightness for a collection $\{M_\gamma, \gamma \in \Gamma\} \subset \mathcal{M}_T(S)$ by requiring that for all $\epsilon > 0$ there exist compact $K_\epsilon \subset S$ such that

$$\inf_{\gamma \in \Gamma} M_\gamma(K_\epsilon) \geq T - \epsilon,$$

then subsets of $\mathcal{M}_T(S)$ are relatively compact in the topology of weak convergence if and only if they are tight.

We finish this section by recalling the *Skorokhod representation*. Suppose we are given a sequence of random variables X_n tending weakly to a limit X . We will see many times in the sequel that the evaluation of limits of certain integrals associated with the X_n are essential for the purposes of characterizing the limit X . These calculations would be simpler if all the random variables were defined on the same probability space and if the convergence were actually convergence with probability one. The following theorem allows us assume this is actually the case when computing the integrals. A proof can be found in [42].

Theorem 1.7. *Let S be a separable metric space, and assume the probability measures $P_n \in \mathcal{P}(S)$, $n < \infty$, tend weakly to $P \in \mathcal{P}(S)$. Then there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ on which there are defined random variables \tilde{X}_n , $n < \infty$, and \tilde{X} such that for all Borel sets B and all $n < \infty$,*

$$\tilde{P} \left\{ \tilde{X}_n \in B \right\} = P_n(B), \quad \tilde{P} \left\{ \tilde{X} \in B \right\} = P(B),$$

and such that

$$\tilde{X}_n \rightarrow \tilde{X}$$

with probability one.

9.2 Criteria for Tightness in $D^k[0, \infty)$

In the previous section we discussed the notion of tightness of a set of probability measures on a metric space S , which turned out to be equivalent to precompactness of the set for spaces S that are complete and separable. In this section we examine the particular case $S = D^k[0, \infty)$. We will consider sets $\{P_\gamma, \gamma \in \Gamma\}$ such that each P_γ is the measure induced on $D^k[0, \infty)$ by a process $x^\gamma(\cdot)$. To simplify the notation, we will assume that all the processes $\{x^\gamma, \gamma \in \Gamma\}$ are defined on a common probability space (Ω, \mathcal{F}, P) . Statements when this is not the case are obvious modifications of the ones given here. The criteria for tightness described below, which will be quite simple to apply for our problems, are due to Aldous and Kurtz [71, Theorem 2.7b]. Recall that the random time τ is an \mathcal{F}_t -stopping time if $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \in [0, \infty)$.

Theorem 2.1. *Consider an arbitrary collection of processes $\{x^\gamma, \gamma \in \Gamma\}$ defined on the probability space (Ω, \mathcal{F}, P) and taking values in $D^k[0, \infty)$. Assume that for each rational $t \in [0, \infty)$ and $\delta > 0$ there exists compact $K_{t,\delta} \subset \mathbb{R}^k$ such that $\sup_{\gamma \in \Gamma} P\{x^\gamma(t) \notin K_{t,\delta}\} \leq \delta$. Define \mathcal{F}_t^γ to be the σ -algebra generated by $\{x^\gamma(s), s \leq t\}$. Let T_T^γ be the set of \mathcal{F}_t^γ -stopping*

times which are less than or equal to T w.p.1, and assume for each $T \in [0, \infty)$ that

$$\lim_{\delta \rightarrow 0} \sup_{\gamma \in \Gamma} \sup_{\tau \in T_T^\gamma} E(1 \wedge |x^\gamma(\tau + \delta) - x^\gamma(\tau)|) = 0. \quad (2.1)$$

Then $\{x^\gamma, \gamma \in \Gamma\}$ is tight.

9.3 Characterization of Processes

As remarked in Section 9.1, our approach to proving the convergence of numerical schemes will be based on proving the weak convergence of a sequence of stochastic processes to an appropriate limit process. In Section 9.2 a useful condition for precompactness of sequences of processes was given. In this section we will give characterizations of Wiener processes and Poisson random measures. These characterizations will be used as part of a rather straightforward method for identifying the limits of convergent sequences of processes, where the elements of the sequence have been chosen to approximate the solution to some given SDE. Together with the precompactness, this will imply the convergence of processes in the form we need. The basic approach to be used is as follows. By rewriting the dynamical equations of the prelimit processes in an appropriate way, the “parts” of the processes corresponding to the limiting process’ Wiener process and Poisson random measure are identified and the appropriate convergences demonstrated. Then, by using a simple approximation argument, it is demonstrated under a weak sense uniqueness assumption that the sequence of approximations to the SDE converge weakly to the solution of the SDE. This basic method will be used frequently in the sequel. An elementary but detailed example of its application is given in the next section for the case of an uncontrolled diffusion process.

An alternative method of characterization that has come to be widely used is the *martingale problem* method of Stroock and Varadhan. This method could be used here as well and, in fact, has been applied to these problems in the past [79, 76]. We have chosen not to use it here because it does not seem to provide for more general results and requires a separate statement of the appropriate martingale problem for each of the problem classes we consider. However, it should be noted that the martingale problem formulation can be a useful tool in establishing the weak sense uniqueness which we assume. Other methods are also available (e.g., semigroup based approaches) and a comparison and discussion of various features may be found in [42].

Wiener Process. Consider a process $w(\cdot)$ defined on a probability space (Ω, \mathcal{F}, P) which takes values in $C^n[0, \infty)$ (for some $n < \infty$), and which

satisfies $w(0) = 0$ w.p.1. Let \mathcal{L}_w denote the differential operator defined by

$$\mathcal{L}_w f(x) = \frac{1}{2} \sum_{i=1}^n f_{x_i x_i}(x)$$

for $f \in C_0^2(\mathbb{R}^n)$. Define

$$M_f(t) = f(w(t)) - f(0) - \int_0^t \mathcal{L}_w f(w(s)) ds.$$

Suppose that \mathcal{F}_t is a filtration such that $w(\cdot)$ is an n -dimensional \mathcal{F}_t -Wiener process. Then Itô's formula implies that $M_f(t)$ is an \mathcal{F}_t -martingale for all $f \in C_0^2(\mathbb{R}^n)$. A useful fact is that the converse is also true. If there is a filtration \mathcal{F}_t defined on (Ω, \mathcal{F}, P) such that $M_f(t)$ is an \mathcal{F}_t -martingale for all $f \in C_0^2(\mathbb{R}^n)$, then $w(\cdot)$ is an n -dimensional \mathcal{F}_t -Wiener process. In particular, if $w(t)$ is a continuous local martingale whose quadratic variation is It , then $w(\cdot)$ is a Wiener process with respect to the filtration it generates.

Poisson Random Measure. Consider an integer valued random measure $N(\cdot)$ defined on a probability space (Ω, \mathcal{F}, P) which for each $\omega \in \Omega$ is a measure on the Borel subsets of $[0, \infty) \times \Gamma$. For $\theta \in C(\Gamma)$, set $\theta_N(t) = \int_0^t \int_\Gamma \theta(\rho) N(ds d\rho)$, and let \mathcal{L}_N^θ denote the operator defined by

$$\mathcal{L}_N^\theta f(x) = \lambda \int_\Gamma [f(x + \theta(\rho)) - f(x)] \Pi(d\rho)$$

for $f \in C_0(\mathbb{R})$. Define

$$M_f^\theta(t) = f(\theta_N(t)) - f(0) - \int_0^t \mathcal{L}_N^\theta f(\theta_N(s)) ds.$$

If there is a filtration \mathcal{F}_t defined on (Ω, \mathcal{F}, P) such that $M_f^\theta(t)$ is an \mathcal{F}_t -martingale for all $f \in C_0(\mathbb{R})$ and $\Theta \in C(\Gamma)$, then $N(\cdot)$ is an \mathcal{F}_t -Poisson random measure with intensity measure $\lambda dt \times \Pi(d\rho)$ on $[0, \infty) \times \Gamma$.

Lastly, we note the important fact that a Wiener process and a Poisson random measure that are defined on the same probability space and with respect to the same filtration are mutually independent [61, Theorem 6.3].

9.4 An Example

It is instructive to see how the results outlined so far in this chapter yield convergence of numerical schemes in a simple example. Because the example is purely motivational, and since the full control problem will be

treated in Chapter 10, we consider a numerical problem for a simple one dimensional diffusion. Although simple, the example will illustrate the way that weak convergence methods can be used to justify numerical approximations. The problem considered can in some cases be more easily treated by classical methods from numerical analysis. However, it will expose some important points and will illustrate the typical use of the material presented in the last three sections. This section will also serve as a reference point for our more involved use of the basic methods later in the book.

We consider a problem with a diffusion process that is the solution to the one dimensional SDE

$$dx = b(x)dt + \sigma(x)dw, \quad x(0) = x. \quad (4.1)$$

The problem of interest will be the approximation of the function

$$W(x) = E_x \left[\int_0^\tau k(x(s))ds + g(x(\tau)) \right],$$

where $\tau = \inf\{t : x(t) \in \{0, a\}\}$ and $W(x) = g(x)$ for $x \in \{0, a\}$. This is the problem considered in Example 3 of Section 5.1. For simplicity, we assume the existence of $c > 0$ such that $\sigma^2(x) \geq c$. We also assume that $b(\cdot)$ and $\sigma(\cdot)$ are Lipschitz continuous. Recall that by Theorem 3.1 this implies the weak sense uniqueness of the solution to (4.1).

Selection of an Approximating Chain. Let h be such that a/h is an integer. For our choice of approximating Markov chain we can use any of the chains developed in Section 5.1 for this model. Using the notation of Chapter 5, we define

$$p^h(x, x \pm h) = \frac{\sigma^2(x)/2 + hb^\pm(x)}{\sigma^2(x) + h|b(x)|}, \quad \Delta t^h(x) = \frac{h^2}{\sigma^2(x) + h|b(x)|}.$$

Let ξ_n^h be a Markov chain with transition probabilities $p^h(\cdot, \cdot)$ and $\xi_0^h = x$. As usual, we define

$$\xi^h(t) = \xi_n^h \text{ for } t \in [t_n^h, t_{n+1}^h), \text{ where } t_n^h = \sum_{i=0}^{n-1} \Delta t^h(\xi_i^h).$$

We also define the stopping times $\tau_h = t_{N_h}^h$, $N_h = \inf\{n : \xi_n^h \in \{0, a\}\}$. The approximation to $W(x)$ is then given by

$$\begin{aligned} W^h(x) &= E_x^h \left[\sum_{i=0}^{N_h-1} k(\xi_i^h) \Delta t^h(\xi_i^h) + g(\xi_{N_h}^h) \right] \\ &= E_x^h \left[\int_0^{\tau_h} k(\xi^h(s))ds + g(\xi^h(\tau_h)) \right] \end{aligned}$$

for points x of the form $x = ih, i \in \{0, \dots, a/h\}$. Recall that $W^h(x)$ satisfies the relation (5.1.13), which is the equation used for numerically computing $W^h(\cdot)$.

In this example, we have used the processes $\xi^h(\cdot)$, rather than $\psi^h(\cdot)$, in the representation for $W^h(x)$. This was done in order to slightly simplify the notation. It is worth noting that it is often more convenient to work with the $\psi^h(\cdot)$ processes. See, for example, Chapter 11. In general, we will work with the processes that are most convenient for the problem at hand.

An Outline of the Proof of Convergence. We will prove that $\xi^h(\cdot) \Rightarrow x(\cdot)$, $\tau_h \Rightarrow \tau$, and that the mapping

$$(\phi(\cdot), t) \rightarrow \int_0^t k(\phi(s))ds + g(\phi(t))$$

is continuous (w.p.1) with respect to the measure induced by $(x(\cdot), \tau)$. By Theorem 1.5 this implies

$$\int_0^{\tau_h} k(\xi^h(s))ds + g(\xi^h(\tau_h)) \Rightarrow \int_0^\tau k(x(s))ds + g(x(\tau)).$$

Under a uniform integrability condition that is also proved below, the last equation implies the convergence of the expectations, and therefore $W^h(x) \rightarrow W(x)$. Most of the effort involves showing $\xi^h(\cdot) \Rightarrow x(\cdot)$. To do this, we first prove tightness. We then take any subsequence of $\{\xi^h(\cdot), h > 0\}$, extract a convergent subsequence, and identify the limit as a solution to (4.1). The weak sense uniqueness of solutions to (4.1) then gives $\xi^h(\cdot) \Rightarrow x(\cdot)$.

Theorem 4.1. *The collection $\{\xi^h(\cdot), h > 0\}$ is tight.*

Proof. We must show the assumptions of Theorem 2.1 hold for the process $\xi^h(\cdot)$. Recall that E_n^h denotes expectation conditioned on $\mathcal{F}(\xi_i^h, i \leq n)$ and that $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$. By construction, the chain satisfies the local consistency conditions, and by a calculation given in Section 5.1

$$\begin{aligned} E_n^h \Delta\xi_n^h &= b(\xi_n^h) \Delta t^h(\xi_n^h), \\ E_n^h (\Delta\xi_n^h - E_n^h \Delta\xi_n^h)^2 &= [\sigma^2(\xi_n^h) + O(h)] \Delta t^h(\xi_n^h). \end{aligned} \tag{4.2}$$

Let $N_h(t) = \max\{n : t_n^h \leq t\}$. Using (4.2), we compute

$$\begin{aligned} E_x^h |\xi^h(t) - x|^2 &= E_x^h \left| \sum_{i=0}^{N_h(t)-1} [E_i^h \Delta\xi_i^h + (\Delta\xi_i^h - E_i^h \Delta\xi_i^h)] \right|^2 \\ &\leq 2E_x^h \left(\left| \sum_{i=0}^{N_h(t)-1} E_i^h \Delta\xi_i^h \right|^2 + \left| \sum_{i=0}^{N_h(t)-1} (\Delta\xi_i^h - E_i^h \Delta\xi_i^h) \right|^2 \right) \end{aligned}$$

$$\begin{aligned}
&\leq 2E_x^h \left| \sum_{i=0}^{N_h(t)-1} b(\xi_i^h) \Delta t^h(\xi_i^h) \right|^2 \\
&\quad + 2E_x^h \sum_{i=0}^{N_h(t)-1} [\sigma(\xi_i^h) + O(h)] \Delta t^h(\xi_i^h) \\
&\leq 2K^2 t^2 + 2(K + O(h)) t,
\end{aligned}$$

where K is a bound for $|b(x)| \vee |\sigma(x)|$ for all $x \in [0, a]$. Together with Chebyshev's inequality, this yields the first condition assumed in Theorem 2.1.

We must also prove (2.1). In the present context, this condition may be rewritten as

$$\lim_{\delta \rightarrow 0} \limsup_{h \rightarrow 0} \sup_{\tau \in \mathcal{T}_T^h} E_x^h (1 \wedge |\xi^h(\tau + \delta) - \xi^h(\tau)|) = 0, \quad (4.3)$$

where \mathcal{T}_T^h be the set of \mathcal{F}_t^h -stopping times which are less than or equal to T w.p.1 and \mathcal{F}_t^h is the σ -algebra generated by $\{\xi^h(s), s \leq t\} = \{\xi_i^h, i \leq N_h(t)\}$. The limit (4.3) can be proved by calculations similar to those of the previous paragraph. Using the strong Markov property of the process $\{\xi_i^h, i < \infty\}$, we have

$$\begin{aligned}
E_x^h (1 \wedge |\xi^h(\tau + \delta) - \xi^h(\tau)|) &\leq \left(E_x^h |\xi^h(\tau + \delta) - \xi^h(\tau)|^2 \right)^{1/2} \\
&\leq (2K^2 \delta^2 + 2(K + O(h)) \delta)^{1/2},
\end{aligned}$$

for any $\tau \in \mathcal{T}_T^h$. This implies (4.3). ■

Theorem 4.2. *The processes $\xi^h(\cdot)$ converge weakly to a solution $x(\cdot)$ of equation (4.1).*

Proof. In the arguments to follow, we will extract subsequences several times. To keep the notation reasonable, we will abuse notation and retain the index h for each successive subsequence. Consider any subsequence $\{\xi^h(\cdot), h > 0\}$. By tightness, we may extract a weakly convergent subsequence, again referred to as $\{\xi^h(\cdot), h > 0\}$. We will prove $\xi^h(\cdot) \Rightarrow x(\cdot)$. By the usual argument by contradiction, this proves that the original sequence converges weakly to $x(\cdot)$.

To prove that $\xi^h(\cdot)$ converges weakly to the solution of (4.1), we essentially “construct” the Wiener process appearing in the representation (4.1) for the limit process. The local consistency condition (4.2) gives

$$\xi^h(t) - x = \sum_{i=0}^{N_h(t)-1} [E_i^h \Delta \xi_i^h + (\Delta \xi_i^h - E_i^h \Delta \xi_i^h)]$$

$$= \sum_{i=0}^{N_h(t)-1} b(\xi_i^h) \Delta t^h(\xi_i^h) + \sum_{i=0}^{N_h(t)-1} \sigma(\xi_i^h) [(\Delta \xi_i^h - E_i^h \Delta \xi_i^h) / \sigma(\xi_i^h)]$$

This suggests the definition

$$w^h(t) = \sum_{i=0}^{n-1} (\Delta \xi_i^h - E_i^h \Delta \xi_i^h) / \sigma(\xi_i^h) \quad \text{for } t \in [t_n^h, t_{n+1}^h]. \quad (4.4)$$

From the calculations which were used to prove the tightness of the sequence $\{\xi^h(\cdot), h > 0\}$ we obtain tightness of the sequence $\{w^h(\cdot), h > 0\}$. Let $\{(\xi^h(\cdot), w^h(\cdot)), h > 0\}$ be a convergent subsequence, and denote the limit by $(x(\cdot), w(\cdot))$. We first prove that $w(\cdot)$ is indeed a Wiener process.

Let us fix $t \geq 0, \tau > 0, q < \infty, t_i \in [0, t]$ with $t_{i+1} > t_i$ for $i \in \{0, \dots, q\}$, and any bounded continuous function $H : I\!\!R^{2 \times q} \rightarrow I\!\!R$. Let $f \in C_0^2(I\!\!R)$ and let \mathcal{L}_w be the differential operator of the Wiener process, i.e., $\mathcal{L}_w f(x) = (1/2)f_{xx}(x)$. From the definition of $w^h(\cdot)$,

$$\begin{aligned} & f(w^h(t+\tau)) - f(w^h(t)) - \int_t^{t+\tau} \mathcal{L}_w f(w^h(s)) ds \\ &= \sum_{i=N_h(t)}^{N_h(t+\tau)-1} [f(w^h(t_{i+1}^h)) - f(w^h(t_i^h))] \\ &\quad - \sum_{i=N_h(t)}^{N_h(t+\tau)-1} \frac{1}{2} f_{xx}(w^h(t_i^h)) \Delta t^h(\xi_i^h) + O(h^2) \\ &= \sum_{i=N_h(t)}^{N_h(t+\tau)-1} f_x(w^h(t_i^h)) [\Delta \xi_i^h - E_i^h \Delta \xi_i^h] / \sigma(\xi_i^h) \\ &\quad + \frac{1}{2} \sum_{i=N_h(t)}^{N_h(t+\tau)-1} f_{xx}(w^h(t_i^h)) [\Delta \xi_i^h - E_i^h \Delta \xi_i^h]^2 / \sigma^2(\xi_i^h) \\ &\quad - \frac{1}{2} \sum_{i=N_h(t)}^{N_h(t+\tau)-1} f_{xx}(w^h(t_i^h)) \Delta t^h(\xi_i^h) + \epsilon^h + O(h^2), \end{aligned}$$

where $E^h|\epsilon^h| \rightarrow 0$ as $h \rightarrow 0$. By using this expression together with the consistency condition (4.2) we have

$$\begin{aligned} & E^h H(\xi^h(t_i), w^h(t_i), 1 \leq i \leq q) \times \\ & \left[f(w^h(t+\tau)) - f(w^h(t)) - \int_t^{t+\tau} \mathcal{L}_w f(w^h(s)) ds \right] \\ & \leq E^h |\epsilon^h| + O(h). \end{aligned} \quad (4.5)$$

At this point we would like to take limits in (4.5) in order to obtain

$$\begin{aligned} EH(x(t_i), w(t_i), 1 \leq i \leq q) & \left[f(w(t + \tau)) - f(w(t)) - \int_t^{t+\tau} \mathcal{L}_w f(w(s)) ds \right] \\ & = 0. \end{aligned} \tag{4.6}$$

If all of the processes $\xi^h(\cdot)$, $w^h(\cdot)$, $x(\cdot)$, and $w(\cdot)$ were defined on a common probability space and if, rather than weak convergence, the actual sense of convergence were $(\xi^h(\cdot), w^h(\cdot)) \rightarrow (x(\cdot), w(\cdot))$ w.p.1., then (4.6) would follow from the dominated convergence theorem. But for purposes of simply computing the expectation in (4.6) it is only the distributions of the processes that are important and not the probability space on which the processes $(x(\cdot), w(\cdot))$ are defined. By the Skorokhod representation (Theorem 1.7) there exists a probability space on which there are defined random processes $(\tilde{\xi}^h(\cdot), \tilde{w}^h(\cdot))$ and $(\tilde{\xi}(\cdot), \tilde{w}(\cdot))$ such that for $h > 0$, $(\tilde{\xi}^h(\cdot), \tilde{w}^h(\cdot))$ has the same distribution as $(\xi^h(\cdot), w^h(\cdot))$, $(\tilde{\xi}(\cdot), \tilde{w}(\cdot))$ has the same distribution as $(\xi(\cdot), w(\cdot))$, and $(\tilde{\xi}^h(\cdot), \tilde{w}^h(\cdot)) \rightarrow (\tilde{\xi}(\cdot), \tilde{w}(\cdot))$ w.p.1. By replacing $\xi^h(\cdot)$, $w^h(\cdot)$, $x(\cdot)$, and $w(\cdot)$ in (4.5) and (4.6) by $\tilde{\xi}^h(\cdot)$, $\tilde{w}^h(\cdot)$, $\tilde{\xi}(\cdot)$, and $\tilde{w}(\cdot)$, and taking limits as $h \rightarrow 0$, we obtain (4.6) as written.

Let $\mathcal{F}_t = \mathcal{F}(x(s), w(s), s \leq t)$. \mathcal{F}_t is also the σ -algebra generated by all random variables of the form $H(x(t_i), w(t_i), 1 \leq i \leq q)$, where $H(\cdot)$ is any bounded and continuous function of $2q$ variables, $t_i \in [0, t]$ for $i \leq q$, and q is any finite integer. By (4.6)

$$E \left[f(w(t + \tau)) - f(w(t)) - \int_t^{t+\tau} \mathcal{L}_w f(w(s)) ds \right] I_A = 0$$

for all $A \in \mathcal{F}_t$, which is equivalent to the statement that

$$E \left[f(w(t + \tau)) - f(w(t)) - \int_t^{t+\tau} \mathcal{L}_w f(w(s)) ds \middle| \mathcal{F}_t \right] = 0$$

w.p.1. We therefore conclude that $f(w(t)) - f(w(0)) - \int_0^t \mathcal{L}_w f(w(s)) ds$ is an \mathcal{F}_t -martingale for all $f \in C_0^2(\mathbb{R}^k)$. By construction the jumps of $h^h(\cdot)$ converge to zero uniformly, which implies that the limit process $w(\cdot)$ has continuous sample paths. Hence, the characterization given in Section 9.3 implies $w(\cdot)$ is an \mathcal{F}_t -Wiener process.

We next identify the process $x(\cdot)$. In the same way that it was used in the next to last paragraph we can employ the Skorokhod representation and assume that $\xi^h(\cdot) \rightarrow x(\cdot)$ with probability one. For each $\delta > 0$ define

$$\left. \begin{aligned} \xi_\delta^h(t) &= \xi^h(j\delta) \\ x_\delta(t) &= x(j\delta) \end{aligned} \right\} \text{ for } t \in [j\delta, j\delta + \delta).$$

Then $\xi_\delta^h(\cdot) \rightarrow x_\delta(\cdot)$ with probability one in $D[0, \infty)$. From the definition

of the process $w^h(\cdot)$,

$$\begin{aligned}\xi^h(t) - x &= \sum_{i=0}^{N_h(t)-1} [E_i^h \Delta \xi_i^h + (\Delta \xi_i^h - E_i^h \Delta \xi_i^h)] \\ &= \int_0^t b(\xi^h(s))ds + \sum_{i=0}^{N_h(t)-1} \sigma(\xi_i^h) [w^h(t_{i+1}^h) - w^h(t_i^h)] + O(h^2).\end{aligned}$$

Using this representation, the continuity and boundedness of $b(\cdot)$ and $\sigma(\cdot)$, and the tightness of $\{\xi^h(\cdot), h > 0\}$, we can write

$$\begin{aligned}\xi_\delta^h(t) - x &= \int_0^t b(\xi_\delta^h(s))ds + \sum_{j=0}^{[t/\delta]} \sigma(\xi_\delta^h(j\delta)) [w^h(j\delta + \delta) - w^h(j\delta)] \\ &\quad + O(h^2) + \epsilon_{\delta,t}^h,\end{aligned}$$

where $[s]$ denotes the integer part of s and where $E|\epsilon_{\delta,t}^h| \rightarrow 0$ as $\delta \rightarrow 0$, uniformly in $h > 0$ and t in any bounded interval. Taking the limit as $h \rightarrow 0$ yields

$$x_\delta(t) - x = \int_0^t b(x_\delta(s))ds + \sum_{j=0}^{[t/\delta]} \sigma(x_\delta(j\delta)) [w(j\delta + \delta) - w(j\delta)] + \epsilon_{\delta,t},$$

where $E|\epsilon_{\delta,t}| \rightarrow 0$ as $\delta \rightarrow 0$. For each j , the random variable $x_\delta(j\delta) = x(j\delta)$ is independent of the random variables $\{w(s) - w(j\delta), s \geq j\delta\}$ (since $w(\cdot)$ is an \mathcal{F}_t -Wiener process). The boundedness of $\sigma(\cdot)$ and properties of the Wiener process imply

$$x_\delta(t) - x = \int_0^t b(x_\delta(s))ds + \int_0^t \sigma(x_\delta(s))dw(s) + \bar{\epsilon}_{\delta,t},$$

where $E|\bar{\epsilon}_{\delta,t}| \rightarrow 0$ as $\delta \rightarrow 0$. By (1.3.3),

$$\int_0^t \sigma(x_\delta(s))dw(s) \rightarrow \int_0^t \sigma(x(s))dw(s)$$

as $\delta \rightarrow 0$. Therefore, $x(\cdot)$ solves

$$x(t) = x + \int_0^t b(x(s))ds + \int_0^t \sigma(x(s))dw(s)$$

as claimed. ■

A Topology for the Set $[0, \infty]$. Because the cost $W^h(x)$ also involves the stopping times τ_h , we must also consider weak convergence for sequences of random variables with values in $[0, \infty]$. We consider $[0, \infty]$ as the one

point compactification of $[0, \infty)$, i.e., the point $\{\infty\}$ is appended to the set $[0, \infty)$ as the limit point of any increasing and unbounded sequence. Since the set $[0, \infty]$ is compact, any sequence of random variables taking values in this set, and in particular the sequence of stopping times $\{\tau_h, h > 0\}$, is tight.

Theorem 4.3. *Under the assumptions of this section, we have $W^h(x) \rightarrow W(x)$.*

Proof. Consider the pair $(\xi^h(\cdot), \tau_h)$. Let $\{(\xi^h(\cdot), \tau_h), h > 0\}$ be a convergent subsequence, with limit denoted by $(x(\cdot), \bar{\tau})$. Using the Skorokhod representation, we can assume that the convergence is

$$(\xi^h(\cdot), \tau_h) \rightarrow (x(\cdot), \bar{\tau}) \quad (4.7)$$

with probability one. Before we can apply Theorem 1.5 to show $W^h(x) \rightarrow W(x)$, there are several issues to resolve. Recall that $\tau = \inf\{t : x(t) \in \{0, a\}\}$. From the definitions of $W^h(x)$ and $W(x)$ we see that to prove $W^h(x) \rightarrow W(x)$ we will need

$$\bar{\tau} = \tau \quad (4.8)$$

w.p.1. Furthermore, it must be demonstrated that the mapping

$$(\phi(\cdot), t) \rightarrow \int_0^t k(\phi(s))ds + g(\phi(t)) \quad (4.9)$$

is continuous (w.p.1) with respect to the measure induced by $(x(\cdot), \tau)$ on $D[0, \infty) \times [0, \infty]$.

Consider the mapping from $D[0, \infty)$ to $[0, \infty]$ given by

$$\hat{\tau}(\phi(\cdot)) = \inf\{t : \phi(t) \in \{0, a\}\}.$$

Since $k(\cdot)$ and $g(\cdot)$ are continuous and since the sample paths of $x(\cdot)$ are continuous w.p.1, sufficient conditions for the continuity w.p.1 of (4.9) are

$$\hat{\tau}(x(\cdot)) < \infty \text{ and } \hat{\tau}(x(\cdot)) \text{ continuous} \quad (4.10)$$

w.p.1. Furthermore, if (4.10) holds w.p.1, then (4.7) implies

$$\bar{\tau} = \lim_h \tau_h = \lim_h \hat{\tau}(\xi^h(\cdot)) = \hat{\tau}(x(\cdot)) = \tau$$

w.p.1. By the usual argument by contradiction, we have $\tau_h \Rightarrow \tau$ for the original sequence. Thus, (4.10) implies (4.8).

Under (4.10), equation (4.7) implies

$$\int_0^{\tau_h} k(\xi^h(s))ds + g(\xi^h(\tau_h)) \rightarrow \int_0^\tau k(x(s))ds + g(x(\tau))$$

w.p.1. Assume that the collection random variables

$$\left\{ \int_0^{\tau_h} k(\xi^h(s))ds + g(\xi^h(\tau_h)), h \in (0, h_0) \right\} \quad (4.11)$$

is uniformly integrable for some $h_0 > 0$. Then this uniform integrability implies the desired convergence

$$E_x \left[\int_0^{\tau_h} k(\xi^h(s))ds + g(\xi^h(\tau_h)) \right] \rightarrow E_x \left[\int_0^\tau k(x(s))ds + g(x(\tau)) \right].$$

The boundedness of $k(\cdot)$ and $g(\cdot)$ imply that a sufficient condition for the uniform integrability of (4.11) is uniform integrability of

$$\{\tau_h, h \in (0, h_0)\}. \quad (4.12)$$

Except for the proofs of (4.10) and (4.12), the proof of $W^h(x) \rightarrow W(x)$ is complete. ■

Proofs of (4.10) and (4.12). Conditions analogous to (4.10) will appear in virtually all problems we consider with an absorbing boundary. If, in addition, the problem is of interest over a potentially unbounded interval (as here), a condition similar to (4.12) will be required.

We turn first to the proof of (4.10) and consider the continuity of the escape time. We treat only the case $\phi(\hat{\tau}(\phi(\cdot))) = a$, since the case $\phi(\hat{\tau}(\phi(\cdot))) = 0$ is symmetric. Also, we may assume $\phi(\cdot)$ is continuous because $x(\cdot)$ has continuous sample paths w.p.1. Suppose that for all $\delta > 0$ we have $\phi(t) > a$ for some $t \in [\tau, \tau + \delta]$. Then $\phi_n \rightarrow \phi$ implies $\limsup \hat{\tau}(\phi_n(\cdot)) \leq \hat{\tau}(\phi(\cdot))$. On the other hand, the fact that $\phi(t) < a$ for $t < \hat{\tau}(\phi(\cdot))$ implies $\liminf \hat{\tau}(\phi_n(\cdot)) \geq \hat{\tau}(\phi(\cdot))$. Thus, to prove the continuity w.p.1 of $\hat{\tau}(\phi(\cdot))$ under the law of $x(\cdot)$, we need only prove that for all $\delta > 0$, $x(t) > a$ for some $t \in [\tau, \tau + \delta]$, w.p.1. In the present example, this holds because of sample path properties of the stochastic integral, as will now be shown.

We recall the law of the iterated logarithm. If $\tilde{w}(\cdot)$ is a standard Wiener process, then

$$\limsup_{t \downarrow 0} \frac{\tilde{w}(t)}{(2t \log \log t)^{1/2}} = 1$$

w.p.1. Thus, the supremum of $\tilde{w}(s)$ for s in the interval $[0, t]$ behaves like $(2t \log \log t)^{1/2}$ as $t \rightarrow 0$. Suppose $f(t)$ is a Lipschitz continuous function and that $f(0) = 0$. It follows that for all $t_0 > 0$, $w(t) - f(t) > 0$ for some $t \in [0, t_0]$, w.p.1. We next apply this result to $x(\cdot)$. Recall that $a(\cdot) = \sigma^2(\cdot)$. There is a rescaling of time $s \rightarrow t(s)$ satisfying $cs \leq t(s) \leq Cs$ for $c = \inf_{x \in [0, a]} a(x)$ and $C = \sup_{x \in [0, a]} a(x)$, and such that

$$\tilde{w}(s) = \int_\tau^{\tau+t(s)} \sigma(x(r))dw(r)$$

is a standard Wiener process [69, Theorem 3.4.6]. Because $\int_{\tau}^{\tau+t(s)} b(x(r))dr$ is Lipschitz continuous, the law of the iterated logarithm implies $x(t) > a$ w.p.1 for some $t \in [\tau, \tau + \delta]$. Consequently, we have the w.p.1 continuity of $\hat{\tau}(\phi(\cdot))$ under the law of $x(\cdot)$.

Now consider (4.12). Assume there is $T < \infty$ and $h_0 > 0$ such that

$$\inf_{h \in [0, h_0]} \inf_{x \in \{0, h, 2h, \dots, a\}} P_x\{\tau_h < T\} = \delta > 0. \quad (4.13)$$

Then for all $h \in [0, h_0]$ and $x \in \{0, h, 2h, \dots, a\}$, $P_x\{\tau_h \geq T\} \leq (1 - \delta)$. From the Markov property of $\xi^h(\cdot)$, $P_x\{\tau_h \geq iT\} \leq (1 - \delta)^i$ for $i < \infty$. Therefore,

$$E_x(\tau_h)^2 \leq \sum_{i=1}^{\infty} (iT)^2 (1 - \delta)^i < \infty,$$

which implies the uniform integrability of (4.12). The condition (4.13) can be established using only weak convergence arguments and properties of the limit process $x(\cdot)$. First, we note that

$$\inf_{x \in [0, a]} P_x\{\tau < T\} > 0$$

for all $T > 0$. This can be proved in many ways. For example, the inequality can be shown by using the same time change as used in the proof of (4.10) together with the explicit form of the Gaussian distribution. Next, assume that (4.13) is not true. Then there exist $T > 0$, $h_n \rightarrow 0$, and $x_n \rightarrow x \in [0, a]$ such that $\lim_n P_{x_n}\{\tau_{h_n} < T\} = 0$. Consider processes $\xi^{h_n}(\cdot)$ which start at x_n rather than at a fixed point x at time $t = 0$. Because $\{x_n, n < \infty\}$ is compact, the same calculations as those used in the case of a fixed starting position imply $\{\xi^{h_n}(\cdot), n < \infty\}$ is tight, and also that $\xi^{h_n}(\cdot)$ tends weakly to a solution $x(\cdot)$ of (4.1). As shown above, the exit time $\hat{\tau}(\phi(\cdot))$ is a continuous function of the sample paths of $x(\cdot)$ with probability one, so, by Theorem 1.5, $\tau_{h_n} \Rightarrow \tau$. Using part (iii) of Theorem 1.4,

$$\liminf_n P_{x_n}\{\tau_{h_n} < T\} = \liminf_n P_{x_n}\{\tau_{h_n} \in (0, T)\} \geq P_x\{\tau \in (0, T)\} > 0.$$

By contradiction, (4.13) is true. ■

Remarks. In the sequel we will see many variations on the basic ideas of this example. In all cases the required adaptations are suggested directly by the particular properties of the problem under consideration. Here we comment on several aspects of the problem considered in this section.

Part of the power of the approach advocated in this book is the ease with which it handles less stringent assumptions. Aside from weak sense uniqueness assumptions on certain limit processes, which would seem to be expected in any case as a consequence of proper modelling, there is considerable flexibility in weakening the other assumptions. For example,

suppose the condition $\sigma^2(x) \geq c > 0$ on $[0, a]$ is dropped. In order that the process not have any points at which it is “stuck,” let us assume $\inf_{x \in [0, a]} [\sigma^2(x) + |b(x)|] > 0$. Then with some minor modifications and appropriate additional assumptions on $b(\cdot)$, the method still can be used. For example, at points where $\sigma(x) = 0$ the definition (4.4) of the approximation $w^h(\cdot)$ to the limiting Wiener process must be modified. This reflects the fact that we cannot “reconstruct” the Wiener process from the process $x(\cdot)$ at those points. However, minor modifications of the definition of $w^h(\cdot)$ and the associated filtration solve the problem, and the analogous conclusion regarding convergence of $\xi^h(\cdot)$ follows as before. The details may be found in Chapter 10.

If we weaken the nondegeneracy assumption on $\sigma(\cdot)$, then we must also reconsider the proofs of (4.10) and (4.12). These conditions are not simply technical nuisances whose validity is not directly related to the convergence of the schemes. For example, suppose $\sigma(x) = 0$ in neighborhoods of both 0 and a . Suppose also that $b(a) \leq 0$. Then clearly the process $x(\cdot)$ does not exit through the point a . An analogous statement holds if $b(0) \geq 0$. In such a case, (4.12) clearly fails. However, it will also be true that for reasonable choices of $k(\cdot)$ that $W(x) = \infty$. Thus, we can have seemingly reasonable approximations (i.e., local consistency holds) and only the failure of (4.12) indicates a difficulty with problem formulation. Precise conditions for such degenerate problems and verification of the analogues of (4.10) and (4.12) will be described in detail in Chapter 10.

9.5 Relaxed Controls

In the last section it was demonstrated that the convergence properties of suitably scaled and interpolated Markov chains could be used to establish the convergence of a numerical scheme corresponding to an uncontrolled diffusion. For the analogous problem involving controlled processes, it will be necessary to consider sequences of controlled Markov chains. In this case, besides the convergence properties of the processes we must also deal with the convergence of the controls. When we choose an approximating Markov chain, we essentially force the limit process to be of the desired type by building the limiting properties directly into the chain, e.g., the consistency conditions. However, in general, we can do little to force the limits of the controls that are optimal for the chains to take any preassigned form, such as feedback. This does not reflect any technical shortcoming, but is in part due to the fact that the infimum will not be attained in many optimal control problems within a given class of controls unless it is compact in an appropriate sense. An example of such lack of compactness was given in Section 4.6 for a deterministic optimal control problem.

We will often take simultaneous limits of processes and controls in the convergence proofs that are to come. In order to guarantee the existence

of limits, it will be necessary to work with a space of controls that have the appropriate closure property and which yield the same minimum value function for the optimization problem we seek to approximate. The *relaxed controls* form a class with such properties. Relaxed controls were first introduced by L. C. Young to establish existence of a (generalized) minimizing control for problems from the calculus of variations [127]. They were later extended to the stochastic setting [45] and have since found several uses, especially with regard to convergence properties [81]. The controls computed for the approximating numerical problem are always of the feedback form.

Deterministic Relaxed Controls. Consider a compact subset \mathcal{U} of some finite dimensional Euclidean space. For convenience we shall assume that each control we deal with is defined on the interval $[0, \infty)$, although it may actually only be applied on some bounded subset. We recall some of the definitions and notation introduced in Section 4.6. The σ -algebras $\mathcal{B}(\mathcal{U})$ and $\mathcal{B}(\mathcal{U} \times [0, \infty))$ are defined as the collection of Borel subsets of \mathcal{U} and $\mathcal{U} \times [0, \infty)$, respectively. A relaxed control is then a Borel measure $m(\cdot)$ such that $m(\mathcal{U} \times [0, t]) = t$ for all $t \geq 0$. We can define a derivative $m_t(\cdot)$, such that

$$m(B) = \int_{\mathcal{U} \times [0, \infty)} I_{\{(u, t) \in B\}} m_t(d\alpha) dt \quad (5.1)$$

for all $B \in \mathcal{B}(\mathcal{U} \times [0, \infty))$ [i.e., $m(d\alpha dt) = m_t(d\alpha)dt$] and such that for each t , $m_t(\cdot)$ is a measure on $\mathcal{B}(\mathcal{U})$ satisfying $m_t(\mathcal{U}) = 1$. For example, we can define $m_t(\cdot)$ in any convenient way for $t = 0$ and as the left hand derivative for $t > 0$:

$$m_t(A) = \lim_{\delta \rightarrow 0} \frac{m(A \times [t - \delta, t])}{\delta}$$

for $A \in \mathcal{B}(\mathcal{U})$. Let $\mathcal{R}(\mathcal{U} \times [0, \infty))$ denote the set of all relaxed controls on $\mathcal{U} \times [0, \infty)$.

A Topology for the Space of Relaxed Controls. The space $\mathcal{R}(\mathcal{U} \times [0, \infty))$ can be metrized in a convenient way in terms of the Prohorov metric of Section 9.1. Let $\pi^T(\cdot)$ denote the Prohorov metric on $\mathcal{P}(\mathcal{U} \times [0, T])$. For $m_1, m_2 \in \mathcal{R}(\mathcal{U} \times [0, \infty))$, define $\bar{m}_i^T, i = 1, 2$, to be the normalized restrictions of these measures to $\mathcal{U} \times [0, T]$; i.e.,

$$\bar{m}_i^T(B) = \frac{1}{T} m_i(B)$$

for all $B \in \mathcal{B}(\mathcal{U} \times [0, T])$. Thus, each \bar{m}_i^T is always a probability measure. We define a metric on $\mathcal{R}(\mathcal{U} \times [0, \infty))$ by

$$d(m_1, m_2) = \sum_{j=1}^{\infty} 2^{-j} \pi^j(\bar{m}_1^j, \bar{m}_2^j).$$

Under this metric, a sequence $m_n(\cdot)$ in $\mathcal{R}(\mathcal{U} \times [0, \infty))$ converges to $m(\cdot) \in \mathcal{R}(\mathcal{U} \times [0, \infty))$ if and only if $\bar{m}_n^j \Rightarrow \bar{m}^j$ for all $j < \infty$. Note that this is equivalent to

$$\int \phi(\alpha, s)m_n(d\alpha ds) \rightarrow \int \phi(\alpha, s)m(d\alpha ds)$$

for any continuous function $\phi(\cdot, \cdot)$ on $\mathcal{U} \times [0, \infty)$ having compact support. Since $\mathcal{P}(\mathcal{U} \times [0, j])$ is complete, separable, and compact for all $j < \infty$, these properties are inherited by the space $\mathcal{R}(\mathcal{U} \times [0, \infty))$ under this metric. It follows that any sequence of relaxed controls has a convergent subsequence. This key property will be used often in the sequel. We will write $m_n \Rightarrow m$ for convergence in this “weak–compact” topology.

Relation to Ordinary Deterministic Controls. We recall the problem considered in Section 4.5. Minimize the cost

$$W(x, u) = \int_0^\infty e^{-\beta s} k(x(s), u(s))ds$$

over all \mathcal{U} –valued measurable control processes $u(\cdot)$, where $x(\cdot)$ solves

$$\dot{x}(t) = b(x(t), u(t)), \quad x(0) = x.$$

Recall also that $\beta > 0$, and that $b(\cdot)$ and $k(\cdot)$ are assumed bounded and continuous. As in Section 4.6, any ordinary control $u(\cdot)$ has the relaxed control representation $m(\cdot)$, where we define

$$m(A \times [0, t]) = \int_0^t I_A(u(s))ds,$$

in the sense that the cost function and dynamics may be rewritten as

$$W(x, u) = W(x, m) = \int_0^\infty \int_{\mathcal{U}} e^{-\beta s} k(x(s), \alpha)m(d\alpha ds), \quad (5.2)$$

$$x(t) = x + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha)m(d\alpha ds). \quad (5.3)$$

Let $m^n(\cdot)$ denote a sequence of relaxed controls such that the associated costs converge to the infimum over all relaxed controls. Compactness of the space of $\mathcal{R}(\mathcal{U} \times [0, \infty))$ and boundedness of $b(\cdot)$ imply relative compactness of the sequence $(x^n(\cdot), m^n(\cdot))$, where $x^n(\cdot)$ is the solution to the controlled ODE under $m^n(\cdot)$. Suppose we retain n as the index of a convergent subsequence and suppose also that the limit is denoted by $(x(\cdot), m(\cdot))$. Owing to the weak convergence $m^n(\cdot) \Rightarrow m(\cdot)$, $x(\cdot)$ solves (5.3) with the cost (5.2). Thus, the infimum over the relaxed controls is always attained. By the approximation theorem of Section 4.6 any relaxed control may be

arbitrarily well approximated by an ordinary control in the sense that the costs under the two controls may be made arbitrarily close. It follows that the infimum over the class of relaxed controls is the same as that over the ordinary controls.

Stochastic Relaxed Controls. A stochastic relaxed control will be a control process that is a deterministic relaxed control for each element of the underlying probability space, and which also satisfies the nonanticipativeness condition usually assumed of ordinary stochastic controls. Recall that our primary interest is in weak sense solutions of controlled stochastic differential equations. Suppose we are given a probability space (Ω, \mathcal{F}, P) , a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process $w(\cdot)$, and an \mathcal{F}_t -Poisson random measure $N(\cdot)$. Then we say $m(\cdot)$ is an admissible relaxed control for the pair $(w(\cdot), N(\cdot))$, or that the triple $(m(\cdot), w(\cdot), N(\cdot))$ is admissible, if $m(\cdot, \omega)$ is a deterministic relaxed control with probability one and if $m(A \times [0, t])$ is \mathcal{F}_t -adapted for all $A \in \mathcal{B}(\mathcal{U})$. There exists a derivative $m_t(\cdot)$ such that $m_t(A)$ is \mathcal{F}_t -adapted for all $A \in \mathcal{B}(\mathcal{U})$ and such that (5.1) holds with probability one [81]. Because the space $\mathcal{R}(\mathcal{U} \times [0, \infty))$ is compact, any collection of relaxed controls in this space is tight. The definitions of weak existence and weak uniqueness of controlled processes under relaxed controls are analogous to those given in Section 1.3 for ordinary stochastic controls. As noted in that section, the techniques and assumptions that are needed to prove weak existence and weak uniqueness of solutions to a SDE with an admissible relaxed control are essentially the same as those for the case without control.

The following convention will simplify the notation in later chapters. Let $m(\cdot)$ be an admissible relaxed control. For $t \geq 0$ let $m(t, \cdot)$ be the random measure with values $m(t, A) = m(A \times [0, t])$ for $A \in \mathcal{B}(\mathcal{U})$. This definition involves an abuse of notation, but it should not cause any confusion.

10

Convergence Proofs

This chapter is the core of the mathematical part of the book. It deals with the approximation and convergence theorems for the basic problem classes: discounted problems with absorbing boundaries; diffusion and jump diffusion models; optimal stopping problems, and problems where we stop on hitting a target set and where there is no discounting. The convergence results for the case of reflecting boundaries and the singular and ergodic control problems will appear in the next chapter.

The chapter starts off with some approximation and limit results for a sequence of controlled jump diffusion problems. These results and the methods which are used provide a base for the later proofs of the convergence of the numerical approximations. The first result, Theorem 1.1, shows that the limit of a sequence of controlled jump diffusions is also a controlled jump diffusion. The method of proof is more direct than the usual martingale problem approach, because we have access to the driving Wiener processes and Poisson measures. The approach is a combination of a classical weak convergence method together with a direct construction. The main points of the theorem are the “admissibility” of the limit controls and the stopping or exit times. The theorem provides the basis for the approximation of relaxed controls by simpler controls. These simpler controls will be applied to the approximating chains to get the convergence results for the numerical methods. In particular, Theorem 1.2 shows that we can approximate a relaxed control by a piecewise constant control which takes values in a finite set.

Theorem 2.1 concerns convergence of sequences of controlled problems to a limit problem, when control stops at the first moment that a given set is exited. It is essentially a consequence of Theorem 1.1, except for the problem of the behavior of the limit path at the first hitting time of the boundary. There is a discussion of this point, and the concept of “randomized stopping” is introduced. Analogous “boundary” problems will appear when dealing with the Markov chain approximations.

Theorem 3.1 shows that we can approximate the optimal control by a “nice” control, which is an appropriate function of the driving Wiener process and Poisson measure at a finite number of time points, and which is continuous in the values of the Wiener process at those time points. Such an approximation will be needed to show that the sequence of optimal costs $V^h(x)$ converges to the optimal cost $V(x)$.

Section 10.4 shows that weak limits of the sequence $\psi^h(\cdot)$ (defined in Subsection 10.4.1 below and Section 4.3) of Markov process interpolations of the approximating Markov chains are actually controlled jump diffusions, and shows that $\liminf_h V^h(x) \geq V(x)$. The proof introduces auxiliary processes $w^h(\cdot)$ and $N^h(\cdot)$ which are approximations to the “limit” Wiener process and Poisson measure. The $\psi^h(\cdot)$ are represented in terms of these auxiliary processes, and this facilitates getting the desired limit. Section 10.5 is concerned with convergence of the costs to an optimal cost. To do this we need to show that $\limsup_h V^h(x) \leq V(x)$. This is done by using the particular form of the ϵ -optimal control for the limit problem which is continuous in the sampled values of $w(\cdot)$, as derived in Section 10.3, together with the optimality of $V^h(x)$ for the approximating chain. In Section 10.6, the convergence of the numerical approximations for the optimal stopping problem is proved.

Only some selected problem classes are treated, but it will be seen that the methods are quite general and have much use elsewhere.

Local consistency is not always possible to get at all points, as seen in the treatment of grid refinement in Section 5.5. It is shown in Theorem 5.3 that one can still get convergence of the numerical algorithms under quite broad conditions.

10.1 Limit Theorems and Approximations of Relaxed Controls

10.1.1 LIMIT OF A SEQUENCE OF CONTROLLED DIFFUSIONS

This section will be the first application of the weak convergence ideas of Chapter 9 and will establish some methods which will be needed later. We will be given a sequence of controlled jump diffusions, each one being “driven” by a possibly different control and Wiener process and Poisson measure. It will be shown that the sequence of solutions and driving processes is tight, and that the limit of any weakly convergent subsequence is a controlled jump diffusion process. Under a uniqueness condition, it will follow from this that a relaxed control can be approximated by an ordinary control which is piecewise constant. This latter result was first shown in [45] and can also be found in [87]. We will then use this result to show that, under a uniqueness condition, a relaxed control can be approximated by an ordinary control which is piecewise constant and takes only finitely many values. The approximation is in the sense of the topology of relaxed controls as well as the closeness of the associated cost functions. These methods will be used later to show convergence of the numerical approximations. The weak convergence terminology will be that of Chapter 9.

The method used in Theorem 1.1 is similar to the method discussed in Section 9.4. It uses a simple “continuity argument” to simplify the more traditional proofs using the so-called martingale formulation. It is more direct than the latter method for our type of problem. Similar methods will be used when proving the convergence $V^h(x) \rightarrow V(x)$ in Sections 10.4 to 10.6.

We will use the controlled jump diffusion model of Section 1.5

$$\begin{aligned} x(t) = & x + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m_s(d\alpha) ds + \int_0^t \sigma(x(s)) dw(s) \\ & + \int_0^t \int_{\Gamma} q(x(s^-), \rho) N(ds d\rho). \end{aligned} \quad (1.1)$$

When writing (1.1) or referring to a solution of (1.1), it is always implied that there exists a probability space, a filtration \mathcal{F}_t , and processes $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$ such that $w(\cdot)$ is a standard \mathcal{F}_t -Wiener process, $N(\cdot)$ is an \mathcal{F}_t -Poisson measure with jump rate λ and jump distribution $\Pi(\cdot)$, $m(\cdot)$ is admissible with respect to $(w(\cdot), N(\cdot))$ [i.e., $m(\cdot)$ is \mathcal{F}_t -adapted], and the solution $x(\cdot)$ is \mathcal{F}_t -adapted. Such a Wiener process and Poisson measure will always be referred to as *standard*. Existence is always in the weak sense, because the probability space and the $w(\cdot), N(\cdot)$ are not specified a priori.

Recall the following definition.

Definition. By weak sense uniqueness for an initial condition x , we mean that the probability law of an admissible triple determines the probability law of any solution $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$ to (1.1) irrespective of the probability space.

We will use the assumptions:

A1.1. *b(·) and σ(·) are bounded and continuous.*

A1.2. *q(·) is measurable and bounded, q(·, ρ) is continuous for each ρ.*

A1.3. *k(·) and g(·) are bounded and continuous.*

A1.4. *Let u(·) be an admissible ordinary control with respect to $(w(\cdot), N(\cdot))$, and suppose that u(·) is piecewise constant and takes only a finite number of values. Then, for each initial condition, there exists a solution to (1.1), where m(·) is a relaxed control representation of u(·) and this solution is unique in the weak sense.*

Remark. (A1.4) is equivalent to the assumption of weak sense existence and uniqueness for each initial condition and constant control.

It is worth noting that the continuity conditions in (A1.1)-(A1.3) can be weakened and it is often important to do so. We will comment on this at the end of Theorem 1.1. Recall the definition $N(t, A) = N([0, t] \times A)$, the number of jumps with values in A by time t . Write $N(t) = N(t, \cdot)$ for simplicity. For future use, note that for each t , $N(t)$ can be constructed from the two processes

$$N(s, \Gamma), \quad \bar{N}(s) = \int_0^s \int_{\Gamma} \rho N(drd\rho), \quad s \leq t. \quad (1.2)$$

The first member is just a Poisson process with rate λ , and it identifies the jump times. The second member can be used to get the jump values.

Notation in Theorem 1.1. In the next theorem, we assume that for each integer n there is a probability space on which are defined a filtration \mathcal{F}_t^n , a pair of processes $(w^n(\cdot), N^n(\cdot))$, an admissible relaxed control $m^n(\cdot)$, and a solution process $x^n(\cdot)$. The $w^n(\cdot), N^n(\cdot)$ are a standard \mathcal{F}_t^n -Wiener process and an \mathcal{F}_t^n -Poisson measure, resp. The filtration satisfies

$$\mathcal{F}_t^n \supset \mathcal{F}(x^n(s), m^n(s), w^n(s), N^n(s), s \leq t).$$

Thus, (1.3) holds.

$$\begin{aligned} x^n(t) &= x^n(0) + \int_0^t \int_{\mathcal{U}} b(x^n(s), \alpha) m_s^n(d\alpha) ds \\ &\quad + \int_0^t \sigma(x^n(s)) dw^n(s) + \int_0^t \int_{\Gamma} q(x^n(s^-), \rho) N^n(ds d\rho) \end{aligned} \quad (1.3)$$

Remark on the Notation in the Theorem Statement. Note that we let the probability space vary with the control. The basic reason for this is that we are working with weak sense solutions and cannot always define the solution process as an explicit function of the “driving forces,” which are the control, the Wiener process, and the Poisson measure. First, consider the most classical case, where the functions $b(\cdot, \alpha)$ and $\sigma(\cdot)$ satisfy a Lipschitz condition [which is assumed to be uniform in α for $b(\cdot, \alpha)$]. Then, given an admissible triple $(m(\cdot), w(\cdot), N(\cdot))$, one can construct a strong solution $x(\cdot)$ on the same probability space and as an explicit function of the triple. However, we want to work with the largest class of controls possible, provided that they are admissible. It is not a priori obvious that the optimal control will in all cases be representable as a function of only $(w(\cdot), N(\cdot))$. Thus, we might at least have to augment the probability space, in a way that depends on the control. It should be understood that the underlying Wiener process and Poisson measure are fictions to a large extent. They are very useful to represent and study the processes, but when calculating cost

functions and their limits, only the distributions of the processes are important, and not the actual probability spaces or the representation of the solution to the SDE. Under the Lipschitz condition, given the probability law of $(m(\cdot), w(\cdot), N(\cdot))$, and the initial condition x , the probability law of $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$ is uniquely determined, and that probability law is all that is important in computing the cost functions. Let $(m^n(\cdot), w(\cdot), N(\cdot))$ be a sequence of admissible triples, all defined on the same probability space with the same standard Wiener process and Poisson measure. The sequence (or some subsequence) will not generally converge to a limit with probability one (or even in probability) on the original probability space, and weak convergence methods might have to be used to get appropriate limits. But then the “limit processes” will not, in general, be definable on the original probability space either. Because we do not want to worry about the actual probability space, we often let it vary with the control.

If the Lipschitz condition does not hold, but the Girsanov transformation method is used to get the controlled process from an uncontrolled process via a transformation of the measure on the probability space, then the Wiener process is not fixed a priori, and its construction depends on the control. The considerations raised in the last paragraph also hold here. These comments provide a partial explanation of the indexing of the Wiener process and Poisson measure by n .

Theorem 1.1. *Assume (A1.1) and (A1.2). Let $x^n(0) \Rightarrow x_0$ and let ν^n be an \mathcal{F}_t^n -stopping time. Then $\{x^n(\cdot), m^n(\cdot), w^n(\cdot), N^n(\cdot), \nu^n\}$ is tight. Let $(x(\cdot), m(\cdot), w(\cdot), N(\cdot), \nu)$ denote the limit of a weakly convergent subsequence. Define*

$$\mathcal{F}_t = \mathcal{F}(x(s), m(s), w(s), N(s), \nu I_{\{\nu \leq t\}}, s \leq t).$$

Then $w(\cdot)$ and $N(\cdot)$ are a standard \mathcal{F}_t -Wiener process and \mathcal{F}_t -Poisson measure, resp., ν is an \mathcal{F}_t -stopping time, $m(\cdot)$ is admissible with respect to $(w(\cdot), N(\cdot))$, $x(0) = x_0$, and $x(\cdot)$ satisfies (1.1).

Proof. *Tightness.* The criterion of Theorem 9.2.1 will be used. Let $T < \infty$, and let $\tilde{\nu}_n$ be an arbitrary \mathcal{F}_t^n -stopping time satisfying $\tilde{\nu}_n \leq T$. Then, by the properties of the stochastic integral and the boundedness of $q(\cdot)$ and the jump rate λ ,

$$E[|x^n(\tilde{\nu}_n + \delta) - x^n(\tilde{\nu}_n)|^2 | \mathcal{F}_{\tilde{\nu}_n}^n] = O(\delta),$$

where the order $O(\delta)$ is uniform in $\tilde{\nu}_n$. Thus, by Theorem 9.2.1, the sequence $\{x^n(\cdot)\}$ is tight. The sequences of controls $\{m^n(\cdot)\}$ and stopping times $\{\nu^n\}$ are tight because their range spaces $[\mathcal{R}(\mathcal{U} \times [0, \infty))$ and $[0, \infty]$, resp.] are compact. Clearly $\{w^n(\cdot), N^n(\cdot)\}$ is tight and any weak limit has the same law as each of the $(w^n(\cdot), N^n(\cdot))$ pairs has.

Characterization of the limit processes. Now that we have tightness, we can extract a weakly convergent subsequence and characterize its limit.

For notational convenience, let the original sequence converge weakly, and denote the limit by $(x(\cdot), m(\cdot), w(\cdot), N(\cdot), \nu)$. Because the processes $w^n(\cdot)$ have continuous paths with probability one, so will $w(\cdot)$. It follows from the weak convergence that $m(t, \mathcal{U}) = t$ for all t . We want to show that the limit $x(\cdot)$ is a solution to a stochastic differential equation with driving processes $(m(\cdot), w(\cdot), N(\cdot))$. This will be done by a combination of a fairly direct method and a use of the martingale method.

Let $\delta > 0$ and let k be a positive integer. For any process $z(\cdot)$ with paths in $D^k[0, \infty)$, define the piecewise constant process $z_\delta(\cdot)$ by

$$z_\delta(t) = z(i\delta), \quad t \in [i\delta, i\delta + \delta).$$

By the tightness of $\{x^n(\cdot)\}$ and the boundedness and continuity in (A1.1), (A1.2), we can write (1.3) as

$$\begin{aligned} x^n(t) &= x^n(0) + \int_0^t \int_{\mathcal{U}} b(x^n(s), \alpha) m_s^n(d\alpha) ds \\ &\quad + \int_0^t \sigma(x_\delta^n(s)) dw^n(s) + \int_0^t \int_{\Gamma} q(x_\delta^n(s^-), \rho) N^n(ds d\rho) + \varepsilon_{\delta, t}^n \end{aligned} \tag{1.4}$$

where

$$E|\varepsilon_{\delta, t}^n| \xrightarrow{\delta} 0$$

uniformly in n and in t in any bounded interval.

For the rest of the proof, we assume that the probability spaces are chosen as required by the Skorokhod representation (Theorem 9.1.7), so that we can suppose that the convergence of $\{x^n(\cdot), m^n(\cdot), w^n(\cdot), N^n(\cdot), \nu^n\}$ to its limit is with probability one in the topology of the path spaces of the processes. Thus,

$$\int_0^t \int_{\mathcal{U}} b(x^n(s), \alpha) m_s^n(d\alpha) ds - \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m_s(d\alpha) ds \xrightarrow{n} 0,$$

uniformly on any bounded time interval with probability one. The sequence $\{m^n(\cdot)\}$ converges in the “compact-weak” topology. In particular, for any continuous and bounded function $\phi(\cdot)$ with compact support,

$$\int_0^\infty \int_{\mathcal{U}} \phi(\alpha, s) m^n(d\alpha ds) \rightarrow \int_0^\infty \int_{\mathcal{U}} \phi(\alpha, s) m(d\alpha ds).$$

Now the Skorokhod representation and weak convergence imply that

$$\int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m_s^n(d\alpha) ds - \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m_s(d\alpha) ds \xrightarrow{n} 0$$

uniformly in t on any bounded interval with probability one.

Owing to the fact that the $x_\delta^n(\cdot)$ are constant on the intervals $[i\delta, i\delta + \delta)$, the third and fourth terms on the right side of (1.4) converge to, resp.,

$$\int_0^t \sigma(x_\delta(s)) dw(s),$$

$$\int_0^t \int_{\Gamma} q(x_\delta(s^-), \rho) N(ds d\rho),$$

which are well defined with probability one since they can be written as finite sums. Thus, we can write

$$\begin{aligned} x(t) &= x_0 + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m_s(d\alpha) ds \\ &\quad + \int_0^t \sigma(x_\delta(s)) dw(s) + \int_0^t \int_{\Gamma} q(x_\delta(s^-), \rho) N(ds d\rho) + \varepsilon_{\delta,t}, \end{aligned} \tag{1.5}$$

where $E|\varepsilon_{\delta,t}| \xrightarrow{\delta} 0$, uniformly on bounded t -intervals.

We next characterize $w(\cdot)$. We know that it is a Wiener process, but we need to show that it is a \mathcal{F}_t -Wiener process. Let $H(\cdot)$ be a real-valued and continuous function of its arguments, and with compact support. Let $\phi(\cdot)$ and $\phi_j(\cdot)$ be real-valued and continuous functions of their arguments and with compact support. Define the function¹

$$(\phi, m)_t = \int_0^t \int_{\mathcal{U}} \phi(\alpha, s) m(d\alpha ds).$$

Let $p, t, u, t_i, i \leq p$, be given such that $t_i \leq t \leq t + u$, $i \leq p$ and $P\{\nu = t_i\} = 0$. For $q = 1, 2, \dots$, let $\{\Gamma_j^q, j \leq q\}$ be a sequence of nondecreasing partitions of Γ such that $\Pi(\partial\Gamma_j^q) = 0$ for all j and all q , where $\partial\Gamma_j^q$ is the boundary of the set Γ_j^q . As $q \rightarrow \infty$, let the diameters of the sets Γ_j^q go to zero. Let the Wiener processes be $\mathbb{R}^{r'}$ -valued, and let $f(\cdot) \in C_0^2(\mathbb{R}^{r'})$.

Define the differential operator \mathcal{L}_w of the Wiener process

$$\mathcal{L}_w f(w) = \frac{1}{2} \sum_{i=1}^{r'} f_{w_i w_i}(w).$$

¹The function $(\phi, m)_t$ of $m(\cdot)$ is introduced, since any continuous function of $m(\cdot)$ can be arbitrarily well approximated by continuous functions of the type

$$h((\phi_j, m)_{t_i}, j \leq q, i \leq p),$$

for appropriate $\{t_i\}$ and continuous $h(\cdot)$ and $\phi_j(\cdot)$ with compact support.

By the fact that $w^n(\cdot)$ is a \mathcal{F}_t^n -Wiener process, we have

$$\begin{aligned} & EH(x^n(t_i), w^n(t_i), (\phi_j, m^n)_{t_i}, N^n(t_i, \Gamma_j^q), j \leq q, i \leq p, \nu^n I_{\{\nu^n \leq t\}}) \\ & \times \left[f(w^n(t+u)) - f(w^n(t)) - \int_t^{t+u} \mathcal{L}_w f(w^n(s)) ds \right] = 0. \end{aligned} \quad (1.6)$$

By the probability one convergence which is implied by the Skorokhod representation

$$E \left| \int_t^{t+u} \mathcal{L}_w f(w^n(s)) ds - \int_t^{t+u} \mathcal{L}_w f(w(s)) ds \right| \rightarrow 0.$$

Using this result and taking limits in (1.6) yields

$$\begin{aligned} & EH(x(t_i), w(t_i), (\phi_j, m)_{t_i}, N(t_i, \Gamma_j^q), j \leq q, i \leq p, \nu I_{\{\nu \leq t\}}) \\ & \times \left[f(w(t+\tau)) - f(w(t)) - \int_t^{t+\tau} \mathcal{L}_w f(w(s)) ds \right] = 0. \end{aligned} \quad (1.7)$$

The set of random variables

$$H(x(t_i), w(t_i), (\phi_j, m)_{t_i}, N(t_i, \Gamma_j^q), j \leq q, i \leq p, \nu I_{\{\nu \leq t\}}),$$

as $H(\cdot), p, q, \phi_j(\cdot), \Gamma_j^q$, and t_i vary over all possibilities, induces the σ -algebra \mathcal{F}_t . Thus, (1.7) implies that

$$f(w(t)) - \int_0^t \mathcal{L}_w f(w(s)) ds$$

is an \mathcal{F}_t -martingale for all $f(\cdot)$ of the chosen class. Thus, $w(\cdot)$ is a standard \mathcal{F}_t -Wiener process.

We now turn our attention to showing that $N(\cdot)$ is an \mathcal{F}_t -Poisson measure. Let $\theta(\cdot)$ be a continuous function on Γ , and define the process

$$\Theta_N(t) = \int_0^t \int_\Gamma \theta(\rho) N(ds d\rho).$$

By an argument which is similar to that used for the Wiener process above, if $f(\cdot)$ is a continuous function with compact support, then

$$\begin{aligned} & EH(x(t_i), w(t_i), (\phi_j, m)_{t_i}, N(t_i, \Gamma_j^q), j \leq q, i \leq p, \nu I_{\{\nu \leq t\}}) \\ & \times \left[f(\Theta_N(t+u)) - f(\Theta_N(t)) \right. \\ & \left. - \lambda \int_t^{t+u} \int_\Gamma [f(\Theta_N(s) + \theta(\rho)) - f(\Theta_N(s))] \Pi(d\rho) ds \right] = 0. \end{aligned} \quad (1.8)$$

Equation (1.8) and the arbitrariness of $H(\cdot), p, q, t_i, \phi_j(\cdot), \Gamma_j^q, f(\cdot)$, and $\theta(\cdot)$ imply that $N(\cdot)$ is an \mathcal{F}_t -Poisson measure.

The facts that $w(\cdot)$ and $N(\cdot)$ are an \mathcal{F}_t -Wiener process and Poisson measure, resp., implies their mutual independence. By construction, $\{\nu \leq t\} \subset \mathcal{F}_t$, which implies that ν is a \mathcal{F}_t -stopping time. Since for each t , \mathcal{F}_t measures $\{x(s), m(s), w(s), N(s), s \leq t\}$, the control $m(\cdot)$ is admissible and $x(\cdot)$ is nonanticipative with respect to $(w(\cdot), N(\cdot))$. We can now take limits as $\delta \rightarrow 0$ in (1.5) to complete the proof of the theorem. ■

Remark on Discontinuous Dynamical Terms. Theorem 9.1.5 allows us to weaken the conditions on $b(\cdot)$ and $a(\cdot)$. Suppose that $(x^n(\cdot), m^n(\cdot), w^n(\cdot), N^n(\cdot))$ converges weakly to $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$. Let us use the Skorokhod representation so that we can assume probability one convergence. Even in the absence of continuity of the $b(\cdot)$ and $a(\cdot)$, if for each t

$$\begin{aligned}\int_0^t b(x^n(s), \alpha) ds &\rightarrow \int_0^t b(x(s), \alpha) ds, \\ \int_0^t a(x^n(s)) ds &\rightarrow \int_0^t a(x(s)) ds\end{aligned}$$

uniformly in α with probability one, then (1.1) will hold. By Theorem 9.1.5, it is enough to assume that the function $b(\phi(\cdot), \alpha)$ from $D^r[0, \infty)$ to $D^r[0, \infty)$ will be continuous with probability one (uniformly in α) with respect to the measure induced by the limit process $x(\cdot)$, and with an analogous assumption for $a(\cdot)$. For the control problems below there are analogous weakenings of the conditions on the cost functions $k(\cdot)$ and $g(\cdot)$. Similar considerations apply to the convergence of the numerical approximations. See, for example, Theorem 5.3.

10.1.2 AN APPROXIMATION THEOREM FOR RELAXED CONTROLS

Suppose that $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$ solves (1.1), where the relaxed control $m(\cdot)$ is admissible with respect to $(w(\cdot), N(\cdot))$. In some of the technical arguments in the sequel it is important to be able to approximate a relaxed control by a simple ordinary control. Under a uniqueness condition, this can be done with the approximation arguments of Theorem 1.1, with appropriate choices of the $m^n(\cdot)$.

The proof of the following theorem is in [81, Theorem 3.5.2]. In the theorem, the processes indexed by ϵ and the limit process are assumed to be defined on the same probability space, but this can always be done via the Skorokhod representation. The value of the cost function depends on the joint distribution of $(x(\cdot), m(\cdot))$. In order to simplify the notation, we write the cost function only as a function of $m(\cdot)$ and the initial condition x . The brevity of the notation should not cause any confusion.

Theorem 1.2. Assume (A1.1) – (A1.4) and let $x(0) = x$. For a given admissible triple $(m(\cdot), w(\cdot), N(\cdot))$, let the solution $x(\cdot)$ to (1.1) exist and be unique in the weak sense. For any finite T and $\beta > 0$, define the cost functions

$$W_T(x, m) = E_x^m \int_0^T \int_{\mathcal{U}} k(x(s), \alpha) m(d\alpha ds) + E_x^m g(x(T)), \quad (1.9)$$

$$W(x, m) = E_x^m \int_0^\infty \int_{\mathcal{U}} e^{-\beta s} k(x(s), \alpha) m(d\alpha ds). \quad (1.10)$$

Given $\epsilon > 0$, there is a finite set $\{\alpha_1^\epsilon, \dots, \alpha_{p_\epsilon}^\epsilon\} = \mathcal{U}_\epsilon \subset \mathcal{U}$, and a $\delta > 0$ with the following properties. There is a probability space on which are defined processes

$$(x^\epsilon(\cdot), u^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot)), \quad (1.11)$$

where $w^\epsilon(\cdot)$ and $N^\epsilon(\cdot)$ are our standard Wiener process and Poisson measure, resp., and $u^\epsilon(\cdot)$ is an admissible \mathcal{U}_ϵ -valued ordinary control which is constant on the intervals $[i\delta, i\delta + \delta]$. Furthermore, the processes (1.11) satisfy (1.1) and

$$\begin{aligned} P_x^m \left\{ \sup_{t \leq T} |x^\epsilon(t) - x(t)| > \epsilon \right\} &\leq \epsilon \\ |W(x, m) - W(x, u^\epsilon)| &\leq \epsilon, \\ |W_T(x, m) - W_T(x, u^\epsilon)| &\leq \epsilon. \end{aligned} \quad (1.12)$$

10.2 Existence of an Optimal Control: Absorbing Boundary

By Theorem 1.1, there exists an optimal control for the cost functions (1.9) and (1.10) because Theorem 1.1 shows that the limit of a weakly convergent minimizing sequence of controlled processes is also a controlled process, and the sequence of costs converges to that of the limit process. When the controlled process is stopped at the time of first exit from some set or first entrance into some target or absorbing set, an additional complication arises, because the first exit times of the sequence of approximating processes might not converge to the first exit time of the limit process. Similar questions will arise when dealing with the convergence of the costs for the numerical algorithm. This problem was also discussed in Chapter 9. In this section, we will work with the cost function

$$W(x, m) = E_x^m \int_0^\tau \int_{\mathcal{U}} e^{-\beta s} k(x(s), \alpha) m_s(d\alpha) ds + E_x^m e^{-\beta \tau} g(x(\tau)), \quad (2.1)$$

where $\beta \geq 0$ and τ is the first escape time from the set G^0 , the interior of the set G satisfying (A2.1) below. The discussion concerning continuity of the escape time below is applicable to general cost functions and not just to the discounted cost function.

Using the notation of Theorem 1.1, let $\{x^n(\cdot), m^n(\cdot), w^n(\cdot), N^n(\cdot), \tau^n\}$ be a minimizing sequence, that is,

$$W(x, m^n) \downarrow V(x) \quad (2.2)$$

and $\tau^n = \inf\{t : x^n(t) \notin G^0\}$. By the results of Theorem 1.1 we know that $\{x^n(\cdot), m^n(\cdot), w^n(\cdot), N^n(\cdot), \tau^n\}$ has a weakly convergent subsequence. For notational simplicity, suppose that the original sequence itself converges weakly and that $(x(\cdot), m(\cdot), w(\cdot), N(\cdot), \tilde{\tau})$ denotes the limit. Define the filtration

$$\mathcal{F}_t = \mathcal{F}(x(s), m(s), w(s), N(s), s \leq t, \tilde{\tau} I_{\{\tilde{\tau} \leq t\}}).$$

Then we know from Theorem 1.1 that $(w(\cdot), N(\cdot))$ are a standard \mathcal{F}_t -Wiener process and Poisson measure, resp. Also, $m(\cdot)$ is admissible, (1.1) holds and $\tilde{\tau}$ is an \mathcal{F}_t -stopping time. If either $\beta > 0$ or the $\{\tau^n\}$ are uniformly integrable, then under the continuity and boundedness of $k(\cdot)$ and $g(\cdot)$, it is always the case that

$$\begin{aligned} E_x^{m^n} \int_0^{\tau^n} \int_U e^{-\beta s} k(x^n(s), \alpha) m_s^n(d\alpha) ds \\ \rightarrow E_x^m \int_0^{\tilde{\tau}} \int_U e^{-\beta s} k(x(s), \alpha) m_s(d\alpha) ds, \\ E_x^{m^n} e^{-\beta \tau^n} g(x^n(\tau^n)) \rightarrow E_x^m e^{-\beta \tilde{\tau}} g(x(\tilde{\tau})). \end{aligned} \quad (2.3)$$

If $\tilde{\tau} = \tau$, the time of first exit of the limit $x(\cdot)$ from G^0 , then we would have

$$V(x) = W(x, m), \quad (2.4)$$

and the existence of an optimal control for the cost function (2.1) would be proved.

Continuity of the First Exit Time. Unfortunately, it is not always the case that the limit $\tilde{\tau}$ is the first exit time of $x(\cdot)$. In order to better understand the problem, refer to the deterministic case illustrated in Figure 10.1.

In the figure, the sequence of functions $\phi_n(\cdot)$ converges to the limit function $\phi_0(\cdot)$, but the sequence of first contact times of $\phi_n(\cdot)$ converges to a time which is not the moment of first contact of $\phi_0(\cdot)$ with the boundary line ∂G . From the illustration, we can see that the problem in this case is that the limit function is tangent to ∂G at the time of first contact.

For our control problem, if the values $W(x, m^n)$ are to converge to the value $W(x, m)$, then we need to assure (at least with probability one) that

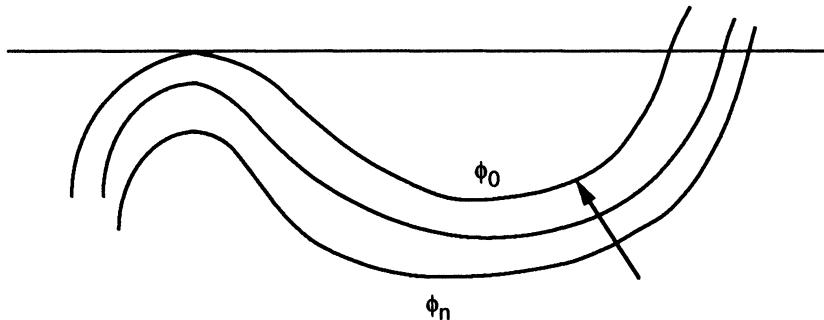


Figure 10.1. Continuity of first exit times.

the paths of the limit $x(\cdot)$ are not “tangent” to ∂G at the moment of first exit from G^0 . Let us now define our requirement more precisely. For $\phi(\cdot)$ in $D^r[0, \infty)$ (with the Skorokhod topology used), define the function $\hat{\tau}(\phi)$ with values in the compactified infinite interval $\overline{\mathbb{R}}^+ = [0, \infty]$ by: $\hat{\tau}(\phi) = \infty$, if $\phi(t) \in G^0$ for all $t < \infty$, and otherwise

$$\hat{\tau}(\phi) = \inf\{t : \phi(t) \notin G^0\}. \quad (2.5)$$

In the example of Figure 10.1, $\hat{\tau}(\cdot)$ is not continuous at the path $\phi_0(\cdot)$.

If the path $\phi_0(\cdot)$ which is drawn in the figure were actually a sample path of a Wiener process, then the probability is zero that it would be “tangent” to ∂G at the point of first contact. This is a consequence of the law of the iterated logarithm for the Wiener process or, more intuitively, because of the “local wild nature” of the Wiener process. It would cross the boundary infinitely often in any small interval about its first point of contact with a smooth boundary. The situation would be similar if the Wiener process were replaced by the solution to a stochastic differential equation with a uniformly positive definite covariance matrix $a(x)$. This was illustrated in Section 9.4, where it was shown that this “tangency” could not happen due to the law of the iterated logarithm. If such a process were the limit of the $\{x^n(\cdot)\}$ introduced above and the boundary ∂G were “smooth” (see the remarks below), then we would have $\tau^n \rightarrow \tau$, where τ is the first hitting time of the boundary. If, in addition, $\beta > 0$ or $\{\tau^n\}$ is uniformly integrable, then

$$W(x, m^n) \rightarrow W(x, m). \quad (2.6)$$

If the original sequence were minimizing, then $W(x, m) = V(x)$. The same “boundary” considerations arise when proving the convergence of the numerical approximations $V^h(x)$ to $V(x)$. We will next give conditions which will guarantee the convergence (2.6).

Convergence of the First Exit Time. The following conditions on the

set G will be used:

A2.1. *The set G is compact and is the closure of its interior G^0 .*

A2.2. *The function $\hat{\tau}(\cdot)$ is continuous (as a map from $D^r[0, \infty)$ to the compactified interval $[0, \infty]$) with probability one relative to the measure induced by any solution to (1.1) for the initial condition x of interest.*

For the purposes of the convergence theorems for the numerical approximations starting in Section 10.3, (A2.2) can be weakened as follows:

A2.2'. *For each $\epsilon_0 > 0$, and initial condition x of interest, there is an ϵ_0 -optimal process $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$ satisfying (1.1), which is unique in the weak sense, and such that $\hat{\tau}(\cdot)$ is continuous with probability $\geq 1 - \epsilon_0$ with respect to the measure of the solution $x(\cdot)$.*

By ϵ_0 -optimality of $m(\cdot)$, we mean that $W(x, m) \leq V(x) + \epsilon_0$.

Remark on (A2.2). (A2.2) and (A2.2') are stated as they are because little is usually known about the ϵ -optimal processes. Such conditions are satisfied in many applications. The main purpose is the avoidance of the “tangency” problem discussed above. The tangency problem would appear to be a difficulty with all numerical methods, since they all depend on some sort of approximation, and implicitly or explicitly one seems to need some “robustness” of the boundary conditions in order to get the desired convergence. In particular, the convergence theorems for the classical finite difference methods for elliptic and parabolic equations generally use a nondegeneracy condition on $a(x)$ in order to (implicitly) guarantee (A2.2).

The nature of the dynamical equation (1.1) often implies the continuity of $\hat{\tau}(\cdot)$ with probability one, owing essentially to the local “wildness” of the Wiener process. Let us consider a classical case. Let $a(x)$ be uniformly positive definite in G , and let ∂G satisfy the following “open cone condition”: There are $\epsilon_0 > 0, \epsilon > 0$, and open (exterior) cones $C(y)$ of radius ϵ_0 at unit distance from the origin such that for each $y \in \partial G$, we have

$$\{x : x - y \in C(y), |y - x| < \epsilon\} \cap G^0 = \emptyset.$$

Then by [40, Theorem 13.8], (A2.2) holds.

The checking of (A2.2) for the degenerate case is more complicated, and one usually needs to take the particular structure of the individual case into account. An important class of degenerate examples is illustrated in [76, pp. 64–66]. The boundary can often be divided into several pieces, where we are able to treat each piece separately. For example, there might be a segment where a “directional nondegeneracy” of the diffusion term $a(x) = \sigma(x)\sigma'(x)$ guarantees the almost sure continuity of the exit times of the

paths which exit on that segment, a segment where the direction of the drift gives a similar guarantee, a segment on which escape is not possible, and the complement of the above parts. Frequently, the last “complementary” set is a finite set of points or a curve of lower dimension than the boundary. Special considerations concerning these points can often resolve the issue there. In the two dimensional example cited above from [76], G is the symmetric square box centered about the origin and the system is ($x = (x_1, x_2)$)

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= u dt + dw, \end{aligned}$$

where the control $u(\cdot)$ is bounded. The above cited “complementary set” is just the two points which are the intersections of the horizontal axis with the boundary, and these points can be taken care of by a test such as that in Theorem 6.1 of [114].

The boundaries in control problems are often somewhat “flexible.” In many cases, they are introduced simply in order to bound the state space. The original control problem might be defined in an unbounded space, but the space truncated for numerical reasons. Even if there is a given “target,” it is often not necessary to fix the target set too precisely. These considerations give us the freedom to vary the boundary slightly. This freedom suggests the “randomized stopping” alternative discussed in the paragraph after the next.

The Girsanov Measure Transformation Technique and Continuity of the Stopping Times. Define the set $A = \{\phi(\cdot) : \hat{\tau}(\cdot) \text{ is continuous at } \phi(\cdot)\}$. Then A is an open (hence measurable) set in $D^r[0, \infty)$ in the Skorokhod topology. Suppose that for some jump diffusion process $x(\cdot)$ satisfying (1.1), the conditions required in Section 1.3 for the Girsanov measure transformation to be usable for modifying the drift are satisfied. Let P denote the measure which induces $x(\cdot)$ on $D^r[0, \infty)$ and suppose that $P\{A\} = 1$. Now let us modify the control via the measure transformation and obtain a new process satisfying (1.1) for that control, with associated measure \tilde{P} satisfying $\tilde{P} \ll P$. Then $\tilde{P}\{A\} = 1$. Thus, if (A2.2) holds for one control it holds for all controls.

An Alternative to (A2.2). If the set G can be altered even very slightly, then there is a satisfactory alternative stopping rule which accomplishes the same purpose as (A2.2). This rule is called *randomized stopping*. Under randomized stopping, the probability of stopping at some time t (if the process has not yet stopped) goes to unity as the state value $x(t)$ at that time approaches ∂G . This will now be formalized.

Randomized Stopping. For some small $\epsilon > 0$, let $\bar{\lambda}(\cdot) > 0$ be a continuous function on the set $N_\epsilon(\partial G) \cap G^0$, where $N_\epsilon(\partial G)$ is the ϵ -neighborhood

of the boundary. Let $\bar{\lambda}(x) \rightarrow \infty$ as x converges to ∂G . Then we will stop the controlled process $x(\cdot)$ at time t with stopping rate $\bar{\lambda}(x(t))$ and stopping cost $g(x(t))$.

As far as the costs are concerned, randomized stopping is equivalent to adding an additional discount factor. The cost is replaced by

$$\begin{aligned} W(x, m) &= E_x^m \int_0^\tau \int_u \exp \left[-\beta s - \int_0^s \bar{\lambda}(x(u)) du \right] k(x(s), \alpha) m_s(d\alpha) ds \\ &\quad + E_x^m \int_0^\tau \exp \left[-\beta s - \int_0^s \bar{\lambda}(x(u)) du \right] \bar{\lambda}(x(s)) g(x(s)) ds. \end{aligned} \tag{2.7}$$

The randomized stopping rule can be applied to the approximating Markov chain $\{\xi_n^h, n < \infty\}$ in the same way: Stop the chain at step n with probability $1 - \exp(-[\bar{\lambda}(\xi_n^h) \Delta t_n^h])$, for $\xi_n^h \in G^0$, and with probability one if $\xi_n^h \notin G^0$. The stopping cost is $g(\xi_n^h)$. The computational problem is altered only slightly. For example, if $\Delta t^h(x, \alpha)$ does not depend on α , then the dynamic programming equation (5.8.3) becomes

$$\begin{aligned} V^h(x) &= (1 - e^{-\bar{\lambda}(x) \Delta t^h(x)}) g(x) \\ &\quad + e^{-\bar{\lambda}(x) \Delta t^h(x)} \times \text{right side of (5.8.3)}, \end{aligned} \tag{2.8}$$

for $x \in G_h$, and with the boundary condition $V^h(x) = g(x)$ for $x \notin G_h$.

The Convergence Theorem. Theorem 1.1 and the above discussion yield the following result:

Theorem 2.1. *Assume (A1.1) – (A1.3), (A2.1), and that the cost (2.1) is used. Let either $\beta > 0$, or $\{\tau^n\}$ be uniformly integrable. Let*

$$\{x^n(\cdot), m^n(\cdot), w^n(\cdot), N^n(\cdot), \tau^n\}$$

be a sequence of solutions to (1.1) and associated stopping times which converges weakly to $(x(\cdot), m(\cdot), w(\cdot), N(\cdot), \tilde{\tau})$. Let (A2.2) hold for the limit. Then $\tilde{\tau} = \tau$ with probability one and (2.6) holds. If the sequence is minimizing, then (2.4) holds. Let $k(x) \geq 0$ and $g(x) = 0$, and let the sequence be minimizing. Then

$$\liminf_n W(x, m^n) \geq W(x, m)$$

and (2.4) holds without (A2.2), where $W(x, m)$ is the cost for the limit process whether or not the solution to (1.1) is unique.

Assume the randomized stopping in lieu of (A2.2), and replace the $\{\tau^n\}$ above with the (no larger) randomized stopping times. Then the assertions in the first paragraph remain true.

10.3 Approximating the Optimal Control

Theorem 1.2 showed that a relaxed control could be approximated by a simple ordinary control. For the proof of convergence of the numerical approximations $V^h(x)$ to the desired limit $V(x)$ in Sections 10.4 to 10.6 below, we will need an approximating control with a “finite dependence” on the “past” of the Wiener process and Poisson measure and which has certain continuity properties. These continuity properties will allow the use of weak convergence methods when applying these “comparison” controls to the Markov chains. The existence of such a control will be proved in Theorem 3.1 below.

Remark on Notation. In the theorem below, we use various standard Wiener processes and Poisson measures, which will be denoted by $w^\gamma(\cdot)$ and $N^\gamma(\cdot)$, resp., for appropriate scalar or vector values of the superscript γ . When a control $m^\gamma(\cdot)$ has the same superscript, it is assumed to be admissible with respect to the first pair. Also, $x^\gamma(\cdot)$ and τ^γ will denote the associated solution process and first exit time from G^0 , resp. The initial condition of all $x^\gamma(\cdot)$ and $x(\cdot)$ is x . For any ordinary stochastic control $u^\gamma(\cdot)$ we use $m^\gamma(\cdot)$ for its relaxed control representation.

Theorem 3.1. *Assume (A1.1) – (A1.4), (A2.1) and (A2.2') and use the cost function (2.1) with $\beta > 0$. Fix $\epsilon_0 > 0$, and let $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$ be an ϵ_0 -optimal solution whose existence is asserted in (A2.2'). Let τ denote the first escape time from G^0 . Then, for each $\epsilon > 0$, there is a $\delta > 0$ and a probability space on which are defined a pair $(w^\epsilon(\cdot), N^\epsilon(\cdot))$, a control $u^\epsilon(\cdot)$ of the type introduced in Theorem 1.2, and a solution $x^\epsilon(\cdot)$ such that*

$$|W(x, m^\epsilon) - W(x, m)| \leq \epsilon. \quad (3.1)$$

There is $\theta > 0$ and a partition $\{\Gamma_j^q, j \leq q\}$ of Γ such that the approximating $u^\epsilon(\cdot)$ can be chosen so that its probability law at any time $n\delta$, conditioned on $\{w^\epsilon(s), N^\epsilon(s), s \leq n\delta, u^\epsilon(i\delta), i < n\}$, depends only on the samples

$$\{w^\epsilon(p\theta), N^\epsilon(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta; u^\epsilon(i\delta), i < n\} \quad (3.2)$$

and is continuous in the $w^\epsilon(p\theta)$ arguments. If the set of stopping times over all controls is uniformly integrable, then the above conclusions continue to hold for $\beta = 0$.

If randomized stopping is used in lieu of the continuity of the exit times in (A2.2'), then the above assertions still hold.

Comment on the Proof. Only the first paragraph will be proved. Inequality (3.1) and the statements above it are essentially Theorem 1.2. The only modifications concern the presence of a stopping boundary. The assertions (3.1) and above yield an approximating control $u^\epsilon(\cdot)$ which takes only

finitely many values and is constant on the time intervals $[n\delta, n\delta + \delta]$. To get the form of the control whose existence is asserted below (3.1), we start with the control $u^\epsilon(\cdot)$ constructed above and modify it in several stages. The desired control will be defined via its conditional probability law, given the past values of the control and the driving Wiener process and Poisson measure. First it is shown, via use of the martingale convergence theorem, that the conditional probability law can be arbitrarily well approximated by a conditional probability law that depends on only a finite number of samples of the driving processes. In order to get the asserted continuity in the samples of the Wiener process, a mollifier is applied to the conditional probability law obtained in the step above.

Proof. Part 1. Let $\epsilon > 0$. By Theorems 1.1 and 1.2, there are $\delta > 0$, a finite set $\mathcal{U}_\epsilon \subset \mathcal{U}$ and a probability space on which are defined a solution to (1.1), namely, $(x^\epsilon(\cdot), u^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot))$, where $u^\epsilon(\cdot)$ is \mathcal{U}_ϵ -valued and constant on the intervals $[n\delta, n\delta + \delta]$. Also, $(x^\epsilon(\cdot), m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot))$ approximates $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$ in the sense that as $\epsilon \rightarrow 0$,

$$(x^\epsilon(\cdot), m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot)) \Rightarrow (x(\cdot), m(\cdot), w(\cdot), N(\cdot)). \quad (3.3)$$

Let $\tau^\epsilon = \inf\{t : x^\epsilon(t) \notin G^0\}$. There is a stopping time $\tilde{\tau}$ such that

$$(x^\epsilon(\cdot), m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot), \tau^\epsilon) \Rightarrow (x(\cdot), m(\cdot), w(\cdot), N(\cdot), \tilde{\tau}). \quad (3.4)$$

By assumption (A2.2'), there is $\epsilon_0 > 0$ which can be supposed to be arbitrary small, such that $\hat{\tau}(\cdot)$ is continuous with probability $\geq 1 - \epsilon_0$ with respect to the measure of the $x(\cdot)$ process. Also, the set

$$\{\phi(\cdot) : \hat{\tau}(\cdot) \text{ is continuous at } \phi(\cdot)\}$$

is open in the Skorokhod topology. For the various values of γ which are used below, let ν_γ denote the probability of the set of paths of the $x^\gamma(\cdot)$ process at which $\hat{\tau}(\cdot)$ is not continuous. By the previous assertions in this paragraph and Theorem 9.1.4 (ii),

$$\limsup_{\epsilon} \nu_\epsilon \leq \epsilon_0. \quad (3.5)$$

Thus, $\tilde{\tau} = \tau = \inf\{t : x(t) \notin G^0\}$ with probability $\geq 1 - \epsilon_0$. This, (3.4), and the discounting imply that

$$\limsup_{\epsilon} |W(x, m^\epsilon) - W(x, m)| \leq \delta_{\epsilon_0}, \quad (3.6)$$

where $\delta_{\epsilon_0} \rightarrow 0$, as $\epsilon_0 \rightarrow 0$.

Part 2. In order to get the representation asserted below (3.1), we will start with a $u^\epsilon(\cdot)$ of the type just described in Part 1 for small ϵ , and modify it slightly to get the desired continuity and “finite dependence” properties.

This will be the $u^{\epsilon\theta\rho}(\cdot)$ defined below. Let $0 < \theta < \delta$. Let the $\{\Gamma_j^q, j \leq q\}$ be partitions of Γ of the type used in the proof of Theorem 1.1, and let $q \rightarrow \infty$ as $\theta \rightarrow 0$. For $\alpha \in \mathcal{U}_\epsilon$, define the function $F_{n\theta}$ as the regular conditional probability

$$\begin{aligned} F_{n\theta}(\alpha; u^\epsilon(i\delta), i < n, w^\epsilon(p\theta), N^\epsilon(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta) \\ = P\{u^\epsilon(n\delta) = \alpha | u^\epsilon(i\delta), i < n, w^\epsilon(p\theta), N^\epsilon(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta\}. \end{aligned} \quad (3.7)$$

By the uniqueness assumption (A1.4), the probability law of $(x^\epsilon(\cdot), m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot))$ is determined by the probability law of $(m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot))$, because the σ -algebra determined by the set $\{u^\epsilon(i\delta), i < n, w^\epsilon(p\theta), N^\epsilon(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta\}$ increases to the σ -algebra determined by $\{u^\epsilon(i\delta), i < n, w^\epsilon(s), N^\epsilon(s), s \leq n\delta\}$ as $\theta \rightarrow 0$, the martingale convergence theorem implies that for each n, α , and δ ,

$$\begin{aligned} F_{n\theta}(\alpha; u^\epsilon(i\delta), i < n, w^\epsilon(p\theta), N^\epsilon(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta) \\ \rightarrow P\{u^\epsilon(n\delta) = \alpha | u^\epsilon(i\delta), i < n, w^\epsilon(s), N^\epsilon(s), s \leq n\delta\} \end{aligned}$$

with probability one, as $\theta \rightarrow 0$.

For given $(w^{\epsilon\theta}(\cdot), N^{\epsilon\theta}(\cdot))$, define the control $u^{\epsilon\theta}(\cdot)$ by the conditional probability

$$\begin{aligned} P\{u^{\epsilon\theta}(n\delta) = \alpha | u^{\epsilon\theta}(i\delta), i < n, w^{\epsilon\theta}(s), N^{\epsilon\theta}(s), s \leq n\delta\} \\ = F_{n\theta}(\alpha; u^{\epsilon\theta}(i\delta), i < n, w^{\epsilon\theta}(p\theta), N^{\epsilon\theta}(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta). \end{aligned} \quad (3.8)$$

By the construction of the control law, as $\theta \rightarrow 0$

$$(m^{\epsilon\theta}(\cdot), w^{\epsilon\theta}(\cdot), N^{\epsilon\theta}(\cdot)) \Rightarrow (m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot)).$$

The solution to (1.1) exists (on some probability space) and is (weak sense) unique when the control is piecewise constant and takes only a finite number of values. Using this, we get

$$(x^{\epsilon\theta}(\cdot), m^{\epsilon\theta}(\cdot), w^{\epsilon\theta}(\cdot), N^{\epsilon\theta}(\cdot)) \Rightarrow (x^\epsilon(\cdot), m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot))$$

as $\theta \rightarrow 0$. Also, by the weak convergence $x^{\epsilon\theta}(\cdot) \Rightarrow x^\epsilon(\cdot)$,

$$\limsup_\theta \nu_{\epsilon\theta} \leq \nu_\epsilon.$$

Hence, we can conclude that

$$|W(x, m^{\epsilon\theta}) - W(x, m^\epsilon)| \leq \delta_2(\nu_{\epsilon\theta}) + \delta_1(\nu_\epsilon) + \eta_\theta, \quad (3.9)$$

where the functions $\delta_i(\cdot)$ introduced here and below go to zero as their arguments go to zero, and η_θ can be made as small as desired by choosing θ small enough.

For $\rho > 0$, define the “mollified” functions $F_{n\theta\rho}(\cdot)$ by

$$\begin{aligned} F_{n\theta\rho}(\alpha; u(i\delta), i < n, w(p\theta), N(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta) \\ = N(\rho) \int \cdots \int F_{n\theta}(\alpha; u(i\delta), i < n, w(p\theta) + z_p, N(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta) \\ \times \prod_p e^{-|z_p|^2/2\rho} dz_p, \end{aligned}$$

where $N(\rho)$ is a normalizing constant such that the integral of the mollifier is unity. The $F_{n\theta\rho}$ are non-negative, their values sum (over $\alpha \in \mathcal{U}_\epsilon$) to unity, and they are continuous in the w -variables. Also, they converge to the unmollified function with probability one as $\rho \rightarrow 0$. The last assertion and the continuity are consequences of the fact that the probability distribution of a normally distributed random variable is absolutely continuous with respect to Lebesgue measure.

Let $u^{\epsilon\theta\rho}(\cdot)$ be the piecewise constant admissible control which is determined by the conditional probability distribution $F_{n\theta\rho}(\cdot)$: In particular, there is a probability space on which we can define $(w^{\epsilon\theta\rho}(\cdot), N^{\epsilon\theta\rho}(\cdot))$ and the control law $u^{\epsilon\theta\rho}(\cdot)$ by the conditional probability

$$\begin{aligned} P\{u^{\epsilon\theta\rho}(n\delta) = \alpha | u^{\epsilon\theta\rho}(i\delta), i < n, w^{\epsilon\theta\rho}(s), N^{\epsilon\theta\rho}(s), s \leq n\delta\} \\ = F_{n\theta\rho}(\alpha; u^{\epsilon\theta\rho}(i\delta), i < n, w^{\epsilon\theta\rho}(p\theta), N^{\epsilon\theta\rho}(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta). \end{aligned}$$

Then, by the construction of the probability law of the controls,

$$(x^{\epsilon\theta\rho}(\cdot), m^{\epsilon\theta\rho}(\cdot), w^{\epsilon\theta\rho}(\cdot), N^{\epsilon\theta\rho}(\cdot)) \Rightarrow (x^{\epsilon\theta}(\cdot), m^{\epsilon\theta}(\cdot), w^{\epsilon\theta}(\cdot), N^{\epsilon\theta}(\cdot))$$

as $\rho \rightarrow 0$. We therefore have

$$\limsup_\rho \nu_{\epsilon\theta\rho} \leq \nu_{\epsilon\theta}$$

and

$$|W(x, m^{\epsilon\theta\rho}) - W(x, m^{\epsilon\theta})| \leq \delta_3(\nu_{\epsilon\theta\rho}) + \delta_2(\nu_{\epsilon\theta}) + \eta_\rho, \quad (3.10)$$

where η_ρ can be made as small as desired by choosing ρ appropriately.

Putting the above arguments together, and noting that ϵ_0 can be chosen arbitrarily small, yields that for each $\epsilon > 0$ there are $\delta > 0, \theta > 0, q, w^\epsilon(\cdot), N^\epsilon(\cdot)$, and an admissible control law which is piecewise constant (on the intervals $[n\delta, n\delta + \delta]$) with values in a finite set $\mathcal{U}_\epsilon \subset \mathcal{U}$, and determined by the conditional probability law

$$\begin{aligned} P\{u^\epsilon(n\delta) = \alpha | u^\epsilon(i\delta), i < n, w^\epsilon(s), N^\epsilon(s), s \leq n\delta\} \\ \equiv F_n(\alpha; u^\epsilon(i\delta), i < n, w^\epsilon(p\theta), N^\epsilon(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta), \end{aligned} \quad (3.11)$$

where the $F_n(\cdot)$ are continuous with probability one in the w -variables, for each value of the other variables, and for which (3.1) holds.

Under the uniform integrability condition, we can restrict our attention to a finite time interval and the same proof works. ■

10.4 The Approximating Markov Chain: Weak Convergence

Let us recall the basic problem to be dealt with. Given an approximation parameter h , the basic computational model is a controlled discrete parameter Markov chain $\{\xi_n^h, n < \infty\}$ which is “locally consistent” with the controlled process (1.1) in the sense used in (4.1.3) (for the process without jumps) and in Subsection 5.6.2 for the general jump diffusion process. Here we are concerned with the cost function with an absorbing boundary or a target set. The problem with reflecting boundaries will be dealt with in Chapter 11. A cost function (2.1) for (1.1) is given, with minimum value $V(x)$. Some appropriate cost function for the chain such as (4.3.6) is given, and its optimal value $V^h(x)$ computed. We wish to prove that $V^h(x) \rightarrow V(x)$ as $h \rightarrow 0$.

The optimal value function for the discrete parameter chain is also an optimal value function for the controlled continuous parameter Markov chain model $\psi^h(\cdot)$ whose properties are defined in Chapters 4 and 5. (If $\beta > 0$, then there might be a small difference in the value functions for the two models, depending on how the discounting is approximated, but the difference goes to zero as $h \rightarrow 0$, as seen in Chapter 4.) The optimal control for the chain $\{\xi_n^h, n < \infty\}$ is a feedback control, which we denote by $\bar{u}^h(x)$. In the proofs to follow, the controls which will be used on these approximating processes will always be ordinary stochastic controls, whether they are of the pure Markov or feedback form or not. The relaxed control terminology is not needed if we are only interested in the behavior of a particular controlled chain. However, as $h \rightarrow 0$ the sequence of ordinary controls might not converge in any traditional sense, and the use of the relaxed control concept enables us to obtain and appropriately characterize limits. We have seen an example of this in Sections 4.5 and 4.6.

The sole reason for using the the approximating chain is to get an approximation to the optimal value function $V(x)$. Under appropriate regularity conditions and also nondegeneracy of $a(x)$, the convergence theorems of numerical analysis can often be used to get the convergence $V^h(x) \rightarrow V(x)$, when the approximations are based on (for example) finite difference schemes. But, owing to the common presence of degeneracy and to other “nonregularities,” this cannot be done as often as one would like. The methods of proof to be employed are purely probabilistic. They rely on the fact that the sequence of “approximating” processes which are

defined by the optimal solutions to the Markov chain control problem can be shown to converge to an optimal process of the original form (1.1). The most useful methods of proof are those of the theory of weak convergence of probability measures, which provide the tools needed to characterize the limits of the sequence of approximating processes. An example was given in Chapter 9. In this section, we will prove the convergence of the continuous parameter interpolations $\psi^h(\cdot)$ to controlled processes of the form of (1.1).

10.4.1 APPROXIMATIONS AND REPRESENTATIONS FOR $\psi^h(\cdot)$

Let us recall the definitions of Section 5.6. The process $\{\xi_n^h, n < \infty\}$ is an approximating controlled Markov chain. The jump times of the interpolation $\psi^h(\cdot)$ are denoted by $\{\tau_n^h, n < \infty\}$. If $\{u_n^h, n < \infty\}$ denotes an admissible control sequence, then the interpolation $u^h(\cdot)$ is defined by $u^h(t) = u_n^h$ on $[\tau_n^h, \tau_{n+1}^h)$. Recall the representation (5.6.11)

$$\psi^h(t) = x + \int_0^t b(\psi^h(s), u^h(s))ds + M^h(t) + J^h(t) + \delta_1^h(t), \quad (4.1)$$

where $E|\delta_1^h(t)| \rightarrow 0$, uniformly on any bounded time interval as $h \rightarrow 0$ and the processes $M^h(\cdot)$ and $J^h(\cdot)$ are defined below (5.6.11).

Define the relaxed control representation $m^h(\cdot)$ of $u^h(\cdot)$ by its derivative

$$m_t^h(A) = I_{\{u^h(t) \in A\}} \quad (4.2)$$

or, equivalently, $m_t^h(\{\alpha\}) = 1$ if $u^h(t) = \alpha$. Then we can write (4.1) in relaxed control notation as

$$\psi^h(t) = x + \int_0^t \int_U b(\psi^h(s), \alpha) m_s^h(d\alpha) ds + M^h(t) + J^h(t) + \delta_1^h(t). \quad (4.3)$$

Recall that the quadratic variation of the martingale $M^h(\cdot)$ is

$$\int_0^t a(\psi^h(s))ds + \delta_2^h(t),$$

where $E \sup_{s \leq t} |\delta_2^h(s)| \rightarrow 0$. [See below (5.6.11).]

The discounted cost function (4.3.6) is

$$\begin{aligned} W^h(x, m^h) &= E_x^{m^h} \int_0^{\tau_h} \int_U e^{-\beta s} k(\psi^h(s), \alpha) m_s^h(d\alpha) ds \\ &\quad + E_x^{m^h} e^{-\beta \tau_h} g(\psi(\tau_h)), \end{aligned} \quad (4.4)$$

and $V^h(x)$ denotes the minimum value. The dynamic programming equation is (4.3.7). Recall that we can approximate the discount factor by any

quantity $d^h(x, \alpha)$ such that $d^h(x, \alpha)/e^{-\beta\Delta t^h(x, \alpha)} \rightarrow 1$, as discussed in Chapter 4. In the next section it will be shown that $V^h(x) \rightarrow V(x)$.

Definition of an Auxiliary Process: An Approximation to the Driving Wiener Process. In the next theorem, we will show that the sequence of processes $\{\psi^h(\cdot)\}$ is tight and that any subsequence has a further subsequence which converges weakly to a controlled diffusion $x(\cdot)$ of the type (1.1), for some driving Wiener process, Poisson measure and admissible relaxed control. For the purposes of proving the desired convergence of the optimal value functions, it is useful to define a process which will converge to the actual Wiener process “driving” the limit $x(\cdot)$. The method is analogous to what is used to get the driving Wiener process for a process which solves the martingale problem. This approximation will be obtained essentially by decomposing and “inverting” $M^h(\cdot)$.

Factor

$$a(x) = \sigma(x)\sigma'(x) = P(x)D^2(x)P'(x),$$

where $P(x)$ is an orthonormal matrix, $D(x)$ is diagonal, and we can assume that each is a measurable function of x . Define $P_h(t) = P(\psi^h(t))$, $D_h(t) = D(\psi^h(t))$. We can factor $a(\psi^h(t)) = \sigma(\psi^h(t))\sigma'(\psi^h(t))$ as

$$a(\psi^h(t)) = P_h(t)D_h^2(t)P_h'(t).$$

Denote the diagonal entries of $D_h(t)$ by $\{d_{h,i}(t), i \leq r\}$. Let $\delta_0(h) \rightarrow 0$ denote the maximum step size of $\xi_{n+1}^h - \xi_n^h$ for the approximating chain for the “diffusion” steps, those excluding the jumps which are due to the approximation of the effects of the Poisson measure. Let $\delta_1(h) > 0$ be such that it goes to zero and $\delta_0(h)/\delta_1(h) \rightarrow 0$. Define the diagonal matrix $D_h^+(t)$ with entries $d_{h,i}^{-1}(t)I_{\{d_{h,i}(t) \geq \delta_1(h)\}}$, and which are defined to equal zero if $d_{h,i}(t) = 0$. Let $\tilde{w}(\cdot)$ be a Wiener process which is independent of $\{\xi_n^h, u_n^h, n < \infty\}$. Define the process $w^h(\cdot)$ by

$$w^h(t) = \int_0^t D_h^+(s)P_h'(s)dM^h(s) + \int_0^t (I - D_h(s)D_h^+(s))d\tilde{w}(s). \quad (4.5)$$

The first term is just a finite (w.p.1) sum, and the second is a stochastic integral. It can be easily verified that the defined process $w^h(\cdot)$ is a martingale. The processes defined by the two terms in (4.5) are orthogonal martingales. The quadratic variation of $w^h(\cdot)$ is just the sum of the quadratic variations of the two components and is

$$\begin{aligned} & \int_0^t D_h^+(s)P_h'(s)[P_h(s)D_h^2(s)P_h'(s)]P_h(s)D_h^+(s)ds + \epsilon^h(t) \\ & + \int_0^t (I - D_h(s)D_h^+(s))(I - D_h(s)D_h^+(s))ds = It + \epsilon^h(t), \end{aligned} \quad (4.6)$$

where I is the identity matrix and $\epsilon^h(t)$ is an error which goes to zero as $h \rightarrow 0$, and is due to the error $a_h(x) - a(x)$.

The second term in (4.5) was constructed to compensate for the degeneracies in the first term; in particular, to assure that the quadratic variation of $w^h(\cdot)$ is close to that of a Wiener process. The first term on the right side of (4.5) is linear between the jump times, and the jumps are bounded above by $\delta_0(h)/\delta_1(h)$ which goes to zero uniformly in all other variables as $h \rightarrow 0$. The truncation level $\delta_1(h)$ was chosen to assure that the jumps in $w^h(\cdot)$ would go to zero as $h \rightarrow 0$, so that any weak limit would have continuous paths with probability one and, in fact, be a standard Wiener process. The fact that any weak limit is a Wiener process is implied by the fact that a continuous local martingale whose quadratic variation function is It must be a Wiener process (Section 9.3). Using the “differential” notation, note that [ignoring the negligible error $\epsilon^h(t)$]

$$\begin{aligned}\sigma(\psi^h(t))dw^h(t) &= P_h(t)D_h(t)dw^h(t) \\ &= P_h(t)D_h(t)[D_h^+(t)P'_h(t)dM^h(t) \\ &\quad + (I - D_h(t)D_h^+(t))d\tilde{w}(t)] \\ &= dM^h(t) + [P_h(t)D_h(t)D_h^+(t)P'_h(t) - I]dM^h(t) \\ &\quad + O(\delta_1(h))d\tilde{w}(t).\end{aligned}$$

Thus, we can write

$$M^h(t) = \int_0^t \sigma(\psi^h(s))dw^h(s) + \epsilon_1^h(t),$$

where, for each t , $E \sup_{s \leq t} |\epsilon_1^h(s)| \rightarrow 0$ as $h \rightarrow 0$. We can now write (4.1) as

$$\psi^h(t) = x + \int_0^t \int_{\mathcal{U}} b(\psi^h(s), \alpha)m_s^h(d\alpha)ds + \int_0^t \sigma(\psi^h(s))dw^h(s) + J^h(t) + \epsilon_2^h(t), \quad (4.7)$$

where for each t , $\lim_{h \rightarrow 0} E \sup_{s \leq t} |\epsilon_2^h(s)| = 0$. Copying (5.6.12), write the jump term $J^h(\cdot)$ as

$$J^h(t) = \sum_{n: \nu_n^h \leq t} q_h(\psi^h(\nu_n^h^-), \rho_n), \quad (4.8)$$

where the terms in (4.8) are defined in Section 5.6. An approximation $N^h(\cdot)$ to a Poisson measure can be written in terms of $\{\nu_n^h, \rho_n\}$. For a Borel set H in Γ , define $N^h(t, H)$ by

$$N^h(t, H) = \sum_{n: \nu_n^h \leq t} I_{\{\rho_n \in H\}}. \quad (4.9)$$

Let \mathcal{F}_t^h denote the minimal σ -algebra which measures

$$\{\psi^h(s), m_s^h(\cdot), w^h(s), N^h(s), s \leq t\}.$$

We are now prepared for the convergence theorem.

10.4.2 THE CONVERGENCE THEOREM FOR THE INTERPOLATED CHAINS

Theorem 4.1. *Assume (A1.1) – (A1.2), and let the approximating chain $\{\xi_n^h, n < \infty\}$ be locally consistent with (1.1). Let $\{u_n^h, n < \infty\}$ denote the admissible sequence of controls which is used. Let $\psi^h(\cdot)$ denote the continuous parameter Markov chain interpolation and $m^h(\cdot)$ a relaxed control representation of $\{u_n^h, n < \infty\}$ for $\psi^h(\cdot)$. Let $\{\tilde{\tau}_h\}$ be a sequence of \mathcal{F}_t^h -stopping times. Then $\{\psi^h(\cdot), m^h(\cdot), w^h(\cdot), N^h(\cdot), \tilde{\tau}_h\}$ is tight. Let the limit of a weakly convergent subsequence be denoted by $(x(\cdot), m(\cdot), w(\cdot), N(\cdot), \tilde{\tau})$, and let \mathcal{F}_t denote the σ -algebra induced by $\{x(s), m(s), w(s), N(s), s \leq t, \tilde{\tau} I_{\{\tilde{\tau} \leq t\}}\}$. Then $w(\cdot)$ and $N(\cdot)$ are a standard \mathcal{F}_t -Wiener process and Poisson measure, resp., $\tilde{\tau}$ is an \mathcal{F}_t -stopping time and $m(\cdot)$ is an admissible control. Let the jump times and jump magnitudes of $N(\cdot)$ be denoted by $\{\nu_n, \rho_n\}$. We also have*

$$x(t) = x + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m_s(d\alpha) ds + \int_0^t \sigma(x(s)) dw + J(t), \quad (4.10)$$

where

$$J(t) = \sum_{n: \nu_n \leq t} q(x(\nu_n^-), \rho_n) = \int_0^t \int_{\Gamma} q(x(s^-), \rho) N(ds d\rho). \quad (4.11)$$

Proof. The direct method of Theorem 1.1 will be used. The sequences $\{m^h(\cdot), \tilde{\tau}_h\}$ are always tight since their range spaces are compact. Let $T < \infty$, and let $\tilde{\nu}_h$ be an \mathcal{F}_t^h -stopping time which is no bigger than T . Then for $\delta > 0$,

$$E_{\tilde{\nu}_h}^{m^h} |w^h(\tilde{\nu}_h + \delta) - w^h(\tilde{\nu}_h)|^2 = O(\delta) + \tilde{\varepsilon}_h,$$

where $\tilde{\varepsilon}_h \xrightarrow{h} 0$ uniformly in $\tilde{\nu}_h$. Thus, by Theorem 9.2.1, the sequence $\{w^h(\cdot)\}$ is tight. A similar argument yields the tightness of $\{M^h(\cdot)\}$. The sequence $\{N^h(\cdot)\}$ is tight (Theorem 9.2.1) because the mean number of jumps on any bounded interval $[t, t+s]$ is bounded by $\lambda s + \delta_1^h(s)$, where $\delta_1^h(s)$ goes to zero as $h \rightarrow 0$, and

$$\lim_{\delta \rightarrow 0} \inf_{h,n} P\{\nu_{n+1}^h - \nu_n^h > \delta | \text{data up to } \nu_n^h\} = 1.$$

This also implies the tightness of $\{J^h(\cdot)\}$. Finally, these results and the boundedness of $b(\cdot)$ implies the tightness of $\{\psi^h(\cdot)\}$.

For $\delta > 0$ and any process $z(\cdot)$, define the process $z_\delta(\cdot)$ by $z_\delta(t) = z(n\delta)$, $t \in [n\delta, n\delta + \delta]$. Then, by the tightness of $\{\psi^h(\cdot)\}$, (4.7) can be written as

$$\begin{aligned}\psi^h(t) &= x + \int_0^t \int_{\mathcal{U}} b(\psi^h(s), \alpha) m_s^h(d\alpha) ds + J^h(t) \\ &\quad + \int_0^t \sigma(\psi_\delta^h(s)) dw^h(s) + \varepsilon_2^{h,\delta}(t),\end{aligned}\quad (4.12)$$

where $\lim_{\delta \rightarrow 0} \limsup_{h \rightarrow 0} E|\varepsilon_2^{h,\delta}(t)| = 0$. We next characterize $w(\cdot)$ and $N(\cdot)$. A slight variation of the proof of Theorem 1.1 will be used. Using the notation of that theorem, we have

$$\begin{aligned}EH(\psi^h(t_i), w^h(t_i), (\phi_j, m^h)_{t_i}, N^h(t_i, \Gamma_j^q), j \leq q, i \leq p, \tilde{\tau}_h I_{\{\tilde{\tau}_h \leq t\}}) \\ \times [w^h(t+u) - w^h(t)] = 0.\end{aligned}\quad (4.13)$$

Abusing notation, let h index a weakly convergent sequence with limit denoted by $(x(\cdot), m(\cdot), w(\cdot), N(\cdot), \tilde{\tau})$. Then, taking weak limits in (4.13) yields

$$\begin{aligned}EH(x(t_i), w(t_i), (\phi_j, m)_{t_i}, N(t_i, \Gamma_j^q), j \leq q, i \leq p, \tilde{\tau} I_{\{\tilde{\tau} \leq t\}}) \\ \times [w(t+u) - w(t)] = 0.\end{aligned}\quad (4.14)$$

Because $w(\cdot)$ is continuous, as in Theorem 1.1, (4.14) implies that $w(\cdot)$ is a continuous \mathcal{F}_t -martingale. An analogous proof yields that

$$\begin{aligned}EH(x(t_i), w(t_i), (\phi_j, m)_{t_i}, N(t_i, \Gamma_j^q), j \leq q, i \leq p, \tilde{\tau} I_{\{\tilde{\tau} \leq t\}}) \\ \times [w(t+u)w'(t+u) - w(t)w'(t) - uI] = 0.\end{aligned}\quad (4.15)$$

Thus, the quadratic variation of the martingale $w(\cdot)$ is just tI , hence it is an \mathcal{F}_t -Wiener process. The proof that $N(\cdot)$ is a \mathcal{F}_t -Poisson measure also follows the line of argument of Theorem 1.1 and the details are omitted.

Taking limits in (4.12) as $h \rightarrow 0$ yields

$$x(t) = x + \int_0^t \int_{\mathcal{U}} b(x_\delta(s), \alpha) m_s(d\alpha) ds + \int_0^t \sigma(x_\delta(s)) dw(s) + J(t) + \varepsilon_\delta(t),$$

where

$$J(t) = \sum_{n: \nu_n \leq t} q(x(\nu_n^-), \rho_n)$$

and where $\lim_{\delta \rightarrow 0} E|\varepsilon_\delta(t)| = 0$. Finally, taking limits in this equation as $\delta \rightarrow 0$ yields (4.10). ■

10.5 Convergence of the Costs: Discounted Costs and Absorbing Boundary

We next treat the convergence of the costs $W^h(x, m^h)$ given by (4.4), where $m^h(\cdot)$ is a sequence of admissible relaxed controls for $\psi^h(\cdot)$. It will also be proved that

$$V^h(x) \rightarrow V(x). \quad (5.1)$$

By the results of Theorem 4.1, with $\tilde{\tau}_h = \tau_h$, we know that each sequence

$$\{\psi^h(\cdot), m^h(\cdot), w^h(\cdot), N^h(\cdot), \tau_h\}$$

of the type used in Theorem 4.1 has a weakly convergent subsequence whose limit processes satisfy (4.10). Abusing notation, let the given sequence converge weakly and denote the limit by $(x(\cdot), m(\cdot), w(\cdot), N(\cdot), \tilde{\tau})$. Let $\beta > 0$. Then, by the weak convergence Theorem 4.1, it is always the case that

$$\begin{aligned} E_x^{m^h} \int_0^{\tau_h} \int_{\mathcal{U}} e^{-\beta s} k(\psi^h(s), \alpha) m_s^h(d\alpha) ds \\ \rightarrow E_x^m \int_0^{\tilde{\tau}} \int_{\mathcal{U}} e^{-\beta s} k(x(s), \alpha) m_s(d\alpha) ds \end{aligned}$$

and

$$E_x^{m^h} e^{-\beta \tau_h} g(\psi^h(\tau_h)) \rightarrow E_x^m e^{-\beta \tilde{\tau}} g(x(\tilde{\tau})).$$

It is not always the case that the limit $\tilde{\tau} = \tau$, the first time of escape of $x(\cdot)$ from G^0 , analogous to the situation in Section 10.2 and Chapter 9. All the considerations discussed in these sections concerning the continuity of the exit times also hold here. Using Theorem 4.1 and following the procedure in Section 10.2 for dealing with the continuity of the first exit time, we have the following theorem which is one half of the desired result (5.1). The last assertion of the Theorem 5.1 follows from the weak convergence, Fatou's lemma and the fact that (using the Skorokhod representation of Chapter 9 for a weakly convergent subsequence so that the convergence is with probability one) $\liminf_h \tau_h \geq \tau$. A criterion for the uniform integrability is given in Theorem 5.2.

Theorem 5.1. *Assume (A1.1) – (A1.3) and (A2.1). Let*

$$\{\psi^h(\cdot), m^h(\cdot), w^h(\cdot), N^h(\cdot), \tau_h\}$$

converge weakly to $(x(\cdot), m(\cdot), w(\cdot), N(\cdot), \tilde{\tau})$. Let the limit process $x(\cdot)$ satisfy (A2.2) or else use the randomized stopping rule. Then $\tilde{\tau} = \tau$. If $\beta > 0$, or if $\{\tau_h\}$ is uniformly integrable, then

$$W^h(x, m^h) \rightarrow W(x, m) \geq V(x), \quad (5.2)$$

where the cost $W(x, m)$ is for the limit process. Also,

$$\liminf_h V^h(x) \geq V(x). \quad (5.3)$$

Let $\beta \geq 0$, $k(\cdot) \geq 0$ and $g(\cdot) \equiv 0$. Then

$$\liminf_h W^h(x, m^h) \geq W(x, m)$$

and (5.3) holds.

The Convergence Theorem. Let $\beta > 0$. In view of (5.3), in order to get the convergence (5.1), we need to prove the reverse inequality

$$\limsup_h V^h(x) \leq V(x). \quad (5.4)$$

The main idea in the proof is to use the minimality of the cost function $V^h(x)$ for the Markov chain control problem. Given an “almost optimal” control for the $x(\cdot)$ process, we adapt it for use on the chain, and then use the minimality of $V^h(x)$ and the weak convergence to get (5.4). Note that τ' in (5.6) below is the infimum of the escape times from the closed set G . It is larger than or equal to τ , the escape time from the interior G^0 . Condition (5.6) holds if there is some i such that for each $x \in G$,

$$\sum_j \sigma_{ij}^2(x) > 0. \quad (5.5)$$

In fact, the proof shows that (5.6) implies the uniform integrability of $\{\tau_h\}$.

Theorem 5.2. Assume (A1.1)–(A1.4) and (A2.1). Let $\beta > 0$, and assume (A2.2'). Then (5.4) holds. If, in addition, (A2.2) holds for all controls or if the randomized stopping rule is used, then (5.1) holds.

Let $\beta = 0$. Then, the assertion continues to hold if $\{\tau_h\}$ is uniformly integrable. Define $\tau' = \inf\{t : x(t) \notin G\}$. Assume that there is $T_1 < \infty$ and $\delta_1 > 0$ such that

$$\inf_{m, y \in G} P_y^m \{\tau' \leq T_1\} \geq \delta_1. \quad (5.6)$$

Then under the other conditions of the $\beta > 0$ case the conclusions continue to hold.

Proof. Let $\beta > 0$ and let ϵ and δ be as in Theorem 3.1. As noted above, we only need to prove (5.4). Let $(m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot))$ be an admissible triple for (1.1), where $m^\epsilon(\cdot)$ is a relaxed control representation of an ordinary control which is determined by the conditional distribution (3.11). Let $x^\epsilon(\cdot)$ denote the associated solution to (1.1). By Theorem 3.1, we can suppose that $(x^\epsilon(\cdot), m^\epsilon(\cdot), w^\epsilon(\cdot), N^\epsilon(\cdot))$ is ϵ -optimal and that

$$P_x^{m^\epsilon} \{\hat{\tau}(\cdot) \text{ not continuous at } x^\epsilon(\cdot)\} = \delta_2(\epsilon) \rightarrow 0,$$

as $\epsilon \rightarrow 0$.

Let $\{\xi_n^h, n < \infty\}$ and $\psi^h(\cdot)$ denote the controlled Markov chain and continuous parameter interpolation, resp., for the control law to be defined below. Similarly, let $w^h(\cdot)$ and $N^h(\cdot)$ be defined by (4.5) and (4.9), resp., for this chain, and let τ_h denote the first escape time from G^0 . The $(w^h(\cdot), N^h(\cdot))$ will replace the $(w(\cdot), N(\cdot))$ in the arguments of the $F_n(\cdot)$ of (3.11). Because the interpolated process $\psi^h(\cdot)$ changes values at random times which might not include the times $\{n\delta\}$ at which the control changes in the discrete time approximation which led to (3.11), we need to alter slightly the timing of the changes of the control values. Let $\{\tau_k^h, k < \infty\}$ denote the jump times of $\psi^h(\cdot)$ and define

$$\sigma_n^h = \min\{k : \tau_k^h \geq n\delta\},$$

the first jump time of $\psi^h(\cdot)$ after or at time $n\delta$. For each n , we have

$$\sigma_{n+1}^h - \sigma_n^h \rightarrow \delta \text{ in probability.} \quad (5.7)$$

We will choose $\{u_n^h, n < \infty\}$ such that $u^h(\cdot)$ will be constant on the intervals $[\sigma_n^h, \sigma_{n+1}^h]$, with values determined by the conditional probability law (3.11). In particular, for k such that $\tau_k^h \in [n\delta, n\delta + \delta)$, use the control law $u_k^h = u^h(\sigma_n^h)$ which is determined by the following conditional distribution at time σ_n^h

$$\begin{aligned} P\{u^h(\sigma_n^h) = \alpha | u^h(\sigma_i^h), i < n, \psi^h(s), w^h(s), N^h(s), s \leq n\delta\} \\ = F_n(\alpha; u^h(\sigma_i^h), i < n, w^h(p\theta), N^h(p\theta, \Gamma_j^q), j \leq q, p\theta \leq \sigma_n^h). \end{aligned} \quad (5.8)$$

Then, by Theorems 4.1 and 5.1, the assumptions concerning the ϵ -optimality, and with probability one continuity properties of $\hat{\tau}(\cdot)$ with respect to the measure of $x^\epsilon(\cdot)$ with $x^\epsilon(0) = x$, we have

$$(\psi^h(\cdot), m^h(\cdot), w^h(\cdot), N^h(\cdot), \tau_h) \Rightarrow (x^\epsilon(\cdot), m^\epsilon(\cdot), w(\cdot), N(\cdot), \tau) \quad (5.9)$$

and

$$V^h(x) \leq W^h(x, m^h) \rightarrow W(x, m^\epsilon) + \delta_3(\epsilon) \leq V(x) + \delta_3(\epsilon) + \epsilon, \quad (5.10)$$

where $\delta_3(\epsilon) \rightarrow 0$, as $\epsilon \rightarrow 0$. This yields the first assertion of the theorem.

Now let $\beta = 0$ and assume (5.6). The main difference between the two cases $\beta > 0$ and $\beta = 0$ concerns the finiteness of the costs and of the “effective stopping time”. We will prove that for each integer k

$$\limsup_h \sup_{x \in G, m} E_x^m(\tau_h)^k < \infty. \quad (5.11)$$

We show first that (5.6) implies that there are $T_2 < \infty$ and $\delta_4 > 0$ such that for small h

$$\sup_{x \in G, m} P_x^m\{\tau_h \leq T_2\} \geq \delta_4. \quad (5.12)$$

Suppose that (5.12) does not hold. Then there are sequences $y_h \in G$, $T_h \rightarrow \infty$, and admissible $m^h(\cdot)$ and associated processes $\psi^h(\cdot)$ with initial conditions $y_h \in G$ such that

$$P_{y_h}^{m^h} \{ \tau_h < T_h \} \rightarrow 0. \quad (5.13)$$

Extract a weakly convergent subsequence (indexed also by h for notational convenience) of $\{\psi^h(\cdot), m^h(\cdot), \tau_h\}$ with limit denoted by $(x(\cdot), m(\cdot), \tau)$. Then $y_h \rightarrow y_0 = x(0) \in G$. By the weak convergence and (5.6),

$$\liminf_h P_{y_h}^{m^h} \{ \tau_h \leq 2T_1 \} \geq P_{y_0}^m \{ \tau' \leq T_1 \} \geq \delta_1$$

which contradicts (5.13). Thus, (5.12) holds.

Now, let $m^h(\cdot)$ be an arbitrary admissible control for the interpolated chain $\psi^h(\cdot)$, and let τ_h denote the associated escape time from G^0 . Let \mathcal{F}_t^h denote the σ -algebra generated by the values of $\psi^h(\cdot)$ and $m^h(\cdot)$ up to time t . Now, by (5.12), for small h we can write

$$\begin{aligned} P_x^{m^h} \{ \tau_h \geq nT_2 + T_2 \} &= E_x^{m^h} I_{\{\tau_h \geq nT_2\}} I_{\{\tau_h \geq nT_2 + T_2\}} \\ &= E_x^{m^h} P_x^{m^h} \{ \tau_h \geq nT_2 + T_2 | \mathcal{F}_{nT_2}^h \} I_{\{\tau_h \geq nT_2\}} \\ &\leq E_x^{m^h} I_{\{\tau_h \geq nT_2\}} (1 - \delta_4/2) \\ &\leq (1 - \delta_4/2)^n, \end{aligned}$$

which implies that $E_x^{m^h} \tau_h \leq T_2 + T_2 / (\delta_4/2)$ and indeed that for each integer k , $E_x^m(\tau_h)^k$ is bounded uniformly in h, m and $x \in G$. ■

Convergence in the Absence of Local Consistency Everywhere. In the discussion of variable grids in Section 5.5, we encountered a problem with local consistency on the set A_0 . Owing to the way that the approximating chain was constructed on A_0 , this absence of local consistency there causes no problem with the convergence. This is a special case of the next theorem, whose proof follows essentially from those of Theorems 4.1 and 5.2.

Theorem 5.3. *Assume the conditions of Theorem 4.1 with the following exceptions. (i) There are sets $\tilde{G}^h \subset G$ and compact \tilde{G} such that $\tilde{G}^h \downarrow \tilde{G}$ and there is local consistency of the approximating chain except possibly on $\tilde{G}^h \cap G_h^0$. (ii) There are bounded $\{\tilde{b}_n^h, \tilde{a}_n^h\}$ such that for $x \in \tilde{G}^h$,*

$$E_{x,n}^{h,\alpha} \Delta \xi_n^h = \tilde{b}_n^h \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)),$$

$$\text{cov}_{x,n}^{h,\alpha} \Delta \xi_n^h = \tilde{a}_n^h \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)).$$

Let h index a weakly convergent subsequence. Then the conclusions of Theorem 4.1 hold, except that the limit take the form

$$x(t) = x + \int_0^t \tilde{b}(s)ds + \int_0^t \tilde{\sigma}(s)dw(s) + J(t), \quad (5.14)$$

where for $x(t) \notin \tilde{G}$, we have

$$\tilde{b}(t) = \int_{\mathcal{U}} b(x(t), \alpha) m_t(d\alpha), \quad \tilde{a}(t) = \tilde{\sigma}(t)\tilde{\sigma}'(t) = a(x(t)).$$

Let $\tilde{B}(y), \tilde{A}(y)$ denote the sets of possible values of $\tilde{b}(t), \tilde{a}(t)$, resp., when $x(t) = y \in \tilde{G}$. Suppose that the solution to (5.14) does not depend on the choices of the “tilde” functions in \tilde{G} within the values allowed by the sets $\tilde{B}(y), \tilde{A}(y)$. Then the limit does not depend on the choices made in the sets \tilde{G}^h .

If the conditions of Theorem 5.2 hold, but with the above exceptions, then the conclusions of that theorem continue to hold.

Remark. The theorem also elaborates the remarks made concerning discontinuous dynamics at the end of Subsection 10.1.1. The theorem can be extended to cover the cases of the next two chapters. The theorem holds for the problem of Section 5.5, where $\tilde{G} = A_0$, and the set $\tilde{A}(x)$ consists of the diagonal matrices whose second diagonal entry is unity, and the first is bounded by $1/6$.

10.6 The Optimal Stopping Problem

We now discuss the optimal stopping problem. The continuous time form was introduced in Section 3.2, and the Markov chain forms in Sections 2.2 and 5.8.

We use the uncontrolled process model

$$x(t) = x + \int_0^t b(x(s))ds + \int_0^t \sigma(x(s))dw(s) + \int_0^t \int_{\Gamma} q(x(s^-), \rho)N(ds d\rho), \quad (6.1)$$

where $w(\cdot)$ and $N(\cdot)$ are our standard Wiener process and Poisson measure, resp., with respect to some filtration \mathcal{F}_t . Let ρ be an \mathcal{F}_t -stopping time. Define the cost

$$W(x, \rho) = E_x \int_0^{\rho} k(x(s))ds + E_x g(x(\rho)). \quad (6.2)$$

A discount factor could be added with no additional difficulty. Also, a continuously acting control can be added, and then we would require a combination of the details of this section and Sections 10.4 and 10.5, but

we will still have the convergence $V^h(x) \rightarrow V(x)$, under the conditions required in Section 10.5 [and if $\beta = 0$, the positivity of the cost rate $k(\cdot)$]. We are also given a compact set G such that the process must stop by the time $\tau = \inf\{t : x(t) \notin G^0\}$ if it has not been stopped earlier. We wish to select the stopping time $\rho \leq \tau$ which minimizes the cost. Define $V(x) = \inf_{\rho \leq \tau} W(x, \rho)$.

The following assumption will be needed:

A6.1. *The solution to (6.1) is unique in the weak sense for each initial condition $x \in G^0$ in that if ρ is an \mathcal{F}_t -stopping time and $x(\cdot)$ is a solution to (6.1), then the probability law of $(w(\cdot), N(\cdot), \rho)$ determines the law of $(x(\cdot), w(\cdot), N(\cdot), \rho)$. Also, either $\hat{\tau}(\cdot)$ is continuous with probability one under the measure of $x(\cdot)$ for each initial condition x of interest, or else the randomized stopping rule is used.*

Remarks on (A6.1). The uniqueness condition holds under a uniform Lipschitz condition on $b(\cdot)$, $\sigma(\cdot)$, and $q(\cdot, \alpha)$ or if the process (6.1) is obtained via a Girsanov measure transformation from a “base” process satisfying (A6.1).

The next theorem gives a condition which guarantees that we need only consider stopping times whose moments are uniformly bounded.

Theorem 6.1. *Assume (A2.1) and (A1.1) – (A1.3) without the control, and let $\inf_{x \in G} k(x) = k_0 > 0$. Assume (A6.1). Then there exists an optimal stopping time $\bar{\rho}$ and*

$$E_x \bar{\rho} \leq 2 \max_{y \in G} g(y)/k_0. \quad (6.3)$$

Comment on the Proof. By the existence of an optimal stopping time we mean that there exists a probability space with a filtration \mathcal{F}_t , an \mathcal{F}_t -Wiener process and Poisson measure with an associated solution to (6.1), and an \mathcal{F}_t -stopping time $\bar{\rho}$, such that $W(x, \bar{\rho}) \leq W(x, \rho)$, where the cost on the right is for any other solution and stopping time on any probability space. The theorem can be proved by a compactness and weak convergence argument as in Theorems 1.1 and 2.1, and we omit the details. The bound (6.3) is a consequence of the fact that if we stop at time ρ instead of at time 0, we must have

$$g(x) \geq k_0 E_x \rho + E_x g(x(\rho)).$$

The Optimal Stopping Problem for the Approximating Markov Chain. The cost function for the approximating discrete parameter chain

and a stopping time N_h is

$$W^h(x, N_h) = E_x \sum_{n=0}^{N_h-1} k(\xi_n^h) \Delta t_n^h + E_x g(\xi_{N_h}^h).$$

For the continuous parameter chain $\psi^h(\cdot)$ and a stopping time ρ_h , the analogous cost is

$$W^h(x, \rho_h) = E_x \int_0^{\rho_h} k(\psi^h(s)) ds + E_x g(\psi^h(\rho_h)).$$

Let $V^h(x)$ denote the optimal value function. Then the dynamic programming equation for both the discrete and continuous parameter chain problem is

$$V^h(x) = \min \left[g(x), \sum_y p^h(x, y) V^h(y) + k(x) \Delta t^h(x) \right], \quad x \in G_h^0,$$

with the boundary conditions $V^h(x) = g(x)$ for $x \notin G_h^0$.

The Convergence Theorem. We have the following result:

Theorem 6.2. *Under the conditions of Theorem 6.1, or if $\sup_{x \in G} E_x \tau' < \infty$ replaces the strict positivity of $k(\cdot)$, where $\tau' = \inf\{t : x(t) \notin G\}$, we have $V^h(x) \rightarrow V(x)$.*

Proof. We work with the first set of conditions only. The proof uses an approximation procedure as in Theorems 4.1, 5.1, and 5.2. Let $(\psi^h(\cdot), \bar{\rho}_h)$ denote the continuous parameter approximating chain and its optimal stopping time, resp., and define $w^h(\cdot)$ and $N^h(\cdot)$ as in (4.5) and (4.9), resp. The sequence

$$(\psi^h(\cdot), w^h(\cdot), N^h(\cdot), \bar{\rho}_h)$$

is tight and we can assume that the $\bar{\rho}_h$ satisfy the bound in (6.3) for all h and $x \in G^0$. By use of the Markov property, as at the end of Theorem 5.2, it can be shown that this boundedness implies that $\limsup_h E_x (\bar{\rho}_h)^k < \infty$ for any positive integer k . Thus, the sequence of stopping times is uniformly integrable. Let $(x(\cdot), w(\cdot), N(\cdot), \rho)$ denote the limit of a weakly convergent subsequence. Then, analogous to the situation in Theorem 4.1, (6.1) holds for the limit processes and there is a filtration \mathcal{F}_t such that $w(\cdot)$ is an \mathcal{F}_t -Wiener process, $N(\cdot)$ is an \mathcal{F}_t -Poisson measure, ρ is an \mathcal{F}_t -stopping time and $x(\cdot)$ is adapted to \mathcal{F}_t . By the uniform integrability and the weak convergence,

$$W^h(x, \bar{\rho}_h) = V^h(x) \rightarrow W(x, \rho) \geq V(x).$$

To get the reverse inequality, we proceed as in Theorem 5.2 and use a “nice” ϵ -optimal stopping rule for (6.1) and apply it to the chain. Then

a weak convergence argument and the fact that $V^h(x)$ is optimal for the chain yields the desired reverse inequality. Let $\epsilon > 0$. First note that there are $\delta > 0$ and $T < \infty$ such that we can restrict the stopping times for (6.1) to take only the values $\{n\delta, n\delta \leq T\}$ and increase the cost (6.2) by at most ϵ . Let ρ_ϵ be an optimal stopping time for (6.1), (6.2) with this restriction. Proceeding as in Section 10.5, we can assume that this ϵ -optimal stopping time is defined by functions $F_n(\cdot)$ which are continuous in the w -variables for each value of the other variables and such that the probability law of ρ_ϵ is determined by $P\{\rho_\epsilon = 0\}$ and, for $n > 1$,

$$\begin{aligned} P\{\rho_\epsilon = n\delta | w(s), N(s), s \leq n\delta, \rho_\epsilon > n\delta - \delta\} \\ = F_n(w(p\theta), N(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta), \end{aligned}$$

where the partitions of Γ are as in Theorem 3.1.

As in Section 10.5, the comparison stopping times for the approximating chain are defined via these functions. Let ρ_h be the stopping time [for $\psi^h(\cdot)$] which is analogous to ρ_ϵ . That is, define σ_n^h as above (5.8) and let the probability law of ρ_h [which will take values $\{\sigma_n^h, n < \infty\}$] be determined by $P\{\rho_h = 0\} = P\{\rho_\epsilon = 0\}$, and, for $n > 1$,

$$\begin{aligned} P\{\rho_h = \sigma_n^h | \psi^h(s), w^h(s), N^h(s), s \leq \sigma_n^h, \rho_h > \sigma_{n-1}^h\} \\ = F_n(w^h(p\theta), N^h(p\theta, \Gamma_j^q), j \leq q, p\theta \leq \sigma_n^h). \end{aligned}$$

As in Theorem 5.2, the proof is completed by a weak convergence argument and use of the uniqueness of the solution to (6.1). ■

11

Convergence for Reflecting Boundaries, Singular Control and Ergodic Cost Problems

The development of the convergence proofs of Chapter 10 is continued, but applied to the problem classes of Chapters 7 and 8. The reflecting boundary and discounted cost problem is covered in Section 11.1. The primary mathematical difficulty with which we must contend is the proof of tightness of the “reflecting process.” The problem is avoided by use of a time rescaling method, under which all the processes are tight. After proving the weak convergence of the rescaled processes and characterizing the limits, the rescaling is inverted to obtain the desired results. This “inversion” is possible due to the conditions imposed on the allowable reflection directions. The time rescaling idea appears to be a rather powerful tool.

In Section 11.2, the rescaling method is used for the singular control problem. We treat the singular control problem with a reflecting boundary, but an absorbing boundary could be used as well. The “ergodic” or average cost per unit time problem is covered in Section 11.3. Here, we work with the “ordinary” control case, but the extension to the singular control case is straightforward. It is also noted that the weak limit of any weakly convergent sequence of stationary measures for the $\psi^h(\cdot)$ processes [recall the $\mu^h(\cdot)$ of (7.5.7)] is a stationary measure for the limit process $x(\cdot)$.

11.1 The Reflecting Boundary Problem

11.1.1 THE SYSTEM MODEL AND MARKOV CHAIN APPROXIMATION

Recall the system model with reflecting boundary from Sections 1.4 and 5.7. Rewrite (5.7.1) (with a jump term added) in relaxed control notation as follows

$$\begin{aligned} x(t) = & x + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) m_s(d\alpha) ds + \int_0^t \sigma(x(s)) dw(s) \\ & + \int_0^t \int_{\Gamma} q(x(s^-), \rho) N(ds d\rho) + z(t), \end{aligned} \tag{1.1}$$

where $m(\cdot)$ is an admissible control and the reflection term satisfies

$$\begin{aligned} |z|(t) &= \text{variation of } z(\cdot) \text{ on } [0, t] = \int_0^t I_{\partial G}(x(s)) d|z|(s), \\ z(t) &= \int_0^t \gamma(s) d|z|(s), \quad \gamma(s) \in r(x(s)), \end{aligned} \tag{1.2}$$

and $z(\cdot)$ is continuous. The function $r(\cdot)$ was defined in Sections 1.4 and 5.7.

In order to avoid trivialities in defining the “reflection” or “projection” properties if a jump of $N(\cdot)$ takes $x(\cdot)$ out of G , as in Chapter 1 we simply assume that $x + q(x, \rho) \in G$ for $x \in G$ and any $\rho \in \Gamma$.

Throughout this section, the smoothness assumptions (A10.1.1) to (A10.1.3) will be assumed to hold, where applicable, as will be the assumptions on the set G and on the reflection directions and approximating chain given in Section 5.7. The conditions on G and the reflection directions and an argument similar to that in Theorem 1.7 can be used to show that $z(\cdot)$ is continuous.

The cost function of interest is (5.8.18) which we rewrite in relaxed control notation as

$$W(x, m) = E_x^m \int_0^\infty \int_{\mathcal{U}} e^{-\beta s} [k(x(s), \alpha) m_s(d\alpha) ds + c'(x(s)) dz(s)], \tag{1.3}$$

where we suppose that $c'(x)\gamma \geq 0$ for all $\gamma \in r(x)$ and all x .

An Auxiliary Result. The following theorem on the integrability properties of the variation of the reflection process will be useful later when uniform integrability of the variation of the increments of $z(\cdot)$ and $z^h(\cdot)$ is needed.

Theorem 1.1. *Under the conditions of this section,*

$$\lim_{T \rightarrow 0} \sup_{m, x \in G} E_x^m |z|^2(T) = 0. \tag{1.4}$$

Also, for each $T < \infty$

$$\sup_{m, x \in G} E_x^m |z|^2(T) < \infty. \tag{1.4'}$$

Proof. Inequality (1.4') follows from (1.4). We start by using the upper bound on the variation given in [36, Theorem 3.4], with a different notation. For any set D and $\delta > 0$, define the neighborhood $N_\delta(D) = \{x : \inf_{y \in D} |x - y| < \delta\}$. By the proof of the cited theorem, our conditions on the boundary ∂G and reflection directions in conditions (i)-(v) of Section 5.7 imply the following: There are $\delta > 0, L < \infty$, open sets $D_i, i \leq L$, and vectors

$v_i, i \leq L$, such that $\cup_i D_i \supset G$, and if $x \in N_\delta(D_i) \cap \partial G$, then $v'_i r > \delta$ for all $r \in r(x)$.

The jump and drift terms are unimportant in the proof and we drop them henceforth. Then, write the simplified (1.1) as

$$x(t) = x + R(t) + z(t), \quad R(t) = \int_0^t \sigma(x(s)) dw(s). \quad (1.5)$$

Fix $T < \infty$. Define a sequence of stopping times β_n and indices i_n recursively as follows: Set $\beta_0 = 0$. Let i_0 be such that $R(0) = x(0) \in D_{i_0}$. Set $\beta_1 = \inf\{t : x(t) \notin N_\delta(D_{i_0})\}$. Define i_1 such that $x(\beta_1) \in D_{i_1}$. Continue in the same way to define all β_n and i_n . By the definition of the β_i ,

$$|x(\beta_i) - x(\beta_{i-1})| \geq \delta/2.$$

By the proof of [36, Theorem 3.4], the conditions in the first paragraph imply that

$$\begin{aligned} v'_{i_m-1}(x(\beta_m) - x(\beta_{m-1})) - v'_{i_m-1}(R(\beta_m) - R(\beta_{m-1})) \\ \geq \delta(|z|(\beta_m) - |z|(\beta_{m-1})). \end{aligned} \quad (1.6)$$

Define $N_T = \min\{n : \beta_n \geq T\} - 1$. Then, using the fact that the $x(t)$ are uniformly bounded, there is $\delta_1 < \infty$ such that

$$|z|(T) \leq \delta_1 N_T - \sum_{n=1}^{N_T} v'_{i_m-1} [R(\beta_m \wedge T) - R(\beta_{m-1} \wedge T)] + \delta_1 |x(T) - x|, \quad (1.7)$$

where $x = x(0)$. Then, using the boundedness of $x(t)$, there are $\delta_2 < \infty$ and a nonanticipative and bounded process $\tilde{\sigma}(\cdot)$ such that

$$E_x^m |z|^2(T) \leq \delta_2 E_x^m N_T^2 + \delta_2 E_x^m \left| \int_0^T \tilde{\sigma}(s) dw(s) \right|^2 + \delta_2 E_x^m |x(T) - x|^2. \quad (1.8)$$

We need only estimate N_T and the right hand term in (1.8). There is $\delta_3 < \infty$ such that

$$P_x^m \{ \sup_{s \leq T} |R(s)| \geq \delta/4 \} \leq 16 E_x^m |R(T)|^2 / \delta^2 \leq \delta_3 T. \quad (1.9a)$$

An argument analogous to the last part of the proof of Theorem 1.7 below implies that there is $\epsilon_1(\cdot)$ such that

$$\sup_{m,x} P_x^m \{ |z|(T) \geq \delta/4 \} \leq \epsilon_1(T), \quad (1.9b)$$

where $\epsilon_1(T)$ goes to zero as $T \rightarrow 0$. Via (1.9a,b) and the boundedness of G , we have

$$\sup_{m,x} E_x^m |x(T) - x|^2 \leq \epsilon_2(T), \quad (1.9c)$$

where $\epsilon_2(T)$ goes to zero as $T \rightarrow 0$. Now given ϵ small and positive, let T be small enough such that $\epsilon_1(T) + \delta_3 T \leq \epsilon$. Then

$$\sup_{m,x} P_x^m \left\{ \sup_{s \leq T} |x(s) - x|^2 \geq \delta/2 \right\} \leq \epsilon.$$

This implies that

$$\sup_{m,x} P_x^m \{N_T = 0\} \geq 1 - \epsilon.$$

By a recursion using the Markov property, we get

$$\sup_{m,x} P_x^m \{N_T = k\} \leq \epsilon^k, \quad k > 0,$$

which yields that $\sup_{m,x} E_x^m (N_T)^k \rightarrow 0$ as $T \rightarrow 0$ for any $k < \infty$. This last fact, together with (1.9c) and (1.8) and yields (1.4). ■

The Approximating Markov Chain. Let $\{\xi_n^h, n < \infty\}$ be a Markov chain which is locally consistent with the reflecting jump diffusion (1.1)-(1.2) in the sense of Section 5.7, and let $u^h = \{u_n^h, n < \infty\}$ be an admissible control sequence. Define the processes $w^h(\cdot)$ as in (10.4.5). Recall the representation (5.7.4) of the continuous parameter Markov process $\psi^h(\cdot)$ which is appropriate for the problem with a reflecting boundary, and which we now rewrite in relaxed control notation

$$\begin{aligned} \psi^h(t) = & x + \int_0^t \int_{\mathcal{U}} b(\psi^h(s), \alpha) m_s^h(d\alpha) ds + M^h(t) + J^h(t) \\ & + z^h(t) + \tilde{z}^h(t) + \delta_1^h(t), \end{aligned} \quad (1.10)$$

where $\tilde{z}^h(\cdot)$ is defined in Section 5.7, and

$$M^h(t) = \int_0^t \sigma(\psi^h(s)) dw^h(s)$$

and $E|\delta_1^h(t)| \rightarrow 0$, uniformly in t in any bounded set. The cost function which we use for the Markov chain interpolation is (5.8.19), and we rewrite it here

$$W^h(x, u^h) = E_x^u \sum_{n=0}^{\infty} e^{-\beta t_n^h} [k(\xi_n^h, u_n^h) \Delta t_n^h + c'(\xi_n^h) \Delta z_n^h]. \quad (1.11)$$

Because the instantaneous reflection states $x \in \partial G_h^+$ do not appear in $\psi^h(\cdot)$, the cost function (1.11) can only be approximated by a cost function for $\psi^h(\cdot)$, if there is a boundary cost. Let us rewrite (1.11) in the following relaxed control form, where $m^h(\cdot)$ is the relaxed control representation of

$u^h(\cdot)$

$$\begin{aligned} W^h(x, m^h) &= E_x^{u^h} \int_0^\infty e^{-\beta s} [k(\psi^h(s), u^h(s))ds + c'(\psi^h(s))dz^h(s) + \epsilon^h] \\ &= E_x^{m^h} \int_0^\infty \int_U e^{-\beta s} [k(\psi^h(s), \alpha)m^h(d\alpha ds) \\ &\quad + c'(\psi^h(s))dz^h(s) + \epsilon^h]. \end{aligned} \tag{1.12}$$

The error term ϵ^h is due to the approximation of the states on ∂G_h^+ by the next state in G_h . By (5.7.3), it satisfies

$$\epsilon^h \leq \int_0^\infty e^{-\beta s} \epsilon_0^h(s) dE|z^h|(s), \tag{1.12'}$$

where $\epsilon_0^h(s) \rightarrow 0$, uniformly in s as $h \rightarrow 0$. As discussed in Chapter 4, (1.11) equals (modulo ϵ^h) (1.12) if the discount factor $e^{-\beta \Delta t^h(x, \alpha)}$ in (1.11) is approximated by $1/(1 + \beta \Delta t^h(x, \alpha))$. Otherwise the two are asymptotically equivalent.

11.1.2 WEAK CONVERGENCE OF THE APPROXIMATING PROCESSES

A Problem with Weak Convergence. The main problem with carrying over the type of weak convergence arguments used in Sections 10.4 to 10.6 is due to the fact that without additional information, we cannot show that the reflection terms $\{z^h(\cdot)\}$ are tight in the Skorokhod topology. In special cases, such as in the analysis of the sort of reflected processes which arise in “heavy traffic” analysis, there are theorems which allow us to represent these reflection terms as continuous functions of the other processes on the right side of (1.10), all of which can be shown to be tight. Perhaps the basic such “continuity theorem” is the so-called reflection mapping theorem of [57, 102] and its extension to sets G which are “boxes” in [95]. Such methods were used in [80, 86, 95] for applications to control problems and have been in common use in the study of the so-called heavy traffic limits [102]. These continuity maps are very convenient where they can be applied; e.g., to the heavy traffic problems. For more general reflecting boundaries, the main results are those of [35, 92]. The continuity maps in [92] were used in the numerical approximations in [80], but do not allow us as much freedom as we would like for the approximation of both the processes on the boundary and the boundary itself.

Here we will take a somewhat more general approach, which will also set the stage for the treatment of the singular control problem. At first, we will not quite get weak convergence of the approximating processes $\psi^h(\cdot)$, but we will be able to get something very close to it, and, more importantly,

we will still be able to prove that $V^h(x) \rightarrow V(x)$. The method to be used involves a (random) rescaling of time such that the processes are tight. We then take limits of the rescaled processes and invert the rescaling to get the desired results. The method seems to have been first used in [95] for a treatment of the routing control problem in heavy traffic and for the numerical problem for singular stochastic control in [85]. This latter reference used the continuous parameter interpolation $\xi^h(\cdot)$ of Section 4.2 and not the Markov chain interpolation $\psi^h(\cdot)$ as used here, but that does not change the results. A slight extension of the proofs for the time rescaling method will be used in Theorem 1.7 to get actual weak convergence.

A Rescaling of Time: The “Stretched Out” Processes. Recall the definition (Section 5.7) of $\Delta\tau_n^h$ and τ_n^h , the interjump and jump times for $\psi^h(\cdot)$; also, $\tau_n^h = \tau_{n+1}^h$ if $\xi_n^h \in \partial G_n^h$ (the reflection states are instantaneous in the interpolation). Recall that a step n of the chain $\{\xi_n^h, n < \infty\}$ is called a reflection step if $\xi_n^h \in \partial G_h^+$, the numerical approximation of the reflecting boundary. Define $\Delta\hat{\tau}_n^h$ by

$$\Delta\hat{\tau}_n^h = \begin{cases} \Delta\tau_n^h & \text{for a nonreflection step,} \\ |\Delta z_n^h| & \text{for a reflection step.} \end{cases} \quad (1.13)$$

Define

$$\hat{\tau}_n^h = \sum_{i=0}^{n-1} \Delta\hat{\tau}_i^h.$$

Define the stretched out time scale $\hat{T}^h(\cdot)$ as follows. $\hat{T}^h(0) = 0$, the derivative of $\hat{T}^h(t)$ is 1 on $[\hat{\tau}_n^h, \hat{\tau}_{n+1}^h]$ if step n is not a reflection step (i.e., $\xi_n^h \in G_h$), and is zero otherwise. Thus, in particular,

$$\hat{T}^h(\hat{\tau}_n^h) = \tau_n^h. \quad (1.14)$$

The new time scale is illustrated in Figure 11.1, where the ξ_3^h and ξ_4^h are in ∂G_h^+ . Now define the rescaled or “stretched out” processes (denoted as the “hat” processes) $\hat{\psi}^h(\cdot)$, $\hat{u}^h(\cdot)$, etc., by

$$\hat{\psi}^h(t) = \psi^h(\hat{T}^h(t)), \quad \hat{u}^h(t) = u^h(\hat{T}^h(t)), \quad \hat{m}(d\alpha, t) = m(d\alpha, \hat{T}^h(t)),$$

etc. The time scale is stretched out at the reflection steps by an amount equal to the absolute value of the conditional mean value of the increment $\Delta\xi_n^h$, namely by $|\Delta z_n^h|$. For the case in the figure, $|\hat{z}^h|(t) = 0$ until τ_3^h , then it jumps by $|\Delta z_3^h| + |\Delta z_4^h|$, then it is constant until the next reflection step.

With the above definitions and (1.10), we have

$$\begin{aligned} \hat{\psi}^h(t) &= x + \int_0^t \int_U b(\hat{\psi}^h(s), \alpha) \hat{m}^h(d\alpha ds) + \hat{M}^h(t) \\ &\quad + \hat{J}^h(t) + \hat{z}^h(t) + \hat{\tilde{z}}^h(t) + \hat{\delta}_1^h(t). \end{aligned} \quad (1.15)$$

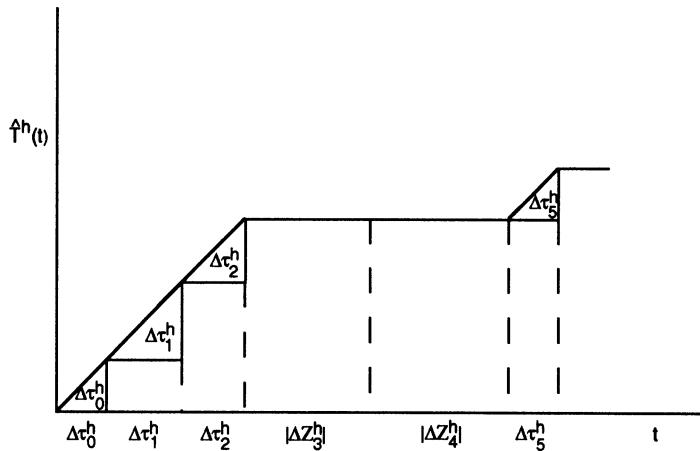


Figure 11.1. The stretched out time scale.

Note that

$$\hat{m}^h(d\alpha ds) = m_{\hat{T}^h(s)}(d\alpha)d\hat{T}^h(s).$$

Thus, it equals zero on the intervals corresponding to reflection steps (i.e., where $\hat{T}^h(\cdot)$ is flat), and takes its usual values otherwise. The process $\hat{M}^h(\cdot)$ is a martingale with quadratic variation

$$\int_0^t a(\hat{\psi}^h(s))d\hat{T}^h(s),$$

and (5.7.5) becomes

$$E \sup_{s \leq t} |\hat{z}^h(s)|^2 = O(h)E|\hat{z}^h|(t). \quad (1.16)$$

Comment. The rescaling stretches the processes in (1.10) so that they are smoother. In fact, the piecewise linear path which connects the points $|\hat{z}^h|(\hat{\tau}_n^h)$ is Lipschitz continuous with coefficient unity. Also, $\hat{T}^h(\cdot)$ is Lipschitz continuous with coefficient unity. Thus, any weak limit of these processes will have Lipschitz continuous paths with Lipschitz constant unity. The processes $z^h(\cdot)$ for which we might have had considerable difficulty in proving tightness are now stretched out enough so that they are quite “tame.” We will do the weak convergence analysis for the rescaled processes. Then, via an inverse time transformation of the limit processes, we obtain the desired results of the type $W^h(x, m^h) \rightarrow W(x, m)$ and $V^h(x) \rightarrow V(x)$. The weak convergence of $z^h(\cdot)$ itself is not required to get convergence of the cost functions of interest.

Weak Convergence of the Approximating Processes.

Theorem 1.2. Assume the conditions of Subsection 11.1.1 and let $\{u_n^h, n < \infty\}$ be an admissible control sequence for the approximating Markov chain. Then the sets of processes

$$\begin{aligned}\hat{Z}^h(\cdot) &= \{\hat{\psi}^h(\cdot), \hat{m}^h(\cdot), \hat{w}^h(\cdot), \hat{T}^h(\cdot), \hat{N}^h(\cdot), \hat{J}^h(\cdot), \hat{z}^h(\cdot)\}, \\ Q^h(\cdot) &= \{m^h(\cdot), w^h(\cdot), N^h(\cdot)\}, \{\hat{z}^h(\cdot), \hat{\delta}_1^h(\cdot)\}\end{aligned}$$

are tight. The $\hat{\delta}_1^h(\cdot)$ and $\hat{z}^h(\cdot)$ converge weakly to the zero process. Let h index a weakly convergent subsequence of $\{\hat{Z}^h(\cdot), Q^h(\cdot)\}$ with limit

$$\hat{Z}(\cdot) = \{\hat{x}(\cdot), \hat{m}(\cdot), \hat{w}(\cdot), \hat{T}(\cdot), \hat{N}(\cdot), \hat{J}(\cdot), \hat{z}(\cdot)\}, Q(\cdot) = \{m(\cdot), w(\cdot), N(\cdot)\}.$$

The pair $(w(\cdot), N(\cdot))$ are a standard Wiener process and Poisson measure, resp., with respect to the natural filtration, and $m(\cdot)$ is admissible. Also, $\hat{x}(t) \in G$. Let $\hat{\mathcal{F}}_t$ denote the σ -algebra which is generated by $\{\hat{Z}(s), s \leq t\}$. Then $w(\hat{T}(t)) = \hat{w}(t)$ and is an $\hat{\mathcal{F}}_t$ -martingale with quadratic variation $\hat{T}(t)I$. Also, $\hat{N}(t) = N(\hat{T}(t))$ and

$$\hat{J}(t) = \int_0^t q(\hat{x}(s^-), \rho) \hat{N}(ds d\rho).$$

The limit processes satisfy

$$\hat{x}(t) = x + \int_0^t \int_{\mathcal{U}} b(\hat{x}(s), \alpha) \hat{m}(d\alpha ds) + \int_0^t \sigma(\hat{x}(s)) d\hat{w}(s) + \hat{J}(t) + \hat{z}(t). \quad (1.17)$$

The process $\hat{z}(\cdot)$ can change only at those t for which $\hat{x}(t) \in \partial G$. It is differentiable and the derivative satisfies

$$\frac{d}{dt} \hat{z}(t) \in r(\hat{x}(t)).$$

Proof. Tightness of $\{\hat{Z}^h(\cdot)\}$ is proved as tightness was proved in Theorem 10.4.1, and similarly for the limiting properties of $\hat{\delta}_1^h(\cdot)$. Let h index a weakly convergent subsequence. Equation (1.16) and the boundedness of $\{E|\hat{z}^h|(t)\}$ imply the zero limit of $\hat{z}^h(\cdot)$. Theorem 10.4.1 implies that $w(\cdot)$ and $N(\cdot)$ are as asserted, as well as the admissibility of $m(\cdot)$. Because $\hat{T}(\cdot)$ and $w(\cdot)$ are continuous, we have $\hat{w}(t) = w(\hat{T}(t))$. The proofs of the assertions concerning $\hat{N}(t) = N(\hat{T}(t))$ and the representation of $\hat{J}(\cdot)$ are straightforward and are omitted. Because $\xi_n^h \in G \cup \partial G_h^+$, it follows that $\hat{x}(t) \in G$.

For any function $\phi(\cdot)$, define the piecewise constant approximating function $\phi_\delta(\cdot)$ as in Theorem 10.4.1. Then [analogous to (10.4.12)], we can write

$$\begin{aligned}\hat{\psi}^h(t) &= x + \int_0^t \int_{\mathcal{U}} b(\hat{\psi}^h(s), \alpha) \hat{m}^h(d\alpha ds) + \int_0^t \sigma(\hat{\psi}_\delta^h(s)) d\hat{w}^h(s) \\ &\quad + \hat{J}^h(t) + \hat{z}^h(t) + \tilde{\hat{z}}^h(t) + \hat{\delta}_1^h(t) + \hat{\epsilon}^{h,\delta}(t),\end{aligned} \quad (1.18)$$

where $\lim_{\delta \rightarrow 0} \limsup_{h \rightarrow 0} E|\hat{\epsilon}^{h,\delta}(t)| = 0$, and the convergence is uniform in t on any bounded interval. Taking limits in (1.18) yields

$$\begin{aligned}\hat{x}(t) &= x + \int_0^t \int_U b(\hat{x}(s), \alpha) \hat{m}(d\alpha ds) + \int_0^t \sigma(\hat{x}_\delta(s)) d\hat{w}(s) \\ &\quad + \hat{J}(t) + \hat{z}(t) + \hat{\epsilon}^\delta(t),\end{aligned}\tag{1.19}$$

where $\lim_{\delta \rightarrow 0} E|\hat{\epsilon}^\delta(t)| = 0$. If $\hat{w}(\cdot)$ can be shown to be an $\hat{\mathcal{F}}_t$ -martingale, then we could take limits in (1.19), as was done in Theorem 10.4.1, and get the representation (1.17).

For the “hatted” processes, redefine $(\phi, \hat{m})_t$ to be

$$(\phi, \hat{m})_t = \int_0^t \phi(s) \hat{m}(d\alpha ds).$$

The desired martingale property can be obtained by proceeding as in Theorem 10.4.1 and showing that

$$\begin{aligned}EH(\hat{x}(t_i), \hat{w}(t_i), (\phi_j, \hat{m})_{t_i}, \hat{N}(t_i, \Gamma_j^q), \hat{z}(t_i), j \leq q, i \leq p) \\ \times [\hat{w}(t+u) - \hat{w}(t)] = 0,\end{aligned}\tag{1.20}$$

where the Γ_j^q are as used in Theorem 10.4.1.

By the upper semicontinuity property of $r(\cdot)$ (the condition (v) in Section 5.7.3) and the fact that Δz_n^h is in a direction in $r(x)$ if $\xi_n^h = x \in \partial G_h^+$, we have that $(d/dt)\hat{z}(t) \in r(\hat{x}(t))$. ■

Remark. Note for future reference that

$$\begin{aligned}EH(\hat{x}(t_i), \hat{w}(t_i), (\phi_j, \hat{m})_{t_i}, \hat{N}(t_i, \Gamma_j^q), \hat{z}(t_i), j \leq q, i \leq p) \\ \times [\hat{w}(t+u)\hat{w}'(t+u) - \hat{w}(t)\hat{w}'(t) - (\hat{T}(t+u) - \hat{T}(t))I] = 0.\end{aligned}\tag{1.21}$$

Theorem 1.3. *Under the conditions of Subsection 11.1.1*

$$\lim_{T \rightarrow 0} \limsup_{h \rightarrow 0} \sup_{m^h, x} E_x^{m^h} |z^h|^2(T) = 0.\tag{1.22}$$

Also, for any $T < \infty$

$$\limsup_{h \rightarrow 0} \sup_{m^h, x} E_x^{m^h} |z^h|^2(T) < \infty.\tag{1.23}$$

Remark on the Proof. The proof parallels that of Theorem 1.1. It uses the validity of (1.6) for $\psi^h(\cdot)$ and $z^h(\cdot)$ (for small h) and the fact that for any $\epsilon > 0$ and sequence $\{m^h(\cdot)\}$ of admissible controls,

$$\lim_{T \rightarrow 0} \limsup_{h} \sup_{m^h, x} P_x^{m^h} \{|z^h|(T) \geq \epsilon\} = 0.$$

The latter fact can be proved by an argument for the “stretched out” or “hatted” processes of the type used in the last part of Theorem 1.7 below.

Theorem 1.4. *Assume the conditions of Subsection 11.1.1, and let $(\hat{Z}(\cdot), Q(\cdot))$ denote the limits of a weakly convergent subsequence as in Theorem 1.2. Define the inverse*

$$T(t) = \inf\{s : \hat{T}(s) > t\}.$$

Then, $T(\cdot)$ is right continuous and $T(t) \rightarrow \infty$ as $t \rightarrow \infty$, with probability one. For any process $\hat{\phi}(\cdot)$, define the “inverse” $\phi(t) = \hat{\phi}(T(t))$, and let \mathcal{F}_t denote the minimum σ -algebra which measures $\{Z(s), s \leq t\}$. Then $w(\cdot)$ and $N(\cdot)$ are a standard \mathcal{F}_t -Wiener process and Poisson measure, resp. Also, $m(\cdot)$ is admissible with respect to $(w(\cdot), N(\cdot))$ and (1.1) and (1.2) hold.

Proof. Inequality (1.23) implies that $\hat{T}(t) \rightarrow \infty$ with probability one as $t \rightarrow \infty$. Thus, $T(t)$ exists for all t and $T(t) \rightarrow \infty$ as $t \rightarrow \infty$ with probability one. By (1.20) and (1.21) we also have

$$EH(x(t_i), w(t_i), (\phi_j, m)_{t_i}, N(t_i, \Gamma_j^q), z(t_i), j \leq q, i \leq p)$$

$$\times [w(t+u) - w(t)] = 0,$$

$$EH(x(t_i), w(t_i), (\phi_j, m)_{t_i}, N(t_i, \Gamma_j^q), z(t_i), j \leq q, i \leq p)$$

$$\times [w(t+u)w'(t+u) - w(t)w'(t) - uI] = 0.$$

Thus, $w(\cdot)$ is an \mathcal{F}_t -Wiener process. We omit the details concerning the fact that $N(\cdot)$ is an \mathcal{F}_t -Poisson measure. It follows that $m(\cdot)$ is admissible with respect to $(w(\cdot), N(\cdot))$, using the filtration \mathcal{F}_t . Finally, a rescaling in (1.17) yields that (1.1) and (1.2) hold. ■

The Limits of the Cost Functions. The next theorem shows that the costs $V^h(x)$ and $W^h(x, m^h)$ converge to the costs for the limit processes and that

$$\liminf_h V^h(x) \geq V(x). \quad (1.24)$$

Given (1.24), in order to complete the proof of convergence

$$V^h(x) \rightarrow V(x), \quad (1.25)$$

we need to use a method which is similar to that in Theorem 10.5.2 and first get a “nice” comparison control. This will be done in Theorem 1.6.

Theorem 1.5. *Assume the conditions of Subsection 11.1.1, and let h index a weakly convergent subsequence of $\{\hat{Z}^h(\cdot), Q^h(\cdot)\}$ with limits as denoted in*

Theorem 1.2. Then

$$\begin{aligned} W(x, m^h) &\rightarrow E_x^m \int_0^\infty \int_{\mathcal{U}} e^{-\beta \hat{T}(t)} [k(\hat{x}(t), \alpha) \hat{m}(d\alpha dt) + c'(\hat{x}(t)) d\hat{z}(t)] \\ &= E_x^m \int_0^\infty \int_{\mathcal{U}} e^{-\beta t} [k(x(t), \alpha) m(d\alpha dt) + c'(x(t)) dz(t)] = W(x, m). \end{aligned} \quad (1.26)$$

Furthermore, (1.24) holds.

Proof. By Theorem 1.3, we have uniform integrability of

$$\{|z^h|(n+1) - |z^h|(n); n, h, m^h, \xi_0^h\}. \quad (1.27)$$

The integrability properties of (1.27) imply that $\hat{T}^h(s) \rightarrow \infty$ as $s \rightarrow \infty$, with probability one. Thus, the cost (1.12) can be written as

$$\begin{aligned} W^h(x, m^h) &= E_x^{m^h} \int_0^\infty \int_{\mathcal{U}} e^{-\beta \hat{T}^h(s)} [k(\hat{\psi}^h(s), \alpha) \hat{m}^h(d\alpha ds) \\ &\quad + c'(\hat{\psi}^h(s)) d\hat{z}^h(s)] + \epsilon^h. \end{aligned}$$

By the uniform integrability of (1.27), $\epsilon^h \rightarrow 0$.

Note that by the definition of $\hat{T}^h(\cdot)$ and the integrability properties of (1.27)

$$E_x^{m^h} \int_T^\infty e^{-\beta \hat{T}^h(s)} d|\hat{z}^h|(s) \leq E_x^{m^h} \int_{\min\{t: \hat{T}^h(t) \geq T\}}^\infty e^{-\beta s} d|z^h|(s) \rightarrow 0,$$

uniformly in h as $T \rightarrow \infty$. Also, due to the tightness and the uniform integrability properties of (1.27), for any T

$$\int_0^T e^{-\beta \hat{T}^h(s)} c'(\hat{\psi}^h(s)) d\hat{z}^h(s)$$

can be arbitrarily well approximated (uniformly in h) by a finite Riemann sum. These facts, the uniform integrability properties of (1.27), and the weak convergence, imply that $W^h(x, m^h)$ converges to

$$\hat{W}(x, \hat{m}) = E_x^m \int_0^\infty \int_{\mathcal{U}} e^{-\beta \hat{T}(s)} [k(\hat{x}(s), \alpha) \hat{m}(d\alpha ds) + c'(\hat{x}(s)) d\hat{z}(s)].$$

By an inverse transformation, this equals

$$W(x, m) = E_x^m \int_0^\infty \int_{\mathcal{U}} e^{-\beta s} [k(x(s), \alpha) m(d\alpha ds) + c'(x(s)) dz(s)].$$

To show (1.24), let the $m^h(\cdot)$ above be a minimizing sequence. Then extract a weakly convergent subsequence, also indexed by h , and with limits denoted as above to get $W^h(x, m^h) \rightarrow W(x, m) \geq V(x)$. ■

In order to complete the development, we need the following analogues of the assumptions used in Chapter 10.

A1.1. *For each $\epsilon > 0$ and initial condition of interest, there is an ϵ -optimal solution $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$ to (1.1), (1.2) which is unique in the weak sense. That is, the distribution of $(m(\cdot), w(\cdot), N(\cdot))$ implies that of $(x(\cdot), m(\cdot), w(\cdot), N(\cdot))$.*

A1.2. *Let $u(\cdot)$ be an admissible ordinary control with respect to $(w(\cdot), N(\cdot))$, and suppose that $u(\cdot)$ is piecewise constant and takes only a finite number of values. Then, for each initial condition there exists a weak sense solution to (1.1) and (1.2), where $m(\cdot)$ is a relaxed control representation of $u(\cdot)$, and this solution is unique in the weak sense.*

Theorem 1.6. *Under the other conditions of this section and (A1.1) and (A1.2), (1.25) holds.*

Proof. The proof is similar to that of Theorem 10.5.2. Use the assumed uniqueness conditions and the uniform integrability of

$$\{|z|(n+1) - |z|(n); n, m, x(0)\} \quad (1.28)$$

(which follows from Theorem 1.1) to get a comparison control of the type used in Theorem 10.5.2. Then use the weak convergence results of Theorems 1.2, 1.4 and 1.5. ■

Tightness of $\{Z^h(\cdot)\}$ and Continuity of $T(\cdot)$. Theorem 1.2 actually implies that the original processes $\{Z^h(\cdot)\}$ is tight. This will now be proved because it will be needed in Section 11.3. Theorem 1.7 also implies the continuity of $z(\cdot)$, as well as the existence of a weak sense solution to (1.1) and (1.2).

Theorem 1.7. *Assume the conditions of Subsection 11.1.1. Let h index a weakly convergent sequence of $\{\hat{Z}^h(\cdot)\}$ and define $T(\cdot)$ as in Theorem 1.4. Then $T(\cdot)$ exists for all t . It is continuous with probability one and $Z^h(\cdot) \Rightarrow Z(\cdot)$.*

Proof. Define $T^h(t) = \inf\{s : \hat{T}^h(s) > t\}$. Suppose that $T(t)$ exists for each t and $T(\cdot)$ is continuous, with probability one. Then $\{T^h(\cdot)\}$ must be tight and the weak limit must equal $T(\cdot)$ with probability one. In any case,

$$Z^h(t) = \hat{Z}^h(T^h(t)) = Z^h(\hat{T}^h(T^h(t))).$$

The weak convergence and the continuity of $T(\cdot)$ now yields

$$Z^h(\cdot) \Rightarrow \hat{Z}(T(\cdot)) = Z(\cdot).$$

Thus, we need only prove the existence of $T(t)$ for each t and the continuity of $T(\cdot)$.

For the rest of the proof, we drop the jump term $J(\cdot)$, because we can always work “between” the jumps. Suppose that the inverse $T(t)$ does not exist for all t . Then, loosely speaking, there must be a sequence of intervals whose length does not go to zero and such that $\hat{T}^h(\cdot)$ “flattens” out on them. More particularly, there are $\rho_0 > 0, \epsilon_0 > 0, t_0 > 0$, and a sequence of stopping times $\{v_h\}$ such that for all $\epsilon > 0$

$$\begin{aligned}\liminf_h P\{|\hat{z}^h|(v_h + t_0) - |\hat{z}^h|(v_h) \geq \epsilon_0, \hat{T}^h(v_h + t_0) - \hat{T}^h(v_h) \leq \epsilon\} \\ \geq 2\rho_0.\end{aligned}$$

Extract a weakly convergent subsequence of

$$\hat{A}^h = \{\hat{\psi}^h(v_h + \cdot), \hat{m}^h(v_h + \cdot) - \hat{m}^h(v_h), \hat{w}^h(v_h + \cdot) - \hat{w}^h(v_h), \dots\}$$

with limit

$$(\hat{x}(\cdot), \hat{m}(\cdot), \hat{w}(\cdot), \dots).$$

Then on a set of probability $\geq \rho_0$, we have

$$\hat{x}(t) \in \partial G \text{ on } [0, t_0], \hat{T}(t_0) = \hat{T}(0) = 0, |\hat{z}|(t_0) > 0.$$

The fact that $|\hat{z}|(t_0) > 0$ as asserted above is implied by the weak convergence and the boundary conditions (i)-(v) of Section 5.7.3.

Thus, on this set $d\hat{x}(t) = d\hat{z}(t)$ on $[0, t_0]$ and $\hat{x}(t) \in \partial G$. This violates the conditions on the reflection directions. In particular, the conditions on the boundary and reflection directions (i)-(v) of Section 5.7.3 imply that $\hat{x}(\cdot)$ cannot remain on an edge on any time interval on which $\hat{z}(\cdot)$ is not constant. The possible values of $\hat{z}(s)$ on that interval would force $\hat{x}(\cdot)$ off the edge. Similarly $\hat{x}(\cdot)$ cannot remain on an 2-face, etc. on any time interval on which $\hat{z}(\cdot)$ is not constant. Thus, $T(\cdot)$ exists for all $t < \infty$ with probability one. The same proof can be applied to yield the continuity of $T(\cdot)$. ■

11.2 The Singular Control Problem

The system model is (8.1.22') but we add a jump term $J(\cdot)$. The associated cost is (8.1.23'), where $k_i \geq 0, q_i > 0$, and we use the smoothness assumptions (A10.1.1) to (A10.1.3), where applicable. Throughout this section, suppose that the chain $\{\xi_n^h, n < \infty\}$ is locally consistent with (8.1.22') in the sense of Section 8.3, and the set G is the “box” used in that section. The model can be generalized in many directions via use of the techniques of Chapter 10 or of the last section. For example, one can add a continuously acting control term and also treat the optimal stopping problem. The conditions on the boundary can also be weakened as in the last section,

provided that the conditions of Section 5.7 continue to hold. We prefer to deal with the simpler case because that is the easiest way to expose the basic structure of the method.

Recall that (see below (8.1.22')) $F(\cdot)$ has the form

$$F(\cdot) = \sum_i v_i F^i(\cdot),$$

where the v_i are given vectors and the $F^i(\cdot)$ are real valued processes. The interpolation $\psi^h(\cdot)$ can be represented as

$$\begin{aligned} \psi^h(t) &= x + \int_0^t b(\psi^h(s))ds + M^h(t) + J^h(t) + F^h(t) \\ &\quad + (I - P')Y^h(t) - U^h(t) + \tilde{z}^h(t) + \tilde{F}^h(t) + \delta_1^h(t), \end{aligned} \quad (2.1a)$$

where

$$F^h(t) = \sum_i v_i F^{h,i}(t),$$

the terms $M^h(t)$, $J^h(t)$, $\tilde{z}^h(t)$, $\delta_1^h(t)$ are as in Section 5.7, and $\tilde{F}^h(t)$ satisfies (8.3.4). The cost function for the approximating chain is (8.3.5) and it can be rewritten as

$$W^h(x, F^h) = E_x^{F^h} \int_0^\infty e^{-\beta t} [k(\psi^h(t))dt + \sum_i q_i dF^{h,i}(t) + \sum_i k_i dU^{h,i}(t)]. \quad (2.1b)$$

Many of the results and methods of Section 11.1 can be carried over without much change. By a slight variation of Theorem 1.1, we have the following result.

Theorem 2.1. *Assume the conditions of this section and let $E|F(T)|^2 < \infty$ for each $T < \infty$. Then*

$$\sup_{x \in G} E_x^F (|Y(T)|^2 + |U(T)|^2) < \infty$$

for each $T < \infty$.

The analogue for the approximating chain can be proved along the same lines as used for Theorem 1.1 and is as follows.

Theorem 2.2. *Under the assumptions of this section and the condition that*

$$\limsup_h \sup_n E|F^h(n+1) - F^h(n)|^2 < \infty, \quad (2.2)$$

we have

$$\limsup_h \sup_n E_x^{F^h} (|Y^h(n+1) - Y^h(n)|^2 + |U^h(n+1) - U^h(n)|^2) < \infty.$$

In Section 11.1, it was not a priori obvious that the reflection terms $\{z^h(\cdot)\}$ were tight. We dealt with that problem by use of a stretched out time scale. But in Theorem 1.7, we showed that the $\{Z^h(\cdot)\}$ actually was tight. The situation is more complicated here, because the control sequence $\{F^h(\cdot)\}$ can always be chosen such that neither it nor the associated sequence of solutions is tight. The time rescaling method of Section 11.1 still works well. In order to use it, we will have to redefine the rescaling to account for the singular control terms. Recall the trichotomy used in Chapter 8, where each step of the chain $\{\xi_n^h, n < \infty\}$ is either a control step, a reflection step, or a step where the transition function for the uncontrolled and unreflected case is used (which is called a “diffusion” step). Redefine $\Delta\hat{\tau}_n^h$ by

$$\Delta\hat{\tau}_n^h = \begin{cases} \Delta\tau_n^h & \text{for a diffusion step,} \\ |\Delta Y_n^h| + |\Delta U_n^h| & \text{for a reflection step,} \\ |\Delta F_n^h| & \text{for a control step.} \end{cases} \quad (2.3)$$

Recall the definition of $\hat{T}^h(\cdot)$ above (1.14). Redefine $\hat{T}^h(\cdot)$ such that its slope is unity on the interval $[\hat{\tau}_n^h, \hat{\tau}_{n+1}^h]$ only if n is a diffusion step [that is, if $\xi_n^h \in G_h$ and no control is exercised], and the slope is zero otherwise. Now, redefine the processes $\hat{\psi}^h(\cdot)$, etc., with this new scale, analogously to what was done in Section 11.1. In place of (1.15) we have

$$\begin{aligned} \hat{\psi}^h(t) &= x + \int_0^t b(\hat{\psi}^h(s))d\hat{T}^h(s) + \hat{M}^h(t) + \hat{J}^h(t) + \hat{F}^h(t) \\ &\quad + (I - P')\hat{Y}^h(t) - \hat{U}^h(t) + \hat{\tilde{z}}^h(t) + \hat{\tilde{F}}^h(t) + \hat{\delta}_1^h(t). \end{aligned} \quad (2.4)$$

In the present context, Theorems 1.2 and 1.4 can be rewritten as follows.

Theorem 2.3. *Assume the conditions of this section. Then the sets of processes*

$$\begin{aligned} \hat{Z}^h(\cdot) &= \{\hat{\psi}^h(\cdot), \hat{w}^h(\cdot), \hat{T}^h(\cdot), \hat{N}^h(\cdot), \hat{Y}^h(\cdot), \hat{U}^h(\cdot), \hat{F}^h(\cdot)\}, \\ Q^h(\cdot) &= \{w^h(\cdot), N^h(\cdot)\}, \{\hat{\delta}_1^h(\cdot), \hat{\tilde{Y}}^h(\cdot), \hat{\tilde{U}}^h(\cdot), \hat{\tilde{F}}^h(\cdot)\} \end{aligned}$$

are tight. The third set converges to the zero process. Let h index a weakly convergent subsequence of $\{\hat{Z}^h(\cdot), Q^h(\cdot)\}$ with limit

$$\hat{Z}(\cdot) = (\hat{x}(\cdot), \hat{w}(\cdot), \hat{T}(\cdot), \hat{N}(\cdot), \hat{Y}(\cdot), \hat{U}(\cdot), \hat{F}(\cdot)), Q(\cdot) = (w(\cdot), N(\cdot)).$$

The $w(\cdot)$ and $N(\cdot)$ are the standard Wiener process and Poisson measure, resp., with respect to the natural filtration, and $\hat{x}(t) \in G$. Let $\hat{\mathcal{F}}_t$ denote the σ -algebra which is generated by $\{\hat{Z}(s), s \leq t\}$. Then $w(\hat{T}(t)) = \hat{w}(t)$ and is an $\hat{\mathcal{F}}_t$ -martingale with quadratic variation $\int_0^t a(\hat{x}(s))d\hat{T}(s)$. Also,

$$\hat{J}(t) = \int_0^t q(\hat{x}(s^-), \rho) \hat{N}(ds d\rho).$$

and

$$\begin{aligned}\hat{x}(t) &= x + \int_0^t b(\hat{x}(s))d\hat{T}(s) + \int_0^t \sigma(\hat{x}(s))d\hat{w}(s) + \hat{F}(t) \\ &\quad + \hat{J}(t) + (I - P')\hat{Y}(t) - \hat{U}(t).\end{aligned}\tag{2.5}$$

The process $\hat{Y}^i(\cdot)$, [resp., $\hat{U}^i(\cdot)$] can change only at those t for which $\hat{x}^i(t) = 0$ [resp., $\hat{x}^i(t) = B_i$]. If (2.2) holds, then Theorem 1.4 continues to hold and the rescaled processes satisfy (8.1.22') with the jump term added.

The limit of the costs can be dealt with via the following theorem.

Theorem 2.4. Assume the conditions of this section and (2.2), and let h index a weakly convergent subsequence of $\{\hat{Z}^h(\cdot), Q^h(\cdot)\}$. Then, with the other notation of Theorem 2.3 used,

$$\begin{aligned}W^h(x, F^h) &\rightarrow E_x^F \int_0^\infty e^{-\beta \hat{T}(t)} \left[k(\hat{x}(t))d\hat{T}(t) + q'd\hat{F}(t) + k'd\hat{U}(t) \right] \\ &= E_x^F \int_0^\infty e^{-\beta t} [k(x(t))dT(t) + q'dF(t) + k'dU(t)] \\ &= W(x, F).\end{aligned}\tag{2.6a}$$

Also,

$$\liminf_h V^h(x) \geq V(x).\tag{2.6b}$$

Completion of the Proof that $V^h(x) \rightarrow V(x)$. In order to complete the proof, we need to find a comparison control as in Theorem 1.6 or 10.5.2. The following conditions will be needed.

A2.1. For each $\epsilon > 0$ and initial condition of interest, there is an ϵ -optimal solution $(x(\cdot), F(\cdot), w(\cdot), N(\cdot))$ to (8.1.22') with the jump term added, and it is unique in the weak sense.

A2.2. Let $F(\cdot)$ be an admissible control with respect to $(w(\cdot), N(\cdot))$, and suppose that $F(\cdot)$ is piecewise constant and takes only a finite number of values. Then for each initial condition there exists a weak sense solution to (8.1.22') with the jump term added, and this solution is unique in the weak sense.

To construct the comparison control, we will use the following result, which can be proved by an approximation and weak convergence argument. See [85] for the details of a closely related case. Let F^a denote the a^{th} component of the control vector F .

Theorem 2.5. Assume the conditions of this section and (A2.1), (A2.2). Let $\epsilon > 0$. There is an ϵ -optimal admissible solution $(x(\cdot), F(\cdot), w(\cdot), N(\cdot))$ to (8.1.22') (with the jump term added) with the following properties. (i) There are $T_\epsilon < \infty, \delta > 0, \theta > 0, k_m < \infty$, and $\rho > 0$, such that $F(\cdot)$ is constant on the intervals $[n\delta, n\delta + \delta]$, only one of the components can jump at a time, and the jumps take values in the discrete set $k\rho, k = 0, \dots, k_m$. Also, $F(\cdot)$ is bounded and is constant after time T_ϵ . (ii) The values are determined by the conditional probability law (the expression defines the functions $q_{nka}(\cdot)$)

$$\begin{aligned} P\{dF^a(n\delta) = k\rho, dF^b(n\delta) = 0, b \neq a | F(i\delta), i < n, w(s), N(s), s \leq n\delta\} \\ \equiv q_{nka}(k\rho; F(i\delta), i < n, w(p\theta), N(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta), \end{aligned} \quad (2.7)$$

where the $q_{nka}(\cdot)$ are continuous with probability one in the w -variables, for each value of the other variables.

The final convergence theorem can now be stated.

Theorem 2.6. Under the conditions of this section and (A2.1) and (A2.2), $V^h(x) \rightarrow V(x)$.

Proof. We need to adapt the ϵ -optimal control of Theorem 2.5 for the approximating Markov chain. In preparation for the argument, let us first note the following. Suppose that we are given a control of “impulsive magnitude” $v_i dF^i$ acting at a time t_0 . Let the other components $F^j(\cdot), j \neq i$, of the control be zero. Thus, the associated instantaneous change of state is $v_i dF^i$. We wish to adapt this control for use on the approximating chain. To do this, first define $n_h = \min\{k : \tau_k^h \geq t_0\}$. Then, starting at step n_h of the approximating chain, we approximate $v_i dF^i$ by applying a succession of admissible control steps in conditional mean direction v_i , each of the randomized type described in Section 8.3. In more detail, let $E_k^h \Delta \xi_k^h = v_i \Delta F_k^{h,i}, k \geq n_h$, denote the sequence of “conditional means,” as in Section 8.3. Continue until $\sum_k \Delta F_k^{h,i}$ sums to dF^i (possibly modulo a term which goes to zero as $h \rightarrow 0$). There might be some reflection steps intervening if G is ever exited. Let $F_0^h(\cdot)$ denote the continuous parameter interpolation of the control process just defined. Because the interpolation interval at a control or reflection step is zero, all of the “conditional mean” jumps $v_i \Delta F_k^{h,i}$ occur simultaneously in the interpolation, and the sequence $\{F_0^h(\cdot)\}$ is tight. Also, $\tilde{F}^h(\cdot)$ converges weakly to the zero process. Thus, the weak limit is just the piecewise constant control process with a single jump, which is at time t_0 and has the value $v_i dF^i$.

With the above example in mind, we are now ready to define the adapted form of the ϵ -optimal control $F(\cdot)$ given in Theorem 2.5. Let $F^h(\cdot)$ denote the continuous parameter interpolation of this adaptation. $F^h(\cdot)$ will be

defined so that it is piecewise constant and has the same number of jumps that $F(\cdot)$ has (at most T_ϵ/δ). Each of the jumps of each component of the control is to be realized for the chain in the manner described in the above paragraph. The limit of the associated sequence $(\psi^h(\cdot), F^h(\cdot), w^h(\cdot), N^h(\cdot))$ will have the distribution of the $(x(\cdot), F(\cdot), w(\cdot), N(\cdot))$ of Theorem 2.5. The non zero jumps for the approximating chain are to occur as soon after the interpolated times $n\delta$ as possible, analogous to the situation in Theorem 10.5.2.

Next, determine the jumps $dF_n^{h,i}$ of the control values for the adaptation of $F(\cdot)$ to the approximating chain by the conditional probability law:

$$\begin{aligned} P\{dF_n^{h,a} = k\rho, dF_n^{h,b} = 0, b \neq a | & F^h(\sigma_i^h), i < n, \psi^h(s), w^h(s), N^h(s), \\ & U^h(s), L^h(s), s \leq n\delta\} \\ &= q_{nka}(k\rho; F^h(\sigma_i^h), i < n, w^h(p\theta), N^h(p\theta, \Gamma_j^q), j \leq q, p\theta \leq n\delta). \end{aligned} \quad (2.8)$$

Let $v_a dF^{a,h}(n\delta)$ denote the jump magnitude chosen by the law (2.8) for the chain at time $n\delta$. Now, we must “realize” this value. Following the method of Theorem 10.5.2, define $\sigma_n^h = \min\{k : \tau_k^h > n\delta\}$. Now, starting at step σ_n^h , realize $v_a dF^{a,h}(n\delta)$ by the method of the first paragraph of the proof. Then, continuing to follow the method of Theorem 10.5.2, apply Theorems 2.3 and 2.4 to get that

$$V^h(x) \leq W^h(x, F^h) \rightarrow W(x, F) \leq V(x) + \epsilon,$$

which yields the theorem. ■

11.3 The Ergodic Cost Problem

The system model is the reflected jump diffusion (1.1) and the cost function is

$$\gamma(x, m) = \limsup_T \frac{1}{T} E_x^m \int_0^T \left[\int_{\mathcal{U}} k(x(t), \alpha) m_t(d\alpha) dt + c'(x(t)) dz(t) \right]. \quad (3.1)$$

The limit might depend on x . Let $\bar{\gamma}$ denote the minimal cost, and suppose that it does not depend on the initial condition. Throughout this section, assume the smoothness assumptions (A10.1.1) to (A10.1.3), where applicable, and the assumptions on G of Section 5.7. The chain $\{\xi_n^h, n < \infty\}$ is assumed to be locally consistent with the reflected jump diffusion in the sense of Section 5.7. Assume the conditions on the function $c(\cdot)$ below (5.8.18). We use the notation of Section 11.1. We continue to suppose that $x + q(x, \rho) \in G$ for $x \in G$ and $\rho \in \Gamma$.

In Chapter 7, the dynamic programming equations were defined under specific conditions on the approximating chains. Those conditions are more

than are needed for the convergence theorems, and we need only assume here that the various chains which appear below are stationary. The convergence results are essentially consequences of those of Section 11.1. Let us assume the following condition.

A3.1. *For each small h there is an optimal feedback (x dependent only) control $\bar{u}^h(\cdot)$ for the chain $\{\xi_n^h, n < \infty\}$.*

Condition (A3.1) holds if there is a solution to the Bellman equation (7.5.10) and some conditions which guarantee this are given in Chapter 7. There will be at least one invariant measure under the control $\bar{u}^h(\cdot)$. Choose any one and, in Theorem 3.1 below, let $\{\xi_n^h, n < \infty\}$ and $\psi^h(\cdot)$ be the stationary processes which are associated with the chosen invariant measure and let $m^h(\cdot)$ be the relaxed control representation of $\bar{u}^h(\psi^h(\cdot))$. We write $z(\cdot) = (Y(\cdot), U(\cdot))$.

Theorem 3.1. *Assume the conditions of this section and (A3.1). Then Theorems 1.1 to 1.4 and 1.7 hold. Let h index a weakly convergent subsequence of $\{Z^h(\cdot)\}$ with limit denoted by $Z(\cdot)$. Then $Z(\cdot)$ satisfies (1.1) and (1.2) and the distribution of*

$$x(t), m(t + \cdot) - m(t), z(t + \cdot) - z(t)$$

does not depend on t .

Proof. In view of the cited theorems, only the stationarity needs to be proved. By the stationarity of the chosen $\psi^h(\cdot)$, the distribution of

$$\psi^h(t), m^h(t + \cdot) - m^h(t), z^h(t + \cdot) - z^h(t)$$

does not depend on t . This and the weak convergence (Theorem 1.7) yield the theorem. ■

Convergence of the Costs. Let x be the initial condition of interest. If all states communicate under $\bar{u}^h(\cdot)$, then the invariant measure is unique and the minimum cost $\bar{\gamma}^h$ does not depend on x . Otherwise, the state space $S_h = G_h \cup \partial G_h^+$ is divided into disjoint communicating sets and a set of transient states. In this case, to fix ideas we suppose that x is not a transient state for any h of interest, and use the stationary process associated with the invariant measure for the communicating set in which x lies. Let $\{\xi_n^h, n < \infty\}$ and $\psi^h(\cdot)$ be the optimal stationary discrete and continuous parameter chains, and let h index a weakly convergent subsequence. If x is a transient state for a sequence of $h \rightarrow 0$, then the appropriate stationary process $\psi^h(\cdot)$ can be constructed as a randomization among the stationary processes associated with the various communicating classes, and essentially the same proof can be used.

Using the weak convergence (Theorem 1.7) and the uniform integrability of (1.27), we can write, where $m^h(\cdot)$ is the relaxed control representation of the optimal control $\bar{u}^h(\cdot)$,

$$\begin{aligned}
\bar{\gamma}^h(x) = \gamma^h(x, \bar{u}^h) &= E^{\bar{u}^h} \int_0^1 \left[\int_{\mathcal{U}} k(\psi^h(t), \alpha) m_t^h(d\alpha) dt \right. \\
&\quad \left. + c'(\psi^h(t)) dz^h(t) \right] \\
&\rightarrow E^m \int_0^1 \left[\int_{\mathcal{U}} k(x(t), \alpha) m_t(d\alpha) dt + c'(x(t)) dz(t) \right] \\
&= \lim_T \frac{1}{T} E^m \int_0^T \left[\int_{\mathcal{U}} k(x(t), \alpha) m_t(d\alpha) dt \right. \\
&\quad \left. + c'(x(t)) dz(t) \right] \\
&= \gamma(m) \geq \bar{\gamma},
\end{aligned} \tag{3.2}$$

where $\gamma(m)$ is the cost for the limit stationary process.

As in the previous sections, to complete the proof that

$$\bar{\gamma}^h(x) \rightarrow \bar{\gamma}, \tag{3.3}$$

we need to find a nice comparison control. This is not as easy to do for the ergodic cost problem as it was for the discounted problem or for the other problems where the cost was of interest for essentially a finite time. We will use the following additional condition.

A3.2. *For each $\epsilon > 0$, there is a continuous feedback control $u^\epsilon(\cdot)$ which is ϵ -optimal with respect to all admissible controls and under which the solution to (1.1) and (1.2) is weak sense unique and has a unique invariant measure.*

Remark. The condition does not seem to be too restrictive. For nondegenerate problems, the method of [77] can be applied to the problem with a reflecting boundary to get that there is always a smooth ϵ -optimal control.

Theorem 3.2. *Assume the conditions of this section and (A3.1) and (A3.2). Then (3.3) holds.*

Proof. We use the same conventions concerning the chains and the initial state x as given above (3.2), except that the $u^\epsilon(\cdot)$ of (A3.2) replaces $\bar{u}^h(\cdot)$. Let $\{\xi_n^h, n < \infty\}$ and $\psi^h(\cdot)$ denote the stationary chains under the control $u^\epsilon(\cdot)$. Then, the weak convergence (Theorem 1.7), the uniform integrability

of (1.27), and the stationarity of the limit, yield that

$$\begin{aligned}
\bar{\gamma}^h(x) \leq \gamma^h(x, u^\epsilon) &= E^{u^\epsilon} \int_0^1 [k(\psi^h(t), u^\epsilon(\psi^h(t)))dt \\
&\quad + c'(\psi^h(t))dz^h(t)] \\
&\rightarrow E^{u^\epsilon} \int_0^1 [k(x(t), u^\epsilon(x(t)))dt + c'(x(t))dz(t)] \\
&= \lim_T \frac{1}{T} E^{u^\epsilon} \int_0^T [k(x(t), u^\epsilon(x(t)))dt \\
&\quad + c'(x(t))dz(t)] \\
&= \gamma(u^\epsilon) \leq \bar{\gamma} + \epsilon,
\end{aligned} \tag{3.4}$$

which yields the theorem. ■

Singular Control. Convergence results for the singularly controlled average cost per unit time problem can be obtained by the appropriate combination of the methods of this and of the last section. In [85], an ergodic cost problem of the singular control type originating in the heavy traffic analysis of [123] was dealt with, where there were side “flow through” constraints in addition.

Stationary Measures. The convergence of the stationary processes $\psi^h(\cdot)$ to a stationary process $x(\cdot)$ implies that any weak limit of the stationary measures $\mu^h(\cdot)$ of (7.5.7) is a stationary measure for $x(\cdot)$. This is true for the case where either a fixed continuous feedback control (x dependent only) $u(\cdot)$ is applied to $\{\xi_n^h, n < \infty\}$ for all h , or for the optimally controlled chain. Of course, in the latter case, the limit might not be Markov. If a fixed continuous feedback control $u(\cdot)$ is applied and the stationary measure $\mu(\cdot)$ of $x(\cdot)$ is unique under that control, then $\mu^h(\cdot) \Rightarrow \mu(\cdot)$.

12

Finite Time Problems and Nonlinear Filtering

The problems considered in Chapters 10 and 11 were of interest over an unbounded time interval, and one did not need to keep track of the actual value of the current time. The cost functions were all over an unbounded interval, whether of the discounted form, of the form where control stops when a set is first exited, or of the average cost per unit time type. Time did not enter explicitly into the stopping criterion. All of the results and problem formulations (except for the average cost per unit time formulation) can be readily extended to problems of interest over a given bounded interval only. Owing to the explicit use of a bounded time interval, there are several variations of the locally consistent approximating Markov chains which might be used. They can all be easily derived from those of Chapter 5. The approximating chains are loosely divided into the “explicit” and “implicit” classes, depending on the treatment of time, somewhat analogous to the classification in classical numerical analysis. Section 12.1 contains an example to motivate the general form of the “explicit” method, and the general case is dealt with in Section 12.2. A motivating example for the “implicit” method appears in Section 12.3, and the general case is given in Section 12.4. Various combinations of these methods can be used. An optimal control problem is formulated in Section 12.5, and the numerical questions as well as the convergence of the algorithms is dealt with in Section 12.6. It turns out that the natural analogues of all of the models of control problems of the previous chapters (with the above cited “ergodic” exception) can be dealt with by the previous proofs with little change.

The methods are developed as direct extensions of the Markov chain approximation ideas of the previous chapters. Other possibilities (e.g., operator splitting techniques) have not been explored.

Section 12.7 concerns the problem of numerical approximation for the nonlinear filtering problem. It is shown how the previous Markov chain approximations can be used to get effective numerical approximations under quite weak conditions. For simplicity of notation, the dynamical terms and cost rates will not depend explicitly on time. The alterations required for the time dependent case should be obvious.

12.1 The Explicit Approximation Method: An Example

In Chapters 10 and 11, where Markov chain approximations and algorithms of the type introduced in Chapters 5 and 6, resp., were used, the time interval of interest was essentially unbounded. It was either infinity or else was the time required for the process to exit a given set. The terminal time does not appear explicitly in the algorithms of Section 5.8 or in Chapters 7 and 8. In a sense, the convergence proofs in Chapters 10 and 11 are probabilistic “extensions” of the convergence proofs in numerical analysis for approximations to elliptic partial differential equations. In many cases, the time interval on which the control problem is of interest is bounded by a finite number T , and then the basic numerical problem is analogous to the solution of a (perhaps nonlinear) parabolic partial differential equation. In that case, unless the interpolation interval $\Delta t^h(x, \alpha)$ is independent of x and α , the Markov chain approximations given in Chapter 5 and the algorithms of Chapter 6 cannot be used without some modification.

In classical numerical analysis, the approximations which are in use for elliptic and parabolic equations are closely related, and the common procedures for the parabolic case can be easily obtained from analogous common procedures for the elliptic case. For the parabolic problem the time parameter needs to be accounted for explicitly. The appropriate approximations are easily obtained from those of Chapter 5. When solving parabolic equations by finite difference methods, one can chose either an “explicit” or an “implicit” finite difference approximation. With appropriately chosen difference approximations, each can be viewed as a Markov chain approximation method in much the same sense that the Markov chain approximations of Sections 5.2 and 5.3 could be seen as interpretations of standard finite difference approximations. Loosely speaking, one can divide the Markov chain approximations for the “fixed” terminal time problem into the classes “explicit” and “implicit”. Both classes can be easily obtained from any Markov chain approximation to the underlying diffusion or jump diffusion process which is locally consistent in the sense of Chapter 5.

Let $E_{x,t}^u$ denote the expectation conditioned on $x(t) = x$ and the use of admissible control $u(\cdot)$. In order to facilitate our understanding of how to get the locally consistent Markov chain approximations and algorithms which are appropriate for the finite time problem, we first do a one dimensional example which is a finite time analogue of Example 4 of Section 5.1. The example will motivate the general technique for the explicit method.

Example. Let $T < \infty$ and for $\delta > 0$, let $N_\delta = T/\delta$ be an integer. Let the system $x(\cdot)$ be defined by (5.1.20), and let $k(\cdot)$ and $g(\cdot)$ be smooth and bounded real valued functions. As in Section 5.1, T , $k(\cdot)$, and $g(\cdot)$ play only an auxiliary role in the construction of the approximating chain. For $t < T$

and a feedback control $u(\cdot)$, define the cost function

$$W(x, t, u) = E_{x,t}^u \left[\int_t^T k(x(s), u(x, s)) ds + g(x(T)) \right].$$

More general boundary conditions will be used in Section 12.5 below. By Subsection 3.1.5, the function $W(x, t, u)$ formally satisfies the partial differential equation

$$W_t(x, t, u) + \mathcal{L}^{u(x,t)} W(x, t, u) + k(x, u(x, t)) = 0, \quad t < T, \quad (1.1)$$

with boundary condition $W(x, T, u) = g(x)$, and where the differential operator \mathcal{L}^α is defined by

$$\mathcal{L}^\alpha f(x) = b(x, \alpha) \frac{\partial f(x)}{\partial x} + \frac{1}{2} \sigma^2(x) \frac{\partial^2 f(x)}{\partial x^2}.$$

We next define a finite difference approximation to (1.1), with interval $h > 0$. For concreteness, we follow (5.1.8) and use the “one sided” difference approximation of the first derivative W_x . The two sided difference can also be used if $\sigma^2(x) - h|b(x, u(x))| \geq 0$ for all x , as below (5.1.17). We use the “explicit” forms

$$\begin{aligned} f_t(x, t) &\rightarrow \frac{f(x, t + \delta) - f(x, t)}{\delta} \\ f_x(x, t) &\rightarrow \frac{f(x + h, t + \delta) - f(x, t + \delta)}{h} \text{ if } b(x, u(x, t)) \geq 0, \\ f_x(x, t) &\rightarrow \frac{f(x, t + \delta) - f(x - h, t + \delta)}{h} \text{ if } b(x, u(x, t)) < 0, \\ f_{xx}(x, t) &\rightarrow \frac{f(x + h, t + \delta) + f(x - h, t + \delta) - 2f(x, t + \delta)}{h^2}. \end{aligned} \quad (1.2)$$

Substituting (1.2) into (1.1), letting $W^{h,\delta}(x, t, u)$ denote the solution to the finite difference equation with x an integral multiple of h and $n\delta < T$, collecting terms, and multiplying all terms by δ yields the expression

$$\begin{aligned} W^{h,\delta}(x, n\delta, u) &= W^{h,\delta}(x, n\delta + \delta, u) \left[1 - \sigma^2(x) \frac{\delta}{h^2} - |b(x, u(x, n\delta))| \frac{\delta}{h} \right] \\ &\quad + W^{h,\delta}(x + h, n\delta + \delta, u) \left[\frac{\sigma^2(x)}{2} \frac{\delta}{h^2} + b^+(x, u(x, n\delta)) \frac{\delta}{h} \right] \\ &\quad + W^{h,\delta}(x - h, n\delta + \delta, u) \left[\frac{\sigma^2(x)}{2} \frac{\delta}{h^2} + b^-(x, u(x, n\delta)) \frac{\delta}{h} \right] \\ &\quad + k(x, u(x, n\delta)) \delta, \end{aligned} \quad (1.3)$$

with the boundary condition $W^{h,\delta}(x, T) = g(x)$. The method is called the explicit method, because the equation (1.3) can be solved recursively: The values of $W^{h,\delta}$ at time $n\delta$ can be obtained by a simple iteration from the values at time $n\delta + \delta$. I.e., the time variable on the right side of (1.3) equals $n\delta + \delta$ and it is $n\delta$ on the left side. This is due to the fact that the spatial derivatives are approximated at time $n\delta + \delta$.

Note that the sum of the coefficients of the $W^{h,\delta}$ terms in (1.3) is unity. Suppose that the coefficient

$$\left[1 - \sigma^2(x) \frac{\delta}{h^2} - |b(x, u(x, n\delta))| \frac{\delta}{h} \right]$$

of the $W^{h,\delta}(x, n\delta + \delta, u)$ term is non-negative. Then the coefficients can be considered to be the transition function of a Markov chain. Defining $p^{h,\delta}(x, y|\alpha)$ in the obvious way, we then rewrite (1.3) as

$$W^{h,\delta}(x, n\delta, u) = \sum_y p^{h,\delta}(x, y|u(x, n\delta)) W^{h,\delta}(y, n\delta + \delta, u) + k(x, u(x, n\delta))\delta. \quad (1.4)$$

The $p^{h,\delta}(x, y|\alpha)$ are the transition probabilities for a controlled Markov chain. Let the associated Markov chain be denoted by $\{\xi_n^{h,\delta}, n < \infty\}$. Note that

$$E_{x,n}^{h,\alpha} \Delta \xi_n^{h,\delta} = b(x, \alpha)\delta,$$

$$\text{cov}_{x,n}^{h,\alpha} \Delta \xi_n^{h,\delta} = \sigma^2(x)\delta + O(h\delta).$$

Let $\delta \rightarrow 0$ and $h \rightarrow 0$ together. By analogy to the definition (4.1.3), we say that the “explicit” controlled Markov chain $\{\xi_n^{h,\delta}, n < \infty\}$ with interpolation interval $\delta > 0$ is *locally consistent* with $x(\cdot)$ if

$$\begin{aligned} E_{x,n}^{h,\alpha} \Delta \xi_n^{h,\delta} &= b(x, \alpha)\delta + o(\delta), \\ \text{cov}_{x,n}^{h,\alpha} \Delta \xi_n^{h,\delta} &= \sigma^2(x)\delta + o(\delta). \end{aligned} \quad (1.5)$$

Thus, the constructed chain and interpolation interval are locally consistent with (5.1.20).

Define the piecewise constant continuous parameter interpolation $\xi^{h,\delta}(\cdot)$ by $\xi^{h,\delta}(t) = \xi_n^{h,\delta}$ on the interval $[n\delta, n\delta + \delta]$. (See Section 4.2.) With the boundary condition $W^{h,\delta}(x, T, u) = g(x)$ and $t = n\delta < T$, the solution to (1.4) can be written as

$$\begin{aligned} W^{h,\delta}(x, t, u) &= E_{x,n}^u \left[\sum_n^{N_\delta-1} k(\xi_i^{h,\delta}, u(\xi_i^{h,\delta}, i\delta))\delta + g(\xi_{N_\delta}^{h,\delta}) \right] \\ &= E_{x,t}^u \left[\int_t^T k(\xi^{h,\delta}(s), u(\xi^{h,\delta}(s), s))ds + g(\xi^{h,\delta}(T)) \right]. \end{aligned}$$

This expression is an approximation to $W(x, t, u)$ if the process $\xi^{h,\delta}(\cdot)$ is an approximation to $x(\cdot)$.

One can also define an analogue $\psi^{h,\delta}(\cdot)$ of the continuous parameter Markov chain interpolation $\psi^h(\cdot)$ of Section 4.3 and an appropriate cost function. The details are left to the reader. The continuous parameter interpolation is just an intermediary in getting the convergence of the cost functions, and either interpolation can be used with the same results.

Remarks. The example shows that one can treat the fixed terminal time problem quite similarly to the unbounded terminal time problem. Notice the following important point. Let $p^h(x, y|\alpha)$ be the transition probabilities in Section 5.1, Example 4. Then, for $y \neq x$,

$$p^{h,\delta}(x, y|\alpha) = p^h(x, y|\alpha) \times \text{normalization}(x), \quad (1.6)$$

where

$$\begin{aligned} \text{normalization}(x) &= 1 - p^{h,\delta}(x, x|\alpha) \\ &= \frac{\delta}{h^2} Q^h(x, \alpha), \end{aligned}$$

where $Q^h(x, \alpha) = \sigma^2(x) + h|b(x, \alpha)|$, the normalization in (5.1.22).

A similar relationship exists in the multidimensional problem and will be pursued and formalized in the next section. The equivalence exists because the only difference between this example and Example 4 of Section 5.1 is the presence of the W_t in (1.1), which causes a transition of each x to itself owing to the form of the finite difference approximation of the first line of (1.2). In general, as will be seen in the next section, we need not use a finite difference based method to get the explicit Markov chain approximation. They can be obtained directly from the approximations of Chapter 5.

12.2 The General Explicit Approximation Method

The above one dimensional example and special state space can be easily generalized by using the observations made in Section 12.1 as a guide and applying them to any of the calculated transition probabilities of Chapter 5 whether of the finite difference or of the so-called direct method types. The method is referred to as “explicit,” because of the analogy with the example of Section 12.1. Time advances by δ at each step. Suppose that we are given transition probabilities $p^h(x, y|\alpha)$ and an interpolation interval $\Delta t^h(x, \alpha)$ which are locally consistent with the controlled diffusion (5.3.1) or jump diffusion (5.6.1). Let the desired interpolation interval $\delta > 0$ be given. Let $p^{h,\delta}(x, y|\alpha)$ (to be defined below) denote the transition probabilities for the approximating chain for the explicit method. Then, following the guide of the example in Section 12.1, for $x \neq y$ we wish to have the “proportionality”

$$\frac{p^{h,\delta}(x, y|\alpha)}{1 - p^{h,\delta}(x, x|\alpha)} = p^h(x, y|\alpha). \quad (2.1)$$

That is, under the control parameter α , and given that $\xi_n^{h,\delta} = x$, the probability that $\xi_{n+1}^{h,\delta} = y \neq x$ equals $p^h(x, y|\alpha)$. Let $\delta \rightarrow 0$ and $h \rightarrow 0$ together. Then, the requirement that both pairs $(p^h(x, y|\alpha), \Delta t^h(x, \alpha))$ and $(p^{h,\delta}(x, y|\alpha), \delta)$ be locally consistent implies that, modulo small error terms,

$$\begin{aligned} b(x, \alpha)\delta &= \sum_y (y - x)p^{h,\delta}(x, y|\alpha) \\ &= \sum_y (y - x)p^h(x, y|\alpha)(1 - p^{h,\delta}(x, x|\alpha)) \\ &= b(x, \alpha)\Delta t^h(x, \alpha)(1 - p^{h,\delta}(x, x|\alpha)). \end{aligned} \quad (2.2)$$

Equation (2.2) implies that (modulo small error terms)

$$1 - p^{h,\delta}(x, x|\alpha) = \frac{\delta}{\Delta t^h(x, \alpha)} \in (0, 1]. \quad (2.3)$$

Thus, given $(p^h(x, y|\alpha), \Delta t^h(x, \alpha))$ and a $\delta > 0$ satisfying

$$\delta \leq \min_{x, \alpha} \Delta t^h(x, \alpha), \quad (2.4)$$

we get the transition probabilities $p^{h,\delta}$ for the explicit method and the associated approximating Markov chain $\{\xi_n^{h,\delta}, n < \infty\}$ from (2.1) and (2.3).

12.3 The Implicit Approximation Method: An Example

The fundamental difference between the so-called explicit and implicit approaches to the Markov chain approximation lies in the fact that in the former the time variable is treated differently than the state variables: It is a true “time” variable, and its value increases by a constant δ at each step. In the implicit approach, the time variable is treated as just another state variable. It is discretized in the same manner as are the other state variables: The approximating Markov chain has a state space which is a discretization of the (x, t) -space, and the component of the state of the chain which comes from the original time variable does not necessarily increase its value at each step. The basic approach for the implicit scheme is best illustrated by a comparison of the explicit and implicit finite difference approximation to the solution of (1.1). Instead of (1.2), use the approximations (again, using a one sided difference approximation for W_x for specificity)

$$\begin{aligned}
f_t(x, t) &\rightarrow \frac{f(x, t + \delta) - f(x, t)}{\delta} \\
f_x(x, t) &\rightarrow \frac{f(x + h, t) - f(x, t)}{h} \text{ if } b(x, u(x, t)) \geq 0, \\
f_x(x, t) &\rightarrow \frac{f(x, t) - f(x - h, t)}{h} \text{ if } b(x, u(x, t)) < 0, \\
f_{xx}(x, t) &\rightarrow \frac{f(x + h, t) + f(x - h, t) - 2f(x, t)}{h^2}.
\end{aligned} \tag{3.1}$$

Note that the last three equations of (3.1) use the value t on the right side rather than $t + \delta$ as in (1.2). Using (3.1) and repeating the procedure which led to (1.4), for $n\delta < T$ we get the standard finite difference approximation

$$\begin{aligned}
&\left[1 + \sigma^2(x) \frac{\delta}{h^2} + |b(x, u(x, n\delta))| \frac{\delta}{h} \right] W^{h,\delta}(x, n\delta, u) \\
&= \left[\frac{\sigma^2(x)}{2} \frac{\delta}{h^2} + b^+(x, u(x, n\delta)) \frac{\delta}{h} \right] W^{h,\delta}(x + h, n\delta, u) \\
&+ \left[\frac{\sigma^2(x)}{2} \frac{\delta}{h^2} + b^-(x, u(x, n\delta)) \frac{\delta}{h} \right] W^{h,\delta}(x - h, n\delta, u) \\
&+ W^{h,\delta}(x, n\delta + \delta, u) + k(x, u(x, n\delta))\delta.
\end{aligned}$$

With the obvious definitions of $\hat{p}^{h,\delta}$ and $\Delta\hat{t}^{h,\delta}$, we can rewrite the above expression as

$$\begin{aligned}
W^{h,\delta}(x, n\delta, u) &= \sum_y \hat{p}^{h,\delta}(x, n; y, n|u(x, n\delta)) W^{h,\delta}(y, n\delta, u) \\
&+ \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|u(x, n\delta)) W^{h,\delta}(x, n\delta + \delta, u) \\
&+ k(x, u(x, n\delta)) \Delta\hat{t}^{h,\delta}(x, u(x, n\delta)).
\end{aligned} \tag{3.2}$$

The $\hat{p}^{h,\delta}$ are non-negative and

$$\sum_y \hat{p}^{h,\delta}(x, n\delta; y, n\delta|\alpha) + \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha) = 1.$$

It can be seen from this that we can consider the $\hat{p}^{h,\delta}$ as a one step transition probability of a Markov chain $\{\hat{\zeta}_n^{h,\delta}, n < \infty\}$ on the “ (x, t) –state space”

$$\{0, \pm h, \pm 2h, \dots\} \times \{0, \delta, 2\delta, \dots\}.$$

It is evident that time is being considered as just another state variable. Refer to Figure 12.2b for an illustration of the transitions. Note that, analogously to (1.6), for $x \neq y$ we have

$$\hat{p}^{h,\delta}(x, n\delta; y, n\delta|\alpha) = p^h(x, y|\alpha) \times \text{normalization}(x), \quad (3.3)$$

where the $p^h(x, y|\alpha)$ are the transition probabilities of Example 4 of Section 4.1. Analogously to what was done in Section 12.2, this relationship will be used in Section 12.4 to get a general implicit Markov chain approximation, starting with any consistent (in the sense of Chapter 4) approximation.

Write $\zeta_n^{h,\delta} = (\zeta_n^{h,\delta}, \zeta_{n,0}^{h,\delta})$, where the 0^{th} component $\zeta_{n,0}^{h,\delta}$ represents the time variable, and $\zeta_n^{h,\delta}$ represents the original “spatial” state. Then we have

$$E_{x,n}^{\alpha,h} \Delta \zeta_n^{h,\delta} = b(x, \alpha) \Delta \hat{t}^{h,\delta}(x, \alpha),$$

$$\text{cov}_{x,n}^{\alpha,h} \Delta \zeta_n^{h,\delta} = \sigma^2(x) \Delta \hat{t}^{h,\delta}(x, \alpha) + \Delta \hat{t}^{h,\delta}(x, \alpha) O(h),$$

$$E_{x,n}^{h,\alpha} \Delta \zeta_{n,0}^{h,\delta} = \Delta \hat{t}^{h,\delta}(x, \alpha).$$

Thus, the “spatial” component of the controlled chain is locally consistent with (5.1.20). The conditional mean increment of the “time” component of the state is $\Delta \hat{t}^{h,\delta}(x, \alpha)$. We have constructed an approximating Markov chain via an “implicit” method. It is called an implicit method because (3.2) cannot be solved by a simple backward iteration. At each n , (3.2) determines the values of the $\{W^{h,\delta}(x, n\delta, u)\}$ implicitly.

12.4 The General Implicit Approximation Method

With the illustration of the above section in mind, we now describe the general method. The procedure follows the general idea used in Section 12.2. As in Section 12.2, let $p^h(x, y|\alpha)$ and $\Delta t^h(x, \alpha)$ be a transition function and interpolation interval which are locally consistent with either the controlled diffusion (5.3.1) or the controlled jump diffusion (5.6.1). Let $\delta > 0$ be given. Let $\hat{p}^{h,\delta}$ and $\Delta \hat{t}^{h,\delta}$ denote the (to be defined) transition probability and interpolation interval for the implicit approximation. Analogously to (2.1), the example in Section 12.3 suggests that, for $x \neq y$, we use the relationship

$$p^h(x, y|\alpha) = \frac{\hat{p}^{h,\delta}(x, n\delta; y, n\delta|\alpha)}{1 - \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)}. \quad (4.1)$$

Thus, to get the transition probabilities $\hat{p}^{h,\delta}$, we need only get

$$\hat{p}^{h,\delta}(x, n\delta; x, n\delta + n\delta|\alpha).$$

This will be done via the local consistency requirements on both $(p^h(x, y|\alpha), \Delta t^h(x, \alpha))$ and the corresponding quantities for the implicit approximation. The conditional mean one step increment of the “time” component of $\zeta_n^{h,\delta}$ is

$$E_{x,n}^{h,\alpha} \Delta \zeta_{n,0}^{h,\delta} = \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)\delta, \quad (4.2)$$

and we define the interpolation interval $\Delta \hat{t}^{h,\delta}(x, \alpha)$ by (4.2). Of course, we can always add a term of smaller order. The consistency equation for the spatial component $(\zeta_n^{h,\delta})$ of the chain is (modulo a negligible error term)

$$\begin{aligned} b(x, \alpha) \Delta t^h(x, \alpha) &= \sum_y (y - x) p^h(x, y|\alpha) \\ &= \sum_y (y - x) \frac{\hat{p}^{h,\delta}(x, n\delta; y, n\delta|\alpha)}{1 - \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)} \\ &= b(x, \alpha) \frac{\Delta \hat{t}^{h,\delta}(x, \alpha)}{1 - \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)}. \end{aligned} \quad (4.3)$$

By equating the two expressions (4.2) and (4.3) for the interpolation interval $\Delta \hat{t}^{h,\delta}(x, \alpha)$, we have

$$\Delta \hat{t}^{h,\delta}(x, \alpha) = (1 - \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)) \Delta t^h(x, \alpha) = \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)\delta.$$

Thus,

$$\begin{aligned} \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha) &= \frac{\Delta t^h(x, \alpha)}{\Delta t^h(x, \alpha) + \delta}, \\ \Delta \hat{t}^{h,\delta}(x, \alpha) &= \frac{\delta \Delta t^h(x, \alpha)}{\Delta t^h(x, \alpha) + \delta}. \end{aligned} \quad (4.4)$$

A short calculation shows that this equals what we obtained in the example of Section 12.3.

Note that $\Delta \hat{t}^{h,\delta}(x, \alpha)$ goes to $\Delta t^h(x, \alpha)$ as $\delta/\Delta t^h(x, \alpha)$ goes to infinity. In particular, suppose that $\Delta t^h(x, \alpha) = O(h^2)$ and $\delta = O(h)$, as is common. Then

$$\Delta \hat{t}^{h,\delta}(x, \alpha) = \Delta t^h(x, \alpha)(1 + O(h)).$$

Generally, the implicit method provides a more accurate approximation than the explicit method.

The Solution to (3.2): Vector Case. Because the interpolation times are not necessarily constant, it is not necessarily possible to represent the solution to (3.2) as a sum of functions of a nonrandom number of steps of the approximating chain. But the solution can be represented in terms of the path up to the first time that the time component $\zeta_{n,0}^{h,\delta}$ reaches or exceeds the value T . Define the stopping time

$$N^{h,\delta}(T) = \min\{n : \zeta_{n,0}^{h,\delta} \geq T\}.$$

Let $u = \{u_n^{h,\delta}, n < \infty\}$ be an admissible control sequence for $\{\zeta_n^{h,\delta}, n < \infty\}$. Define $\Delta \hat{t}_n^{h,\delta} = \Delta \hat{t}^{h,\delta}(\zeta_n^{h,\delta}, u_n^{h,\delta})$, and let $E_{x,n}^u$ denote the expectation given use of u and $\zeta_n^{h,\delta} = x, \zeta_{n,0}^{h,\delta} = n\delta$. Then the solution to (3.2) with the boundary condition $W^{h,\delta}(x, T, u) = g(x)$ can be written as

$$W^{h,\delta}(x, n\delta, u) = E_{x,n}^u \left[\sum_{i=n}^{N^{h,\delta}(T)-1} k(\zeta_i^{h,\delta}, u_i^{h,\delta}) \Delta \hat{t}_i^{h,\delta} + g(\zeta_{N^{h,\delta}(T)}^{h,\delta}) \right]. \quad (4.5)$$

An equivalent representation to (4.5) is

$$W^{h,\delta}(x, n\delta, u) = E_{x,n}^u \left[\sum_{i=n}^{N^{h,\delta}(T)-1} k(\zeta_i^{h,\delta}, u_i^{h,\delta}) [\zeta_{i+1,0}^{h,\delta} - \zeta_{i,0}^{h,\delta}] + g(\zeta_{N^{h,\delta}(T)}^{h,\delta}) \right]. \quad (4.6)$$

Next, define the “interpolated” times $\hat{t}_n^{h,\delta} = \sum_0^{n-1} \Delta \hat{t}_i^{h,\delta}$, and define the continuous parameter interpolations $\hat{\zeta}^{h,\delta}(\cdot) = (\zeta^{h,\delta}(\cdot), \zeta_0^{h,\delta}(\cdot))$ and $u^{h,\delta}(\cdot)$ by

$$u^{h,\delta}(t) = u_n^{h,\delta}, \quad \hat{\zeta}^{h,\delta}(t) = \hat{\zeta}_n^{h,\delta} \quad \text{for } t \in [\hat{t}_n^{h,\delta}, \hat{t}_n^{h,\delta} + \Delta \hat{t}_n^{h,\delta}). \quad (4.7)$$

Then, for $t = n\delta$, we can write (4.5) and (4.6) as, resp.,

$$\begin{aligned} & E_{x,t}^u \left[\int_t^T k(\zeta^{h,\delta}(s), u^{h,\delta}(s)) ds + g(\zeta^{h,\delta}(T)) \right], \\ & E_{x,t}^u \left[\int_t^T k(\zeta^{h,\delta}(s^-), u^{h,\delta}(s^-)) d\zeta_0^{h,\delta}(s) + g(\zeta^{h,\delta}(T)) \right]. \end{aligned}$$

An Alternative Approximating Chain. There is an alternative Markov chain approximation $\{\tilde{\zeta}_n^{h,\delta}\}$ which is useful when there is no control and will be used for the “implicit” approximation to the optimal nonlinear filter in Section 12.7 below. To get this new chain, we just look at $\zeta_n^{h,\delta}$ at the times that $\zeta_{n,0}^{h,\delta}$ changes. In more detail, define $v_0 = 0$, and, for $n > 0$ define

$$v_n = \min\{i > v_{n-1} : \zeta_{i,0}^{h,\delta} - \zeta_{i-1,0}^{h,\delta} = \delta\}.$$

Then define $\tilde{\zeta}_n^{h,\delta} = \zeta_{v_n}^{h,\delta}$. Define the continuous parameter interpolation $\tilde{\zeta}^{h,\delta}(t) = \tilde{\zeta}_n^{h,\delta}$ for $t \in [n\delta, n\delta + \delta]$.

Suppose that there is no control and denote the one step transition probability of $\{\tilde{\zeta}_n^{h,\delta}\}$ by $\tilde{p}^{h,\delta}(x, y)$. This can be calculated as follows. For any function $g(\cdot)$, define $U(\cdot)$ by

$$U(x) = \sum_y \hat{p}^{h,\delta}(x, n\delta; y, n\delta) U(y) + \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta) g(x). \quad (4.8)$$

The solution can be written as

$$U(x) = \sum_y \tilde{p}^{h,\delta}(x, y)g(y). \quad (4.9)$$

Suitable adaptations of the methods of Chapter 6 can be used to get the $\tilde{p}^{h,\delta}(x, y)$, and solutions of equations such as (4.8).

12.5 The Optimal Control Problem: Approximations and Dynamic Programming Equations

The Cost Function. The system will be the controlled diffusion (5.3.1) or the jump diffusion (5.6.1). All of the problem formulations of Chapters 10 and 11 can be carried over to the finite time case and the proofs require no change, given the appropriate locally consistent transition probabilities. However, for illustrative purposes, we next set up the computational problem up for the case of an absorbing boundary. Assume:

A5.1. *G is a compact set in \mathbb{R}^r and is the closure of its interior. $k(\cdot)$ is a bounded continuous function on $G \times [0, T] \times \mathcal{U}$, and $g(\cdot)$ is a bounded and continuous function on $([\mathbb{R}^r - G^0] \times [0, T]) \cup (\mathbb{R} \times \{T\})$.*

Define $\tau = \min\{t : x(t) \notin G^0\}$. For an ordinary admissible control $u(\cdot)$, define the cost function

$$W(x, t, u) = E_{x,t}^u \left[\int_t^{T \wedge \tau} k(x(s), s, u(s))ds + g(x(T \wedge \tau), T \wedge \tau) \right].$$

For an admissible relaxed control $m(\cdot)$, the cost is written as

$$W(x, t, m) = E_{x,t}^m \left[\int_t^{T \wedge \tau} \int_{\mathcal{U}} k(x(s), s, \alpha)m(ds d\alpha) + g(x(T \wedge \tau), T \wedge \tau) \right]. \quad (5.1)$$

Define the optimal cost, where the infimum is over all admissible controls of the indicated type

$$V(x, t) = \inf_u W(x, t, u) = \inf_m W(x, t, m).$$

The formal dynamic programming equation for the minimal cost is

$$V_t(x, t) + \min_{\alpha \in \mathcal{U}} [\mathcal{L}^\alpha V(x, t) + k(x, t, \alpha)] = 0, \quad (5.2)$$

for $x \in G^0$ and $t < T$, with the boundary conditions $V(x, t) = g(x, t)$, $x \notin G^0$, or $t \geq T$.

The Computational Approximation: Explicit Method. Let $\delta > 0$ satisfy (2.4) and let $p^{h,\delta}(x, y|\alpha)$ be the transition function of a locally consistent approximating Markov chain as derived in Section 12.2. Let $T = N_\delta\delta$ for some integer N_δ . Then the dynamic programming equation for the approximating optimal control problem is

$$V^{h,\delta}(x, n\delta) = \min_{\alpha \in \mathcal{U}} \left[\sum_y p^{h,\delta}(x, y|\alpha) V^{h,\delta}(y, n\delta + \delta) + k(x, n\delta, \alpha)\delta \right], \quad (5.3)$$

for $x \in G_h^0$, $n\delta < T$, and with the same boundary condition as for (5.2).

The Computational Approximation: Implicit Method. Let

$$\hat{p}^{h,\delta}(x, n\delta; y, n\delta|\alpha), \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)$$

be the transition function of an approximating Markov chain for the implicit method as derived in Section 12.4, with the interpolation interval $\Delta\hat{t}^{h,\delta}(x, \alpha)$. Then the dynamic programming equation for the approximating optimal control problem is

$$V^{h,\delta}(x, n\delta) = \min_{\alpha \in \mathcal{U}} \left[\sum_y \hat{p}^{h,\delta}(x, n\delta; y, n\delta|\alpha) V^{h,\delta}(y, n\delta) + \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha) V^{h,\delta}(x, n\delta + \delta) + k(x, n\delta, \alpha) \Delta\hat{t}^{h,\delta}(x, \alpha) \right] \quad (5.4)$$

for $x \in G_h^0$ and $n\delta < T$, and with the same boundary condition as for (5.2).

Combinations of the Explicit and Implicit Method. The explicit Markov chain approximation is a special case of the implicit approximation, because we can always add a “time” component (called $\xi_n^{h,\delta}$) to $\xi_n^{h,\delta}$ and define the extended state $\hat{\xi}_n^{h,\delta} = (\xi_n^{h,\delta}, \xi_{n,0}^{h,\delta})$, where $\Delta\xi_{n,0}^{h,\delta} = \delta$ for all n .

The two methods can be combined. One can randomize among them. For example, let $p_0 \in [0, 1]$ and at each step choose the explicit method with probability p_0 and the implicit with probability $(1 - p_0)$. Also, p_0 can depend on the current value of the state. Whether there is any value to such combinations is not clear.

12.6 Methods of Solution, Decomposition and Convergence

Solving the Explicit Equation (5.3). Because the minimum cost function values $V^{h,\delta}(y, n\delta + \delta)$ at time $n\delta + \delta$ are on the right side of (5.3), while the value at state x at time $n\delta$ is on the left side, (5.3) is solved by a simple backward iteration.

Solving the Implicit Equation (5.4). In (5.4), if the time on the left is $n\delta$ and the state there is x , then the only value of the minimum cost function at time $n\delta + \delta$ which is on the right side is $V^{h,\delta}(x, n\delta + \delta)$. The $V^{h,\delta}(y, n\delta)$ on the right, for $y \neq x$, are evaluated at time $n\delta$. Hence, (5.4) cannot be solved by a simple backward iteration. However, all of the methods of Chapter 6 for solving (6.1.1) or (6.1.2) can be used. The general approach is as follows: Let $N\delta = T$. Starting with time N , we have the boundary condition $V^{h,\delta}(x, N\delta) = g(x, N\delta)$. Suppose that $V^{h,\delta}(x, n\delta + \delta)$ is available for some $n < N$. Then there is a $\hat{C}(\cdot)$ such that we can rewrite (5.4) as

$$V^{h,\delta}(x, n\delta) = \min_{\alpha \in \mathcal{U}} \left[\sum_y \hat{p}^{h,\delta}(x, n\delta; y, n\delta | \alpha) V^{h,\delta}(y, n\delta) + \hat{C}^h(x, n+1, \alpha) \right] \quad (6.1)$$

for $x \in G_h$ and with the boundary condition $V^{h,\delta}(x, n\delta) = g(x)$ for $x \notin G_h^0$. The $\hat{C}^h(x, n, \cdot)$ are continuous in α .

By the construction of the transition probabilities, $\hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta | \alpha) > 0$ for all α, x . Hence,

$$\sum_y \hat{p}^{h,\delta}(x, n\delta; y, n\delta | \alpha) < 1,$$

for all α and x . This and the continuity of the transition probabilities implies that the effective transition function in (6.1) is a contraction for all feedback controls, and that the contraction is uniform in the control. Now simply apply to (6.1) any of the methods of Chapter 6 which can be used for (6.1.1) or (6.1.2).

Decomposition Methods. The parallel processing and decomposition methods which can be used for solving (6.1.1) or (6.1.2) can also be used for (5.4) or (6.1). One possible implementation will be illustrated on a one dimensional problem. The state space is illustrated in Figure 12.1. The spatial interval is $[a, c]$ which is a multiple of h and is discretized by grid points h units apart. Also suppose that $(b - a)$ is an integral multiple of h . The explicit method will be used for updating the values of the minimum value function at b . The transitions are illustrated in Figure 12.2a, and are assumed to be to the nearest neighbors only (for our illustrative example).

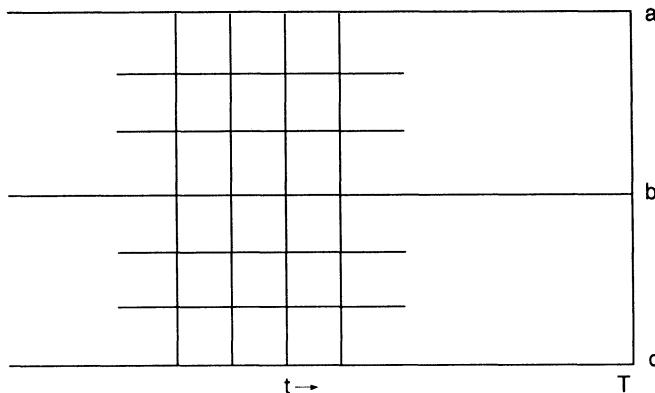
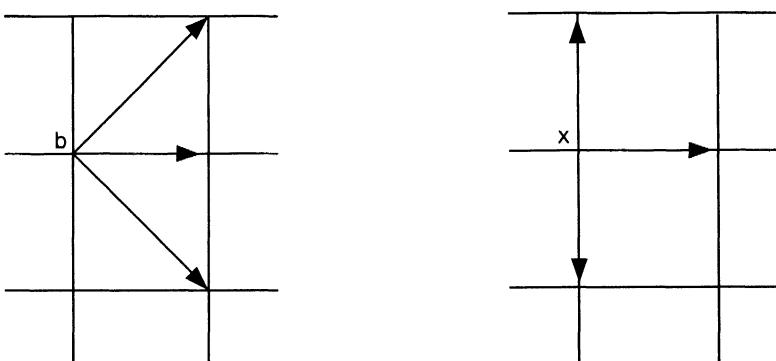


Figure 12.1. The decomposition regions.



(a) The explicit method.

(b) The implicit method; $x \neq a, b, c$.
Figure 12.2.

We suppose that the transition probabilities at points $x \neq b$ are as for the general implicit method of Section 12.4 and are illustrated in Figure 12.2b.

Now the computational procedure is as follows: For given n , let the values $V^{h,\delta}(x, n\delta + \delta)$ be given for all x . Calculate $V^{h,\delta}(b, n\delta)$ in terms of these given values via the explicit method. Then, at time $n\delta$, the values of $V^{h,\delta}(x, n\delta)$ are known, except for $x \in G_h^0 \cap (a, b)$ and $G_h^0 \cap (b, c)$ and the two halves can be worked on simultaneously via the implicit method.

The sequence $V^{h,\delta}(x, n\delta)$ can be shown to converge to the correct value $V(x, t)$ as $h \rightarrow 0$, $\delta \rightarrow 0$, and $n\delta \rightarrow t$. Clearly, more than one subdivision of the state space can be used and the procedure can also be used in higher dimensions. We have no actual computational experience with the method.

Convergence Theorems. Under their respective conditions, the finite

time analogues of all of the theorems in Chapters 10 and 11 hold. We need only extend the definition of the functions which might depend on time, as done in (A5.1) above.

Alternative Methods. In a series of interesting papers, Hanson [21, 22, 23] has explored the use of Galerkin-type procedures, splitting methods and parallel computation for problems where the cost is quadratic in the control and the dynamical term $b(x, \alpha)$ is of the form $b_0(x)\alpha$. The proofs of convergence require more regularity than we need here, but the results show considerable promise. At present, it is not clear when those classes of methods have “Markov chain” interpretations.

12.7 Nonlinear Filtering

The nonlinear filtering problem is one of the basic problems in current stochastic control and communication theory. The Markov chain approximation method is well suited for the computation of approximations to the optimum filters. We first discuss the problem of numerically approximating the distribution of $x(t)$, an uncontrolled jump diffusion. Then, approximations to the optimal nonlinear filter for a jump diffusion model with white noise corrupted observations will be developed.

12.7.1 APPROXIMATION TO THE SOLUTION OF THE FOKKER-PLANCK EQUATION

We start with a few formal remarks. Let $x(\cdot)$ be defined by

$$dx = b(x)dt + \sigma(x)dw, \quad x(0) = x.$$

Let \mathcal{L}^* denote the formal adjoint of the differential operator \mathcal{L} of $x(\cdot)$. Under appropriate regularity conditions [40, 44, 49], $x(t)$ has a density $p(\cdot, \cdot)$ and it is the solution to the Fokker-Planck or “forward” Kolmogorov equation

$$p_t(y, t) = \mathcal{L}^* p(y, t) = \frac{1}{2} \sum_{i,j} \frac{\partial^2 (a_{ij}(y)p(y, t))}{\partial y_i \partial y_j} - \sum_i \frac{\partial (b_i(y)p(y, t))}{\partial y_i}, \quad (7.1)$$

with the initial condition $p(y, 0) = \delta(x - y)$, where $\delta(\cdot)$ is the Dirac delta function; that is, $x(0)$ is concentrated at the initial point x . There is a similar equation for the jump diffusion case

$$dx = b(x)dt + \sigma(x)dw + \int_{\Gamma} q(x(t^-), \gamma)N(dtd\gamma). \quad (7.2)$$

For many problems of interest, (7.1) has only a formal meaning, because either the density does not exist or it is not smooth enough to satisfy (7.1).

Henceforth, when referring to the density, we mean the weak sense density or (equivalently) the distribution function. We next discuss the problem of the approximate calculation of the density. For numerical purposes, it is usually necessary to work in a bounded state space. The state space G might be bounded either naturally or because a bound was imposed for numerical reasons. For example, the process dynamics can be such that the process never leaves some bounded set, even without reflection or absorption: The process might be reflected back from the boundary of G , or it might be stopped or killed once it leaves the interior G^0 . The exact form of the process is not important here, provided only that we work with appropriately consistent approximations. Because the process is uncontrolled, we drop the control parameter in the notation for the approximating transition probability $p^{h,\delta}$. For purposes of simplicity in the exposition and because the boundary condition plays only a secondary role, we make the assumption in the general form (A7.1) below. The details can be filled in by referring to the weak convergence results for the various cases of Chapters 10 and 11.

A7.1. *For $\delta > 0$ and $h > 0$, the function $p^{h,\delta}(x, y)$ is a transition function for an uncontrolled locally consistent explicit Markov chain approximation to a diffusion or jump diffusion $x(\cdot)$ of the type discussed in Section 12.2. The process $x(\cdot)$ might be constrained to the bounded set G by a reflecting boundary condition, or it might be killed on first exit from G^0 . The continuous parameter interpolation $\xi^{h,\delta}(\cdot)$ converges weakly to $x(\cdot)$ as $h, \delta \rightarrow 0$.*

Let $p^{h,\delta}(x, n\delta, y)$ denote the n -step transition function. Then, as $(\delta, h) \rightarrow 0$ and $n\delta \rightarrow t \geq 0$, we have

$$p^{h,\delta}(x, n\delta, \cdot) \rightarrow p(x, t, \cdot)$$

in the sense that for any bounded and continuous real valued function $\phi(\cdot)$,

$$\sum_y \phi(y) p^{h,\delta}(x, n\delta, y) \rightarrow \int \phi(y) p(x, t, y) dy. \quad (7.3)$$

This is just a consequence of the weak convergence. Thus, explicit Markov chain approximations of the type introduced in Section 12.2 can be used to get approximations to the weak sense density under quite broad conditions. It follows from Section 11.3 that the limit of any weakly convergent subsequence of invariant measures for the chains $\{\xi_n^{h,\delta}, n < \infty\}$ is an invariant measure of $x(\cdot)$. The alternative implicit chains $\{\tilde{\zeta}_n^{h,\delta}, n < \infty\}$ defined in Section 12.4 can also be used.

12.7.2 THE NONLINEAR FILTERING PROBLEM: INTRODUCTION AND REPRESENTATION

A standard nonlinear filtering problem [41, 73, 93, 128] uses a jump diffusion process model $x(\cdot)$ and where the data available at time t is $\mathcal{Y}(t) \equiv \{y(s), s \leq t\}$, where

$$dy = g(x)dt + dw_0, \quad (7.4)$$

where $g(\cdot)$ is bounded and continuous and $w_0(\cdot)$ is a standard Wiener process which is independent of $x(\cdot)$. We will develop a computational approximation to this filtering problem. Condition (A7.1) will be assumed. First, we will define the optimal filter via the so-called representation theorem. Then the optimal nonlinear filter for an approximating process will be defined, and finally the numerical method will be given. The development is an extension and simplification of that in [76].

For $\phi(\cdot)$ an arbitrary bounded and continuous real valued function, define the conditional expectation operator E_t by

$$E_t \phi(x(t)) = E[\phi(x(t))|\mathcal{Y}(t)].$$

One of the most important results in the theory of nonlinear filtering is the so-called representation theorem, which is a “limit” form of Bayes rule. We will use it in the form in which it was originally derived by [73], which is particularly convenient for the types of approximation and weak convergence methods which will be used. That reference derived the result for the diffusion case. The jump process case was first derived by [108] and [126]. More modern developments via measure transformation and martingale techniques are in [41, 93], but the representations obtained by those techniques are the same as used here for the processes of concern here.

Let $\tilde{x}(\cdot)$ be a process with the same probability law as $x(\cdot)$ has, but which is independent of $(x(\cdot), y(\cdot))$. Define

$$R(t) = \exp \left[\int_0^t g(\tilde{x}(s))' dy(s) - \frac{1}{2} \int_0^t |g(\tilde{x}(s))|^2 ds \right].$$

Then we have the representation [73]

$$E_t \phi(x(t)) = \frac{E[R(t)\phi(\tilde{x}(t))|\mathcal{Y}(t)]}{E[R(t)|\mathcal{Y}(t)]}. \quad (7.5)$$

Except for a few special cases, the evaluation of (7.5) is not a finite calculation. The best known cases where the calculation is (essentially) finite are (i) the Kalman-Bucy filter, where the process is not reflected or killed, $g(\cdot)$ and $b(\cdot)$ are linear functions of x , $q(x, \gamma) \equiv 0$, and $\sigma(\cdot)$ does not depend on x and (ii) $x(\cdot)$ is a finite state Markov chain. Numerical methods for the approximate evaluation of the conditional distribution were given

in [76] and related variations were in [33] and [78]. Robustness (locally Lipschitz continuity) of the numerical approximations with respect to the observation process was shown in [78] and this property is enjoyed by the algorithm to be described in Subsection 12.7.3. Recently, [4] described the development of a special purpose computer chip for the solution of two dimensional problems via the Markov chain approximation method.

It is important to keep in mind that all approximations to the nonlinear filtering problem are actually approximations to some representation of Bayes rule. Other approximations to Bayes rule for the problem at hand might lead to preferable procedures in particular cases. The procedures to be described below are of the same type as in these references, but can use more general approximating processes. The basic idea depends on two facts. First, if the signal $x(\cdot)$ were a finite state discrete time Markov chain, then the conditional distribution can be obtained in principle by a finite calculation. Then, if the Markov chain is a locally consistent approximation to $x(\cdot)$ and the actual physical data from (7.4) is used, then the weak convergence results from previous sections can be used to show the convergence of the approximation to the true filter.

We will next state the analogue of (7.5) for a finite state Markov chain. Then we use (A7.1) to approximate the process $x(\cdot)$ by a locally consistent Markov chain such as that defined in Section 12.2. The approximating filter to be defined in Subsection 12.7.3 will be that for the approximating chain, but will use the actual physical observations (7.4). It will turn out that the resulting filter will converge to the values given by (7.5) as h and δ go to zero. Either the explicit or the implicit approximations can be used.

The Optimal Filter for a Markov Chain Signal Process. Let $\{\xi_n, n < \infty\}$ be a finite state Markov chain with one step transition probabilities $p(x, y)$. Let v be a positive real number and suppose that $\{\psi_n, n < \infty\}$ is a sequence of mutually independent normally distributed random variables with mean zero and covariance vI , and which is also independent of the $\{\xi_n, n < \infty\}$. Suppose that we observe the white noise corrupted data $y_n = \tilde{g}(\xi_n) + \psi_n$ at time step n , for some bounded function $\tilde{g}(\cdot)$. Define $\mathcal{Y}_n = \{y_i, i \leq n\}$ and the conditional distribution

$$\tilde{Q}_n(x) = P\{\xi_n = x | \mathcal{Y}_n\}.$$

We now use Bayes rule to define a convenient recursive formula for $\tilde{Q}_n(x)$. Let the expression $P\{y_n | \xi_n = x\} = P\{y_n | \xi_n = x, \xi_{n-1} = y\}$ denote the conditional (normal with mean $\tilde{g}(x)$ and covariance vI) density of the observation at the value y_n . Note that

$$\begin{aligned}\tilde{Q}_n(x) &= \sum_y P\{\xi_n = x | \xi_{n-1} = y, y_n\} \tilde{Q}_{n-1}(y) \\ &= \frac{\sum_y P\{y_n | \xi_n = x, \xi_{n-1} = y\} P\{\xi_n = x | \xi_{n-1} = y\} \tilde{Q}_{n-1}(y)}{\text{normalization}}.\end{aligned}$$

Substituting in the normal conditional density function of the observation y_n , we can rewrite the last expression as

$$\frac{\sum_y \exp \left[\frac{1}{v} \tilde{g}(x)' y_n - \frac{1}{2v} |\tilde{g}(x)|^2 \right] p(y, x) \tilde{Q}_{n-1}(y)}{\text{normalization}}, \quad (7.6)$$

where, in both cases, the normalization is just the numerator summed over x .

The Optimal Filter for the “Explicit” Chain $\{\xi_n^{h,\delta}, n < \infty\}$ of Sections 12.1 and 12.2. Let us specialize the result (7.6) to the chain introduced in Sections 12.1 and 12.2. Thus, ξ_n is replaced by $\xi_n^{h,\delta}$, and the transition probability $p(x, y)$ is replaced by $p^{h,\delta}(x, y)$. Let the observation at the n^{th} time step be

$$\Delta y_n^{h,\delta} = g(\xi_n^{h,\delta})\delta + [w_0(n\delta + \delta) - w_0(n\delta)],$$

where $g(\cdot)$ and $w_0(\cdot)$ are as in (7.4). Thus, $v = \delta$ and $\tilde{g}(x) = g(x)\delta$. Define $\mathcal{Y}_n^{h,\delta} = \{\Delta y_i^{h,\delta}, i \leq n\}$, and

$$\tilde{Q}_n^{h,\delta}(x) = P\{\xi_n^{h,\delta} = x | \mathcal{Y}_n^{h,\delta}\}.$$

Define the expression

$$R^{h,\delta}(x, \Delta y_n^{h,\delta}) = \exp \left[g(x)' \Delta y_n^{h,\delta} - \frac{1}{2} |g(x)|^2 \delta \right].$$

Then we can write (7.6) as

$$\tilde{Q}_n^{h,\delta}(x) = \frac{\sum_y R^{h,\delta}(x, \Delta y_n^{h,\delta}) p^{h,\delta}(y, x) \tilde{Q}_{n-1}^{h,\delta}(y)}{\text{normalization}}, \quad (7.7)$$

where the normalization is the numerator summed over x , $\tilde{Q}_0^{h,\delta}(x)$ is the a priori probability that $\xi_0^{h,\delta} = x$, and is a weak sense approximation to the weak sense density of $x(0)$. One can also write (7.7) in the unnormalized form

$$\tilde{q}_n^{h,\delta}(x) = \sum_y R^{h,\delta}(x, \Delta y_n^{h,\delta}) p^{h,\delta}(y, x) \tilde{q}_{n-1}^{h,\delta}(y), \quad (7.8)$$

where $\tilde{q}_0^{h,\delta}(x) = \tilde{Q}_0^{h,\delta}(x)$ and $\tilde{q}_m^{h,\delta}(x)$ equals $\tilde{Q}_m^{h,\delta}(x)$ times a normalizing factor which depends only on the data $\mathcal{Y}_m^{h,\delta}$.

A Representation of the Form of (7.5). For the purposes of showing convergence, it is useful to put (7.7) and (7.8) into the “functional” form (7.5). Let $\{\tilde{\xi}_n^{h,\delta}, n < \infty\}$ be a Markov chain with the same transition

probabilities $p^{h,\delta}(x, y)$ and which is independent of $\{\xi_n^{h,\delta}, \Delta y_n^{h,\delta}, n < \infty\}$. Define

$$\tilde{R}_n^{h,\delta} = \exp \left[\sum_{i=0}^n g(\tilde{\xi}_i^{h,\delta})' \Delta y_i^{h,\delta} - \frac{1}{2} \sum_{i=0}^n |g(\tilde{\xi}_i^{h,\delta})|^2 \delta \right].$$

Then (7.7) can be written as

$$\begin{aligned} \tilde{Q}_n^{h,\delta}(x) &= \frac{E[I_{\{\tilde{\xi}_n^{h,\delta}=x\}} \tilde{R}_n^{h,\delta} | \mathcal{Y}_n^{h,\delta}]}{E[\tilde{R}_n^{h,\delta} | \mathcal{Y}_n^{h,\delta}]} \\ &= \frac{\tilde{q}_n^{h,\delta}(x)}{\sum_y \tilde{q}_n^{h,\delta}(y)}, \end{aligned} \quad (7.9)$$

with the obvious unnormalized version.

The Filter for the “Implicit” Method. Recall the definition of the chain $\{\tilde{\zeta}_n^{h,\delta}\}$ and its one step transition probabilities $\tilde{p}^{h,\delta}(x, y)$ defined at the end of Section 12.4. To get the filter for the implicit method simply replace the $p(x, y)$ in (7.7), (7.8), or (7.9) by $\tilde{p}^{h,\delta}(x, y)$.

12.7.3 THE APPROXIMATION TO THE OPTIMAL FILTER FOR $x(\cdot), y(\cdot)$

The numerical approximation to the optimal filter (7.5) is just either (7.7), (7.8), or (7.9) with the *actual physical observations* $y(n\delta + \delta) - y(n\delta)$ used in place of $\Delta y_n^{h,\delta}$. Both (7.7) and (7.8) provide recursive formulas which can be used for the actual computation, and one representation for the recursion is in [76, pp.132-133]. The initial condition $\tilde{Q}_0^{h,\delta}(\cdot)$ is any approximation to the a priori weak sense density of $x(0)$ and which converges weakly to that density as h and δ go to zero. As noted above, these equations can be used for the implicit method also.

The Convergence Proof. Consider the representations (7.7), (7.8) or (7.9). For $t = n\delta$, define

$$R^{h,\delta}(t) = \exp \left[\sum_{i=0}^n g(\tilde{\xi}_i^{h,\delta})' (y(n\delta + \delta) - y(n\delta)) - \frac{1}{2} \sum_{i=0}^n |g(\tilde{\xi}_i^{h,\delta})|^2 \delta \right].$$

Let $\xi^{h,\delta}(\cdot)$ be the continuous parameter interpolation defined in Section 12.2. Let $\phi(\cdot)$ be a continuous and bounded real valued function. The value that the numerical approximation (7.7) or (7.9) gives for the estimate of the conditional expectation $E_t \phi(x(t))$ is the right hand side of

$$E_t \phi(x(t)) \approx \frac{E[\phi(\tilde{\xi}^{h,\delta}(t)) R^{h,\delta}(t) | \mathcal{Y}(t)]}{E[R^{h,\delta}(t) | \mathcal{Y}(t)]}. \quad (7.10)$$

It should be apparent from the forms of (7.10) and (7.5) and a weak convergence argument that (7.10) converges to (7.5) as h and δ converge to zero, provided that the consistency conditions of Section 12.2 hold.

The proof for the “implicit” method is the same.

13

Problems from the Calculus of Variations

A large class of deterministic optimal control problems are special cases of the stochastic optimal control problems considered previously. This is true both with respect to the construction of schemes as well as the proofs of convergence. In fact, the convergence proofs become much simpler in the deterministic setting.

In the present chapter we will consider deterministic optimal control problems which are not special cases of the stochastic control problems dealt with previously. We have two goals in mind. One goal is to show the flexibility of the Markov chain approximation methods with regard to weakening assumptions. The second goal is to discuss several classes of problems that are of current interest but not covered by the results given so far. To be precise, we will consider the construction of numerical schemes and the proofs of their convergence for problems from the calculus of variations. Such problems arise in a wide variety of settings. Well known examples are classical mechanics and geometric optics (e.g., [25, 51]). A more recent example is the theory of large deviations of stochastic processes [48]. In many ways these problems are simpler than most of the problems treated previously. There are, however, some interesting new features that must be dealt with. For example, when we rewrite a calculus of variations problem as a control problem, the space of controls is usually unbounded. It is also frequently the case that we must deal with costs that are discontinuous. Other difficulties arise in the infinite time problems due to the fact that the running cost often is not bounded from below away from zero. Such complicating features might also occur in the more standard stochastic and deterministic control problems. The interested reader can combine the methods used in the present chapter with those introduced previously to treat some of these generalizations.

An outline of the chapter is as follows. In Section 13.1 we will describe two basic classes of problems together with some variations. The first class is the calculus of variations analogue of the control problem treated in Chapter 12. We consider a fixed time interval and consider both a running cost and a terminal cost. This type of cost criteria corresponds to what is called a *Bolza problem* in the calculus of variations. In the second class, we also allow the *time* at which the process is stopped to be a control variable.

In Sections 13.2 and 13.3 we describe the numerical schemes and give the proofs of convergence for the finite time problem and the problem with controlled stopping time, respectively. Of particular interest is the treatment of the case where the running cost is not bounded from below away from zero. Also included in Section 13.3 is an application. We describe a calculus of variations formulation of a problem from the class of problems known as *shape from shading* and discuss some aspects of the numerical implementation. Section 13.4 considers a problem in which the running cost is discontinuous in the state variable. In the case of stochastic control problems, such discontinuities may essentially be ignored (with respect to the convergence of the numerical schemes) if the optimally controlled process induces a measure on path space under which the total cost is continuous w.p.1 (see the remark after Theorem 10.1.1). Here the situation is quite different because the underlying processes are deterministic. In this case, properly dealing with the discontinuity becomes the main focus of interest. We will see that the Markov chain method continues to work well and, in fact, provides a very natural way to deal with a rather difficult deterministic problem.

13.1 Problems of Interest

Let $k : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$ denote a running cost. We will assume throughout this chapter that $k(\cdot, \cdot)$ satisfies the following uniform superlinear growth condition:

$$\liminf_{c \rightarrow \infty} \inf_x \inf_{\alpha: |\alpha|=c} k(x, \alpha) / c = +\infty. \quad (1.1)$$

This condition is natural and holds in most applications. Under this condition there exists a convex function $l : [0, \infty) \rightarrow (-\infty, +\infty]$ which is bounded from below and satisfies

$$k(x, \alpha) \geq l(|\alpha|) \text{ for all } (x, \alpha), \text{ and } \lim_{c \rightarrow \infty} l(c)/c \rightarrow \infty. \quad (1.2)$$

In previous chapters, we have always assumed that the control space is compact. Thus, tightness of the sequence of relaxed controls appearing in the numerical schemes was automatic. This will not be the case in this chapter. The condition (1.1) will be used in lieu of the compactness to force tightness of relaxed controls, at least for the sequence of optimal relaxed controls associated with the approximations to the value function. The function $g(\cdot)$ will be a stopping cost, and we will always assume $g(\cdot)$ is bounded from below.

We next give two representative problem classes.

13.1.1 FINITE TIME PROBLEMS

Fix $T > 0$. In the finite time problem, we seek

$$V(x) = \inf \left[\int_0^T k(\phi(s), \dot{\phi}(s)) ds + g(\phi(T)) \right], \quad (1.3)$$

where the infimum is over all absolutely continuous functions $\phi : [0, T] \rightarrow \mathbb{R}^k$ satisfying $\phi(0) = x$. For notational convenience we rewrite this problem as an optimal control problem. The set of admissible controls for this problem will consist of all measurable functions from $[0, T]$ to \mathbb{R}^k . Let $u(\cdot)$ be any admissible control. The dynamics of the controlled process are then given simply by $\dot{x}(t) = u(t)$, $x(0) = x$, and the cost to be minimized is

$$W(x, u) = \int_0^T k(x(s), u(s)) ds + g(x(T)).$$

More generally, we could consider the problem of approximating

$$V(x, t) = \inf \left[\int_t^T k(\phi(s), \dot{\phi}(s)) ds + g(\phi(T)) \right],$$

where the infimum is over all absolutely continuous functions $\phi : [t, T] \rightarrow \mathbb{R}^k$ satisfying $\phi(t) = x$. This is the calculus of variations analogue of the problems treated in Chapter 12. To simplify the notation, we consider only $V(x)$, but note that the approximation schemes that are derived below actually yield approximations to $V(x, t)$ for all $t \in [0, T]$.

Just as in Chapter 3 it is possible to formally derive a Bellman equation for the cost $V(x, t)$, at least for the case where $k(\cdot, \cdot)$ is continuous. For the problem under consideration, the differential operator takes the particularly simple form

$$\mathcal{L}^\alpha f(x) = f'_x(x)\alpha,$$

where the control variable α takes values in \mathbb{R}^k . The formal Bellman equation is then given by

$$\begin{cases} V_t(x, t) + \inf_{\alpha \in \mathbb{R}^k} [\mathcal{L}^\alpha V(x, t) + k(x, \alpha)] = 0, \\ V(x, T) = g(x). \end{cases}$$

It is sometimes the case that the functions $k(\cdot)$ and $g(\cdot)$ satisfy continuity conditions. However, there are many interesting applications where this is not the case. For example, it may be the case that the path $\phi(\cdot)$ is required to stay in some closed set G over the interval $[0, T]$. This can be incorporated into the problem by defining $k(x, \alpha)$ to be $+\infty$ for $x \notin G$. It is also possible that constraints are placed on $\phi(\cdot)$ or on the location of $\phi(T)$. These can also be incorporated into the problem as given above by suitably

redefining $k(\cdot)$ and $g(\cdot)$. In general, these control and state space constraints can be readily dealt with (when constructing the numerical method and proving its convergence). See, for example, the convergence theorems of Sections 13.2 and 13.3. A case with a discontinuity that is more difficult to deal with appears in Section 13.4.

For numerical purposes we may require that the state space be bounded. One method for bounding that produces an algorithm which is simple to program is to simply stop $\phi(\cdot)$ at the first time τ that it leaves the interior of a suitable set, such as

$$G = \{x : c_i \leq x_i \leq d_i \text{ for } i = 1, 2, \dots, k\},$$

at which time a stopping cost $g(\phi(\tau), \tau)$ may be assessed. This would add a Dirichlet boundary condition to the Bellman equation above. For notational convenience we will combine the terminal and stopping costs as one function $g(x, t)$, and for later purposes we will assume that $g(\cdot, \cdot)$ is defined on all of $\mathbb{R}^k \times [0, T]$. The calculus of variations problem then becomes

$$V(x) = \inf \left[\int_0^{T \wedge \tau} k(\phi(s), \dot{\phi}(s)) ds + g(\phi(T \wedge \tau), T \wedge \tau) \right]. \quad (1.4)$$

Alternatively, one can use the analogue of the “reflecting” boundary condition for diffusions that was introduced in Section 1.4.

Fix a particular $x \in G$. For either of these methods of bounding the state space, a condition that is sufficient to guarantee that the value for the original problem (1.3) and the value of the modified problem (1.4) are the same is that the minimizing trajectories for the two problems be the same and remain in G^0 . It is often the case that properties of the functions $k(\cdot, \cdot)$ [e.g. the lower bound $l(\cdot)$] and $g(\cdot)$ can be exploited to obtain a bound on the range

$$R = \{\phi(t) : 0 \leq t \leq T\},$$

where $\phi(\cdot)$ starts at x at time 0 and is a minimizing trajectory for the problem with no boundary. In such a case, G^0 should be chosen to contain R . By imposing a suitably large stopping cost on the set $\{(x, t) : x \in \partial G, t \in [0, T]\}$, it can be assured that the minimizing trajectories for the original and bounded problems are the same and remain in G^0 .

Owing to its practical importance with regard to numerical implementation, we will use the problem (1.4) as our canonical example of a finite time problem. In particular, the convergence of numerical schemes will be demonstrated for this problem. It is worth noting that under appropriate additional conditions, problems with a time dependent running cost can be handled by the methods developed below.

13.1.2 PROBLEMS WITH A CONTROLLED STOPPING TIME

In this case the problem of interest is to again find

$$V(x) = \inf \left[\int_0^\rho k(\phi(s), \dot{\phi}(s)) ds + g(\phi(\rho)) \right], \quad (1.5)$$

where now the infimum is over all $\rho \geq 0$ and absolutely continuous functions $\phi : [0, \rho] \rightarrow \mathbb{R}^k$ satisfying $\phi(0) = x$. In the same way as in the previous subsection, this problem may be rewritten as an optimal control problem. As with the finite time problem, it is sometimes the case that $k(\cdot)$ and $g(\cdot)$ are discontinuous.

Because the stopping time is directly controlled, the Bellman equation takes the form

$$\begin{cases} \inf_{\alpha \in \mathbb{R}^k} [\mathcal{L}^\alpha V(x) + k(x, \alpha)] = 0, & x \notin B, \\ V(x) = g(x), & x \in B, \quad V(x) \leq g(x), & x \notin B, \end{cases}$$

for some (a priori unknown) stopping set B . Remarks analogous to those for the finite time problem hold concerning methods for modifying (1.5) in order to bound the state space of the process. In particular, we will focus on the modification of (1.5) in which $\phi(\cdot)$ is stopped at the first time τ that it leaves the interior of a given compact set G . The cost now takes the form

$$V(x) = \inf \left[\int_0^{\rho \wedge \tau} k(\phi(s), \dot{\phi}(s)) ds + g(\phi(\rho \wedge \tau)) \right], \quad (1.6)$$

where the infimum is over the same set of paths as (1.5) and all $\rho \geq 0$. Note that $g(\cdot)$ combines the stopping cost and the cost that is added when the set G^0 is exited. Thus, $g(\cdot)$ will often be discontinuous on ∂G .

13.2 Numerical Schemes and Convergence for the Finite Time Problem

We will now set up the Markov chain approximation for the numerical method. The dynamical equation for the optimal control problem of Subsection 13.1.1 is simply $\dot{x}(t) = u(t)$, where the control process $u(\cdot)$ takes values in \mathbb{R}^k . Suppose we fix a grid $S_h \subset \mathbb{R}^k$. Since the dynamics are independent of the state, we will assume $\Delta t^h(x, \alpha)$ and the distribution of $y - x$ under $p^h(x, y | \alpha)$ are independent of x . State dependency of these quantities can be put back in with little effort if it is desired. The local consistency conditions (4.1.3) thus become

$$E_{x,n}^{h,\alpha} \Delta \xi_n^h = \alpha \Delta t^h(\alpha) + o(|\alpha| \Delta t^h(\alpha)), \quad (2.1)$$

$$E_{x,n}^{h,\alpha} [\Delta \xi_n^h - E_{x,n}^{h,\alpha} \Delta \xi_n^h] [\Delta \xi_n^h - E_{x,n}^{h,\alpha} \Delta \xi_n^h]' = o(|\alpha| \Delta t^h(\alpha)), \quad (2.2)$$

where $\{\xi_n^h, n < \infty\}$ is a controlled Markov chain taking values in S_h . Due to the unbounded nature of the control space care must be taken regarding how the “errors” in the local consistency equations depend on α . The precise meaning of (2.1) is that

$$(E_{x,n}^{h,\alpha} \Delta \xi_n^h - \alpha \Delta t^h(\alpha)) / (|\alpha| \Delta t^h(\alpha)) \rightarrow 0$$

uniformly for $\alpha \in \mathbb{R}^k$, and similarly for (2.2). We will also assume throughout the rest of this chapter that $\Delta t^h(\alpha) \rightarrow 0$ for each $\alpha \in \mathbb{R}^k$ and that

$$\lim_{h \rightarrow 0} \sup \{|y - x| : p^h(x, y | \alpha) > 0, x \in S_h, y \in S_h, \alpha \in \mathbb{R}^k\} = 0.$$

Example 2.1. Suppose that we use the grid

$$S_h = \left\{ x : x = h \sum_{i=1}^k e_i n_i, n_i = 0, \pm 1, \pm 2, \dots \right\}.$$

Then an obvious choice for the transition function is

$$p^h(x, y | \alpha) = \begin{cases} \alpha^\pm / \sum_j |\alpha_j| & \text{if } y = x \pm h e_i \\ 0 & \text{otherwise} \end{cases}$$

and $\Delta t^h(\alpha) = h(\sum_j |\alpha_j|)^{-1}$. This definition actually makes sense only when $\alpha \neq 0$, and we take care of the omitted case by setting

$$p^h(x, y | 0) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise.} \end{cases}$$

The sequence $\Delta t^h(0) > 0$ is arbitrary [for the purposes of satisfying (2.1) and (2.2)] and for simplicity we can take $\Delta t^h(0) = h$. ■

We will require for the finite time problem that $\Delta t^h(\alpha)$ be independent of the control. Following the development of Chapter 12, we next derive explicit and implicit schemes from transition probabilities satisfying the local consistency conditions (2.1) and (2.2).

13.2.1 DESCRIPTIONS OF THE NUMERICAL SCHEMES

Explicit Schemes. Because $u(\cdot)$ may be unbounded, care must be taken with respect to the relationship between the time discretization and the spatial discretization.

With the definitions [analogous to (12.2.3) and (12.2.1)]

$$\tilde{p}^{h,\delta}(x, x | \alpha) = 1 - \frac{\delta}{\Delta t^h(\alpha)}, \quad (2.3)$$

$$\tilde{p}^{h,\delta}(x, y|\alpha) = p^h(x, y|\alpha) (1 - \tilde{p}^{h,\delta}(x, x|\alpha)), \quad (2.4)$$

we find that $\tilde{p}^{h,\delta}(x, y|\alpha)$ defines a locally consistent chain in the sense of (2.1) and (2.2) [but with δ replacing $\Delta t^h(\alpha)$ there] if we restrict α to the set

$$\mathcal{U}^{h,\delta} = \left\{ \alpha : \delta/\Delta t^h(\alpha) \leq 1 \right\}. \quad (2.5)$$

If we assume $\Delta t^h(0) \geq \delta$, then for the case of Example 2.1 we have

$$\mathcal{U}^{h,\delta} = \left\{ \alpha : \sum_{j=1}^k |\alpha_j| \leq h/\delta \right\}.$$

It may be that bounds on the optimal control for the original calculus of variations problem are available. For example, it may be that one can calculate B such that

$$\text{ess sup}_{t \in [0, T]} \sum_j |\dot{\phi}_j|(t) \leq B,$$

where ess sup stands for essential supremum and $\phi(\cdot)$ is a minimizing path. Then we must assume conditions which guarantee that the restricted control space $\mathcal{U}^{h,\delta}$ is eventually “big enough.” For example, in the case of Example 2.1 we must assume that the pair (h, δ) are sent to zero in such a way that

$$\liminf_{h, \delta} h/\delta \geq B.$$

If such a bound is not available, we must assume

$$\liminf_{h, \delta} \mathcal{U}^{h,\delta} = \mathbb{R}^k, \quad (2.6)$$

i.e., h and δ must be sent to zero in such a way that given any $u \in \mathbb{R}^k$ we have $u \in \mathcal{U}^{h,\delta}$ for all sufficiently small $h > 0$ and $\delta > 0$.

Now let $G_h = G \cap S_h$ and $G_h^0 = G^0 \cap S_h$. Then the explicit approximation scheme for solving (1.4) is

$$V^{h,\delta}(x, n\delta) = \min_{\alpha \in \mathcal{U}^{h,\delta}} \left[\sum_y \tilde{p}^{h,\delta}(x, y|\alpha) V^{h,\delta}(y, n\delta + \delta) + k(x, \alpha)\delta \right] \quad (2.7)$$

for $x \in G_h^0$ and $n\delta < T$, together with the boundary and terminal condition $V^{h,\delta}(x, n\delta) = g(x, n\delta)$ for $x \notin G_h^0$ and $n\delta \leq T$ or $x \in G_h^0$ and $n\delta = T$.

Implicit Schemes. Following Section 12.4, we define a transition function on the (x, t) -grid $S_h \times \{0, \delta, 2\delta, \dots\}$ by

$$\hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha) = \frac{\Delta t^h(\alpha)}{\Delta t^h(\alpha) + \delta}, \quad (2.8)$$

$$\hat{p}^{h,\delta}(x, n\delta; y, n\delta|\alpha) = p^h(x, y|\alpha) (1 - \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)), \quad (2.9)$$

$$\Delta\hat{t}^{h,\delta}(\alpha) = \frac{\delta\Delta t^h(\alpha)}{\Delta t^h(\alpha) + \delta}, \quad (2.10)$$

where $p^h(x, y|\alpha)$ and $\Delta t^h(\alpha)$ satisfy the local consistency conditions (2.1) and (2.2). We also retain from Chapter 12 the notation $\hat{\zeta}_n^{h,\delta}$ for the controlled Markov chain and $(\zeta_n^{h,\delta}, \hat{\zeta}_n^{h,\delta})$ to denote separately the “spatial” and “temporal” components.

Define $N^{h,\delta}(T) = \min\{n : \zeta_{n,0}^{h,\delta} \geq T\}$. Then the implicit scheme for solving (1.4) is

$$\begin{aligned} V^{h,\delta}(x, n\delta) &= \min_{\alpha \in \mathbb{R}^k} \left[\sum_y \hat{p}^{h,\delta}(x, n\delta; y, n\delta|\alpha) V^{h,\delta}(y, n\delta) \right. \\ &\quad \left. + \hat{p}^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha) V^{h,\delta}(x, n\delta + \delta) + k(x, \alpha) \Delta\hat{t}^{h,\delta}(\alpha) \right] \end{aligned} \quad (2.11)$$

for $x \in G_h^0$ and $n\delta < T$, and with the same boundary condition as for (2.7). Note that for the implicit scheme we need not require that h and δ satisfy any special relationship as they tend to zero.

13.2.2 APPROXIMATIONS AND PROPERTIES OF THE VALUE FUNCTION

In this subsection we state some properties of the controls and the value function that are needed in the convergence proofs. Let $B_a(x) = \{y : |x - y| < a\}$. We will use the following assumptions.

A2.1. *The set G is compact and satisfies interior and exterior cone conditions: There exist $\epsilon > 0$ and continuous functions $v(\cdot)$ and $w(\cdot)$ such that given any $x \in \partial G$, $\cup_{0 < a < \epsilon} B_{\epsilon a}(x + av(x)) \subset G$ and $\cup_{0 < a < \epsilon} B_{\epsilon a}(x + aw(x)) \cap G = \emptyset$.*

A2.2. *The function $k(\cdot, \cdot)$ is continuous and satisfies the superlinear growth condition (1.1). There exist $M < \infty$ and $f : \mathbb{R} \rightarrow [0, \infty)$ satisfying $f(t) \rightarrow 0$ as $t \rightarrow 0$, such that for all x and y ,*

$$\sup_{\alpha} [|k(x, \alpha) - k(y, \alpha)| - f(|x - y|)(M + k(x, \alpha))] \leq 0.$$

A2.3. *The function $g(\cdot, \cdot)$ is uniformly continuous and bounded when restricted to either of the sets $(\mathbb{R}^k - G^0) \times [0, T]$ and $G^0 \times \{T\}$.*

Remarks. The condition (A2.2) occurs often in calculus of variations problems of the type we consider. See, for example, [48, Chapter 5]. The continuity assumption in (A2.2) is intermediate between continuity in x that is uniform in α (which is much too restrictive) and simple continuity. Note

that (A2.3) does not assume continuity of $g(\cdot, \cdot)$. We have chosen this form for $g(\cdot, \cdot)$ because it is common in applications. Owing to the “controllability” of the dynamics $\dot{\phi}(t) = u(t)$, discontinuities in $g(\cdot, \cdot)$ are generally not too difficult to deal with. In particular, it follows from (A2.1) and (A2.2) that the infimum in (1.4) is the same if $g(\cdot, \cdot)$ is replaced by $g_*(\cdot, \cdot)$, where $g_*(x, t) = g(x, t)$ for $t < T$ and $g_*(x, T) = \lim_{\epsilon \rightarrow 0} \inf\{g(y, T) : |y - x| \leq \epsilon\}$.

It is sometimes desirable to weaken (A2.1). For a simple example, consider the case of control until a target set is reached when the target set is a single point. For obvious reasons, such target sets are not typically considered in stochastic control problems. However, they do appear often in deterministic control. Clearly, the interior cone condition is satisfied at the point, but the exterior cone condition is not. In the example of Subsection 13.3.3 we show how to extend the convergence proofs to cover such cases.

Definition. The set of *admissible relaxed controls* for this class of deterministic problems consists of all measures $m(\cdot)$ on the Borel subsets of $\mathbb{R}^k \times [0, \infty)$ which satisfy $m(\mathbb{R}^k \times [0, t]) = t$ for all $t \in [0, \infty)$.

For notational convenience we continue the convention of assuming controls are defined for all time, even if they are only applied over a finite interval.

Theorem 2.2. *Assume the condition (A2.2). Suppose that a relaxed control $m(d\alpha ds) = m_s(d\alpha)ds$ is given such that for*

$$x(t) = x + \int_0^t \int_{\mathbb{R}^k} \alpha m_s(d\alpha) ds,$$

we have

$$\int_0^T \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds < \infty.$$

Let $\epsilon > 0$ be given. Then there exists $\bar{\delta} > 0$ and a finite set $\{\alpha_1^\epsilon, \dots, \alpha_{p_\epsilon}^\epsilon\} = \mathcal{U}_\epsilon \subset \mathbb{R}^k$ with the following properties. There is a function $u^\epsilon : [0, T] \rightarrow \mathcal{U}_\epsilon$ which is constant on intervals of the form $[i\bar{\delta}, i\bar{\delta} + \bar{\delta})$, and such that if

$$x^\epsilon(t) = x + \int_0^t u^\epsilon(s) ds,$$

then

$$\sup_{0 \leq t \leq T} |x^\epsilon(t) - x(t)| \leq \epsilon$$

and

$$\sup_{0 \leq t \leq T} \left| \int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds - \int_0^t k(x^\epsilon(s), u^\epsilon(s)) ds \right| \leq \epsilon.$$

Proof. The theorem is essentially a simpler version of Theorem 10.1.2. The main new difficulty in the current setup is the unboundedness of the control space. We will prove that given $\epsilon > 0$ there is an admissible relaxed control $m^\epsilon(\cdot)$ and a compact set K^ϵ such that $m^\epsilon(\cdot)$ is supported on $K^\epsilon \times [0, T]$, and such that if $x^\epsilon(\cdot)$ is the associated solution, then

$$\sup_{0 \leq t \leq T} \left| \int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds - \int_0^t \int_{\mathbb{R}^k} k(x^\epsilon(s), \alpha) m_s^\epsilon(d\alpha) ds \right| \leq \epsilon$$

and

$$\sup_{0 \leq t \leq T} |x^\epsilon(t) - x(t)| \leq \epsilon.$$

Given the existence of such a control, the proof of existence of a control $u^\epsilon(\cdot)$ satisfying all the conditions of the theorem follows from Theorem 10.1.2.

The lower bound (1.2) implies $k(x, \alpha)$ is non-negative whenever $|\alpha|$ is sufficiently large. By the dominated convergence theorem and (1.2), there exists $c(\epsilon) < \infty$ such that

$$\int_0^T \int_{\mathbb{R}^k} I_{\{|\alpha| \geq c(\epsilon)\}} |k(x(s), \alpha)| m(d\alpha ds) \leq \epsilon/2 \quad (2.12)$$

and

$$\int_0^T \int_{\mathbb{R}^k} I_{\{|\alpha| \geq c(\epsilon)\}} |\alpha| m(d\alpha ds) \leq \epsilon.$$

This last inequality implies

$$\sup_{0 \leq t \leq T} \left| \int_0^t \int_{\mathbb{R}^k} I_{\{|\alpha| \geq c(\epsilon)\}} \alpha m(d\alpha ds) \right| \leq \epsilon. \quad (2.13)$$

We can also assume $c(\epsilon)$ is sufficiently large so that

$$\left(\sup_{x \in G} |k(x, 0)| \right) \left(\int_0^T \int_{\mathbb{R}^k} I_{\{|\alpha| \geq c(\epsilon)\}} m(d\alpha ds) \right) \leq \epsilon/2. \quad (2.14)$$

The desired $m^\epsilon(\cdot)$ can then be obtained by defining

$$m_s^\epsilon(B) = m_s(B \cap \{\alpha : |\alpha| < c(\epsilon)\}) + m_s(\{\alpha : |\alpha| \geq c(\epsilon)\}) I_B(0)$$

for all Borel sets B and $s \geq 0$. For this choice of $m^\epsilon(\cdot)$ we have

$$\sup_{0 \leq t \leq T} |x^\epsilon(t) - x(t)| \leq \epsilon$$

because of (2.13). Therefore, by (A2.2), (2.12), and (2.14),

$$\begin{aligned} \sup_{0 \leq t \leq T} \left| \int_0^t \int_{\mathbb{R}^k} k(x^\epsilon(s), \alpha) m_s^\epsilon(d\alpha) ds - \int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds \right| \\ \leq \epsilon + f(\epsilon) \left(MT + \int_0^T \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds \right). \end{aligned}$$

As was observed in Chapters 9 and 10, for problems with boundary conditions one must pay particular attention to the manner in which controlled trajectories reach the boundary. A previously discussed method for proving the convergence of costs involved showing the continuity of the first exit time of ϕ when ϕ was a sample path of an underlying process (at least w.p.1). See the discussion in Section 10.2. Here, the deterministic nature of the problem makes this approach not applicable. Instead of requiring that the exit time be continuous for all paths x that solve $x(t) = x + \int_0^t u(s)ds$ for arbitrary admissible controls (which is impossible), we shall require only the existence of an ϵ -optimal control such that the exit time is continuous at the associated solution. This condition will always hold under our assumptions [except when we weaken (A2.1) in Section 3.3] and turns out to be sufficient for the proof of the upper bound $\limsup_{h \rightarrow 0} V^h(x) \leq V(x)$. To handle the lower bound $\liminf_{h \rightarrow 0} V^h(x) \geq V(x)$, we will need the following result.

Theorem 2.3. *Assume the conditions (A2.1) and (A2.2). Let $m(\cdot)$ be any admissible relaxed control, $x \in G^0$, and let $x(t) = x + \int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha)m(d\alpha ds)$. Fix any $\sigma < \infty$, and assume that $x(t) \in G$ for $t \in [0, \sigma]$, and that*

$$\int_0^\sigma \int_{\mathbb{R}^k} k(x(s), \alpha)m(d\alpha ds) < \infty. \quad (2.15)$$

Then given $\epsilon > 0$ there exists a control $m^\epsilon(\cdot)$ with associated solution $x^\epsilon(\cdot)$ such that

$$x^\epsilon(t) \in G^0, \text{ for } t \in [0, \sigma], \quad x^\epsilon(0) = x(0), \quad x^\epsilon(\sigma) = x(\sigma),$$

and

$$\left| \int_0^\sigma \int_{\mathbb{R}^k} k(x(s), \alpha)m(d\alpha ds) - \int_0^\sigma \int_{\mathbb{R}^k} k(x^\epsilon(s), \alpha)m^\epsilon(d\alpha ds) \right| < \epsilon.$$

Proof. The basic idea of the proof is to show that if the path $x(\cdot)$ is on ∂G at any time $t \in (0, \sigma)$, then the interior cone condition and the continuity properties of the running cost imply that by “pushing” the path by a small amount in the $v(x(t))$ direction, we obtain a path that is interior to G at time t and that has nearly the same running cost as $x(\cdot)$.

We now proceed to the proof. Assume the theorem is false. Then there exists a control $m(\cdot)$ with associated solution $x(\cdot)$ for which the conclusion is not true. Define A to be the set of all $t \in [0, \sigma)$ such that given any $\epsilon > 0$ there exists a control $m^*(\cdot)$ with associated solution $x^*(\cdot)$ with the properties that $x^*(s) \in G^0$ for all $s \in [0, t)$, $x^*(0) = x(0)$, $x^*(s) = x(s)$ and $m_s^*(\cdot) = m_s(\cdot)$ for $s \in [t, \sigma]$, and

$$\left| \int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha)m(d\alpha ds) - \int_0^t \int_{\mathbb{R}^k} k(x^*(s), \alpha)m^*(d\alpha ds) \right| < \epsilon.$$

Define σ^* to be the supremum of the set A . Our assumption concerning $m(\cdot)$ implies $\sigma^* < \sigma$, which we now show is false. For simplicity we will assume that σ^* is not only the supremum of A but also that $\sigma^* \in A$. The proof when $\sigma^* \notin A$ is essentially the same but notationally more complicated.

By (A2.1) and the fact that $x(t) \in G$ for $t \in [\sigma^*, \sigma]$, there is $\nu \in (0, \sigma - \sigma^*)$ such that

$$x(t) + \gamma v(x(\sigma^*)) \in G^0 \quad (2.16)$$

for all $t \in [\sigma^*, \sigma^* + \nu]$ and $\gamma \in (0, \nu]$.

Let $\epsilon > 0$ be given, and let $m^*(\cdot)$ be an admissible control such that $x^*(t) \in G^0$ for $t \in [0, \sigma^*]$, $x^*(0) = x(0)$, $x^*(t) = x(t)$ and $m_t^*(\cdot) = m_t(\cdot)$ for $t \in [\sigma^*, \sigma]$, and such that

$$\left| \int_0^{\sigma^*} \int_{\mathbb{R}^k} k(x(s), \alpha) m(d\alpha ds) - \int_0^{\sigma^*} \int_{\mathbb{R}^k} k(x^*(s), \alpha) m^*(d\alpha ds) \right| < \epsilon/2.$$

The existence of $m^*(\cdot)$ is guaranteed by the definition of σ^* . There is $\nu^* > 0$ [which can depend on $x^*(\cdot)$] such that

$$x^*(t) + \gamma v(x(\sigma^*)) \in G^0 \quad (2.17)$$

for all $t \in [\sigma^* - \nu^*, \sigma^*]$ and $\gamma \in (0, \nu^*]$. For $B \subset \mathbb{R}^k$ and $x \in \mathbb{R}^k$, let $B + x = \{y + x : y \in B\}$. For each $c > 0$ we define an admissible relaxed control $m^c(\cdot)$ by

$$m_s^c(B) = \begin{cases} m_s^*(B) & \text{for } s \in [0, \sigma^* - \nu^*], \\ m_s^*(B - cv(x(\tau^*))/\nu^*) & \text{for } s \in [\sigma^* - \nu^*, \sigma^*], \\ m_s(B + cv(x(\tau^*))/\nu) & \text{for } s \in [\sigma^*, \sigma^* + \nu], \\ m_s(B) & \text{for } s \in [\sigma^* + \nu, \sigma]. \end{cases}$$

Let $x^c(\cdot)$ be the associated solution. The effect of this control is to force $x^c(t) \in G^0$ for $t \in [\sigma^* - \nu^*, \sigma^* + \nu]$, for all sufficiently small $c > 0$. This follows from (2.16) and (2.17). We also have $x^c(t) = x(t)$ and $m_t^c(\cdot) = m_t(\cdot)$ for $t \in [\sigma^* + \nu, \sigma]$, and $x^c(t) = x^*(t)$ for $t \in [0, \sigma^* - \nu^*]$. Under (A2.2), we have

$$\left| \int_0^{\sigma^* + \nu} \int_{\mathbb{R}^k} k(x^c(s), \alpha) m^c(d\alpha ds) - \int_0^{\sigma^* + \nu} \int_{\mathbb{R}^k} k(x^*(s), \alpha) m^*(d\alpha ds) \right| \rightarrow 0$$

as $c \rightarrow 0$. Because $\epsilon > 0$ is arbitrary, we find that $\sigma^* + \nu \in A$, and thereby obtain a contradiction. ■

13.2.3 CONVERGENCE THEOREMS

We first prove a general convergence result for chains that satisfy the local consistency conditions (2.1) and (2.2) and also have uniformly bounded

running costs. The theorem is stated for chains that have a single discretization parameter h . However, the statement and proof are essentially the same for the two parameter families considered in Subsection 13.2.1 with $(h, \delta) \rightarrow 0$ replacing $h \rightarrow 0$. The result will be used in many places in the chapter and, consequently, is formulated in a general way. We first set up the notation of the theorem. Let $\{\xi_i^h, i < \infty\}$ be a controlled Markov chain. For each $h > 0$, assume that the transition probabilities $p^h(x, y|\alpha)$ and interpolation interval $\Delta t^h(\alpha)$ satisfy (2.1) and (2.2). Let $\{u_i^h, i < \infty\}$ be the sequence of controls applied to the chain and define

$$t_n^h = \sum_{i=0}^{n-1} \Delta t^h(u_i^h)$$

and the interpolated processes

$$\xi^h(t) = \xi_n^h, u^h(t) = u_n^h \text{ for } t \in [t_n^h, t_{n+1}^h).$$

As always, we assume the controls are admissible in the sense of Section 2.3. For $S < \infty$, let $S_h = \inf\{t_i^h : t_i^h \geq S\}$. Define the relaxed control $m^h(\cdot)$ by

$$m_s^h(A) = I_A(u^h(s)), \quad m^h(A \times [0, t]) = \int_0^t m_s^h(A) ds.$$

In the theorem statement, $c(h) \rightarrow 0$ is any sequence for which $c(h)|\alpha|\Delta t^h(\alpha)$ is a bound for the absolute value of the $o(|\alpha|\Delta t^h(\alpha))$ terms appearing in (2.1) and (2.2).

Theorem 2.4. *Assume that (1.1) holds, and that*

$$\limsup_{h \rightarrow 0} E^{m^h} \int_0^{S_h} \int_{\mathbb{R}^k} k(\xi^h(s), \alpha) m^h(d\alpha ds) < \infty$$

for each $S < \infty$. Then the set $\{m^h(\cdot), h > 0\}$ is tight. Assume also that the collection of initial conditions $\{\xi_0^h, h > 0\}$ is tight, and suppose that a subsequence (again indexed by h) is given such that $\{m^h(\cdot), h > 0\}$ converges weakly to a limit $m(\cdot)$ and ξ_0^h converges weakly to x . Then $m(\cdot)$ is an admissible relaxed control and $\{\xi^h(\cdot), h > 0\}$ converges weakly to a limit $x(\cdot)$ that satisfies

$$x(t) - x = \int_0^t \int_{\mathbb{R}^k} \alpha m(d\alpha ds)$$

w.p.1.

Suppose for each $h > 0$ that N_h is a stopping time for the chain $\{\xi_i^h, i < \infty\}$, and let $\tau_h = t_{N_h}^h$. If

$$\lim_{h \rightarrow 0} c(h) E^{m^h} \int_0^{\tau_h} \int_{\mathbb{R}^k} |\alpha| m^h(d\alpha ds) = 0,$$

then the convergence is uniform in the sense that

$$\sup_{0 \leq t \leq \tau_h} \left| \xi^h(t) - \xi^h(0) - \int_0^t \int_{\mathbb{R}^k} \alpha m^h(d\alpha ds) \right| \rightarrow 0$$

in probability.

Proof. Fix $S < \infty$. By using the lower bound (1.2) and the assumptions of the theorem, we obtain

$$\limsup_{h \rightarrow 0} E^{m^h} \left[\int_{\mathbb{R}^k \times [0, S]} l(|\alpha|) m^h(d\alpha ds) \right] < \infty, \quad (2.18)$$

and, by the superlinearity of $l(\cdot)$,

$$\limsup_{b \rightarrow \infty} \limsup_{h \rightarrow 0} E^{m^h} \left[\int_{\mathbb{R}^k \times [0, S]} I_{\{|\alpha| \geq b\}} |\alpha| m^h(d\alpha ds) \right] = 0. \quad (2.19)$$

Because the control space is \mathbb{R}^k , an arbitrary collection of relaxed controls need not have compact closure. However, we have the estimate (2.18) on the “tails” of these measures. A straightforward application of Prohorov’s theorem (Theorem 9.1.2) and Chebyshev’s inequality can then be used to show the sequence $\{m^h(\cdot), h > 0\}$ is tight in $\mathcal{R}([0, \infty) \times \mathbb{R}^k)$.

Now suppose that a subsequence again indexed by h is given such that $\{m^h(\cdot), h > 0\}$ converges to a limit $m(\cdot)$. Define the process $\{e_i^h, i < \infty\}$ by $e_0^h = 0$ and

$$e_{i+1}^h = e_i^h + E [\xi_{i+1}^h | u_i^h] - u_i^h \Delta t^h(u_i^h),$$

and the interpolation $e^h(t) = e_n^h$ for $t \in [t_n^h, t_{n+1}^h]$. The process $e^h(\cdot)$ keeps track of the “error” terms in (2.1). Owing to the definition of this process,

$$E [\xi_{i+1}^h - \xi_i^h - u_i^h \Delta t^h(u_i^h) - (e_{i+1}^h - e_i^h) | u_i^h] = 0.$$

Therefore, if we define

$$\gamma^h(t) = \xi^h(t) - \xi^h(0) - \int_0^t \int_{\mathbb{R}^k} \alpha m^h(d\alpha ds) - e^h(t), \quad (2.20)$$

then the process $\{\gamma^h(t_i^h), i < \infty\}$ is a martingale. A calculation using (2.2) gives

$$E^{m^h} |\gamma^h(S_h)|^2 \leq c(h) E^{m^h} \int_0^{S_h} \int_{\mathbb{R}^k} |\alpha| m^h(d\alpha ds),$$

where $c(h) \rightarrow 0$ as $h \rightarrow 0$. Thus, under the assumptions of the theorem,

$$E^{m^h} \left[\xi^h(S_h) - \xi^h(0) - \int_0^{S_h} \int_{\mathbb{R}^k} \alpha m^h(d\alpha ds) - e^h(S_h) \right]^2 \rightarrow 0$$

and, by the martingale inequality (1.1.2),

$$\sup_{0 \leq t \leq S} \left| \xi^h(t) - \xi^h(0) - \int_0^t \int_{\mathbb{R}^k} \alpha m^h(d\alpha ds) - e^h(t) \right| \rightarrow 0.$$

Using the Skorokhod representation we can assume that the convergence $m^h(\cdot) \rightarrow m(\cdot)$ is w.p.1 in the topology of weak convergence of measures on $\mathbb{R}^k \times [0, \infty)$. By (2.19) and the fact that $m(\mathbb{R}^k \times \{t\}) = 0$ for all $t \in [0, S]$, we obtain

$$\int_0^t \int_{\mathbb{R}^k} \alpha m^h(d\alpha ds) \rightarrow \int_0^t \int_{\mathbb{R}^k} \alpha m(d\alpha ds)$$

w.p.1. It follows from (2.1) that $\sup_{0 \leq t \leq S} |e^h(t)| \rightarrow 0$. Combining these facts with the representation (2.20), we have that $(\xi^h(\cdot), m^h(\cdot))$ converges w.p.1 to $(x(\cdot), m(\cdot))$, where

$$x(t) = x + \int_0^t \int_{\mathbb{R}^k} \alpha m_s(d\alpha) ds$$

and $m_s(d\alpha)ds = m(d\alpha ds)$.

The last sentence of the theorem follows from the same calculations as those given above, save that $\tau_h \wedge S$ replaces S_h . ■

We are now ready to prove the convergence theorems. It will often be convenient to have the chains ξ_i^h, u_i^h defined for all $i < \infty$, and not only up until the chain exits G^0 or is stopped by choice. Unless otherwise stated, it will be our convention that the control $\alpha = 0$ will be used after our interest in the chain stops.

Theorem 2.5. *Assume (A2.1), (A2.2), (A2.3), and that (2.6) holds. Then for the explicit scheme defined by (2.7) we have*

$$V^{h,\delta}(x) \rightarrow V(x).$$

Proof. Let $\{\xi_i^{h,\delta}, i < \infty\}$ be a controlled Markov chain as in Subsection 13.2.1 with the transition probabilities $\tilde{p}^{h,\delta}(x, y|\alpha)$, interpolation interval δ , and initial condition $x \in G^0$. Let $\{u_i^{h,\delta}, i < \infty\}$ denote a sequence of controls that are applied to the chain. We define the interpolated process and control by setting

$$\xi^{h,\delta}(t) = \xi_i^{h,\delta}, \quad u^{h,\delta}(t) = u_i^{h,\delta}, \quad t \in [i\delta, i\delta + \delta).$$

Define $\tau_{h,\delta}$ to be the first time $\xi^{h,\delta}(\cdot)$ leaves G_h^0 . As usual, we use relaxed rather than ordinary controls to prove the convergence of the scheme. We define $m^{h,\delta}(\cdot)$ by setting

$$m_s^{h,\delta}(A) = I_A(u^{h,\delta}(s)), \quad m^{h,\delta}(A \times [0, t]) = \int_0^t m_s^{h,\delta}(A) ds$$

for all Borel sets A in \mathbb{R}^k .

In the proofs of the upper and lower bounds we will have to deal with a number of special cases, due to the possible discontinuity of $g(\cdot, \cdot)$. The various cases are all determined by whether or not the process exits G , and if so, by the time at which it exits. Regrettably, this lengthens the proof somewhat.

Proof of the Lower Bound. For any $\epsilon > 0$, let $\{m^{h,\delta}(\cdot), h > 0, \delta > 0\}$ be the relaxed control representation of an ϵ -optimal sequence of ordinary admissible controls. In proving the lower bound, we can assume without loss that the running costs associated to this sequence are bounded from above, and therefore that Theorem 2.4 applies. Let (h, δ) denote any subsequence, and retain (h, δ) to denote a further subsequence along which $(\xi^{h,\delta}(\cdot), m^{h,\delta}(\cdot), \tau_{h,\delta})$ converges weakly. We have the inequality

$$\begin{aligned} V^{h,\delta}(x) &\geq E_x^{m^{h,\delta}} \int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbb{R}^k} k(\xi^{h,\delta}(s), \alpha) m^{h,\delta}(d\alpha ds) \\ &\quad + E_x^{m^{h,\delta}} g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}) - \epsilon. \end{aligned}$$

Assume we are using the Skorokhod representation, so that the convergence $(\xi^{h,\delta}(\cdot), m^{h,\delta}(\cdot), \tau_{h,\delta}) \rightarrow (x(\cdot), m(\cdot), \tilde{\tau})$ is w.p.1. Let $\tau = \inf\{t : x(t) \in \partial G\}$, and fix an $\omega \in \Omega$ such that the convergence holds. We claim that

$$\liminf_{(h,\delta) \rightarrow 0} \left[\int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbb{R}^k} k(\xi^{h,\delta}(s), \alpha) m^{h,\delta}(d\alpha ds) + g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}) \right] \geq V(x). \quad (2.21)$$

The case $\tilde{\tau} < T, \tau = \tilde{\tau}$. In this case (2.21) follows from the convergence $(\xi^{h,\delta}(\cdot), m^{h,\delta}(\cdot), \tau_{h,\delta}) \rightarrow (x(\cdot), m(\cdot), \tilde{\tau})$, Fatou's lemma, and (A2.2) and (A2.3).

The case $\tilde{\tau} < T, \tau < \tilde{\tau}$. For this case we still have the left hand side of (2.21) bounded below by

$$\int_0^{\tilde{\tau}} \int_{\mathbb{R}^k} k(x(s), \alpha) m(d\alpha ds) + g(x(\tilde{\tau}), \tilde{\tau}).$$

Since $\tau < \tilde{\tau}$, this cost is not the cost associated with $(x(\cdot), m(\cdot))$ in the definition of $V(x)$. However, by Theorem 2.3 there exists a relaxed control $\bar{m}(\cdot)$ with associated solution $\bar{x}(\cdot)$ that starts at x , remains in G^0 for $t \in (0, \tilde{\tau})$, satisfies $\bar{x}(\tilde{\tau}) = x(\tilde{\tau})$, and has running cost arbitrarily close to that of $((m(\cdot), x(\cdot)))$. Thus, (2.21) holds in this case as well.

The cases $\tilde{\tau} > T, \tau = \tilde{\tau}$ and $\tilde{\tau} > T, \tau < \tilde{\tau}$ are handled in the same way as $\tilde{\tau} < T, \tau = \tilde{\tau}$ and $\tilde{\tau} < T, \tau < \tilde{\tau}$, respectively. Finally, we must deal with $\tilde{\tau} = T$. The only difference between this case and the previous cases is that $g(\cdot, \cdot)$ may be discontinuous at $(x(T), T)$. Consider first the case when $\tau = \tilde{\tau}$, i.e., $x(t) \in G^0$ for $t \in [0, T]$. In this case, (2.21) follows from the fact that

$$\int_0^T \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds + g_*(x(T), T) \geq V(x),$$

where $g_*(x, T) = \lim_{\epsilon \rightarrow 0} \inf \{g(y, T) : |y - x| \leq \epsilon\}$. To prove the last inequality, it is enough to note that if $\lim_{\epsilon \rightarrow 0} \inf \{g(y, T) : |x(T) - y| \leq \epsilon, y \in G^0\} < g(x(T), T)$, then by (A2.1) and (A2.2) an admissible relaxed control $\bar{m}(\cdot)$ can be found which has running cost that is arbitrarily close to that of $m(\cdot)$ and for which the associated solution $\bar{x}(\cdot)$ stays in G^0 and terminates at a point $\bar{x}(T)$ which can be made arbitrarily close to $x(T)$. The case $\tau < \tilde{\tau}$ is similar, save that Theorem 2.3 must be applied to deal with any parts of $\bar{x}(\cdot)$ on ∂G for $t \in [0, T]$. Thus, (2.21) holds w.p.1.

By Fatou's lemma $\liminf_{(h, \delta) \rightarrow 0} V^{h, \delta}(x) \geq V(x) - \epsilon$ along the convergent subsequence. To prove $\liminf_{(h, \delta) \rightarrow 0} V^{h, \delta}(x) \geq V(x)$ for the original sequence, we argue by contradiction, and then send $\epsilon \rightarrow 0$.

Proof of the Upper Bound. Fix $\epsilon > 0$ and let $m(\cdot)$ be an ϵ -optimal admissible relaxed control for $V(x)$. Let $x(\cdot)$ denote the associated solution. We first show we can assume by modifying $m(\cdot)$ if need be] that $x(\cdot)$ either exits G^0 in a “regular” way or else remains in G^0 .

Case 1. Suppose $x(t) \in G^0$ for $t \in [0, T]$. In this case no modification is needed, and we leave $m(\cdot)$ unchanged.

Case 2. Suppose that $\tau = \inf\{t : x(t) \in \partial G\} < T$. Assumptions (A2.1) and (A2.2) imply that we may redefine $m_s(\cdot)$ for s in some interval of the form $s \in (\tau, \nu], \nu > \tau$ in such a way that $x(t) \notin G$ for $t \in (\tau, \nu]$ and

$$\int_\tau^t \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds \rightarrow 0$$

as $t \rightarrow \tau$.

Case 3. Suppose that $x(t) \in G^0$ for $t \in [0, T]$ and $x(T) \in \partial G$. Then there are two possibilities. If $\lim_{\epsilon \rightarrow 0} \inf \{g(y, T) : |x(T) - y| \leq \epsilon, y \in G^0\} \leq g(x(T), T)$, then by the same argument as in Theorem 2.3 there exists a control $m^\epsilon(\cdot)$ which is 2ϵ -optimal and for which the associated solution remains in G^0 for all $t \in [0, T]$. Otherwise, we can use the exterior cone condition of (A2.1) and (A2.2) and find a control $m^\epsilon(\cdot)$ which is 2ϵ -optimal

and for which the associated solution exits G^0 before time T . Hence, case 3 reduces to either case 1 or case 2.

The main point of all this is that we can always find an ϵ -optimal control that avoids the problematic case where $x(\cdot)$ exits G^0 at time T . More precisely, we may assume that given any $\epsilon > 0$ there exists an ϵ -optimal control $m(\cdot)$ with associated solution $x(\cdot)$ such that either

$$x(\cdot) \in G^0 \text{ for all } t \in [0, T], \quad (2.22)$$

or else there is $\tau \in [0, T)$ and $\nu > 0$ such that

$$\begin{aligned} x(t) \in G^0 \text{ for } t \in [0, \tau), \quad x(t) \notin G \text{ for } t \in (\tau, \tau + \nu], \\ \int_{\tau}^t \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds \rightarrow 0 \text{ as } t \rightarrow \tau. \end{aligned} \quad (2.23)$$

The case $\tau > T$. Let any $\epsilon_1 > 0$ be given. By Theorem 2.2 there is a finite set $\mathcal{U}_{\epsilon_1} \subset \mathbb{R}^k$, $\bar{\delta} > 0$, and an ordinary control $u^{\epsilon_1}(\cdot)$ with the following properties. $u^{\epsilon_1}(\cdot)$ takes values in \mathcal{U}_{ϵ_1} , is constant on intervals of the form $[j\bar{\delta}, j\bar{\delta} + \bar{\delta})$, and if $x^{\epsilon_1}(\cdot)$ is the associated solution, then

$$\sup_{0 \leq t \leq T} |x^{\epsilon_1}(t) - x(t)| \leq \epsilon_1$$

and

$$\sup_{0 \leq t \leq T} \left| \int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds - \int_0^t k(x^{\epsilon_1}(s), u^{\epsilon_1}(s)) ds \right| \leq \epsilon_1.$$

Under (2.6) we have $\mathcal{U}_{\epsilon_1} \subset \mathcal{U}^{h, \delta}$ whenever $h > 0$ and $\delta > 0$ are sufficiently small. We apply the control $u^{\epsilon_1}(\cdot)$ to the chain $\{\xi_i^{h, \delta}, i < \infty\}$ in the obvious way, namely,

$$u_i^{h, \delta} = u^{\epsilon_1}(j\bar{\delta}) \text{ for } i\delta \in [j\bar{\delta}, j\bar{\delta} + \bar{\delta}).$$

Define $\xi^{h, \delta}(\cdot)$ to be the piecewise constant interpolation of $\{\xi_i^{h, \delta}, i < \infty\}$ with interpolation interval δ . Then, by Theorem 2.4, $\xi^{h, \delta}(\cdot) \rightarrow x^{\epsilon_1}(\cdot)$ w.p.l. Suppose we first let $(h, \delta) \rightarrow 0$ and then send $\epsilon_1 \rightarrow 0$. By applying the dominated convergence theorem and using (A2.2) and (A2.3), we have

$$\limsup_{(h, \delta) \rightarrow 0} V^{h, \delta}(x) \leq V(x) + \epsilon.$$

The case $\tau < T$. Next consider the case of (2.23). As in the case of (2.22) we can assume the existence of piecewise constant ordinary control with the given properties save that T is replaced by ν . Because $\sup_{0 \leq t \leq \nu} |x^{\epsilon_1}(t) - x(t)|$ can be made arbitrarily small, (2.23) implies we can also assume that

$$x^{\epsilon_1}(t) \in G^0 \text{ for } t \in [0, \tau - \epsilon_1), \quad x^{\epsilon_1}(t) \notin G \text{ for } t \in (\tau + \epsilon_1, \nu]. \quad (2.24)$$

We now apply $u^{\epsilon_1}(\cdot)$ as in the previous case. By sending $(h, \delta) \rightarrow 0$ and then $\epsilon_1 \rightarrow 0$, we obtain $\limsup_{(h, \delta) \rightarrow 0} V^{h, \delta}(x) \leq V(x) + \epsilon$ from (A2.2), (A2.3), (2.23), and (2.24). Since $\epsilon > 0$ is arbitrary, the upper bound is proved. ■

A very similar proof gives the following result.

Theorem 2.6. *Assume (A2.1), (A2.2), and (A2.3). Then for the implicit scheme defined by (2.11) we have*

$$V^{h, \delta}(x) \rightarrow V(x).$$

13.3 Problems with a Controlled Stopping Time

We now turn to problems in which the process is stopped at a chosen time ρ or the first exit time from G^0 , whichever is smaller. Owing to the fact that ρ is potentially unbounded, we must impose suitable conditions on the running cost $k(\cdot, \cdot)$ in order to guarantee that the minimal cost is bounded from below. We will make the assumption that $k(x, \alpha) \geq 0$ for all x and α . The mathematical problem (1.6) may still be well posed even if this condition is violated. However, because most current applications satisfy the non-negativity condition we restrict our attention to this case.

Given that $k(\cdot, \cdot)$ is non-negative, there are two essentially different cases. In the first case, we assume the existence of $k_0 > 0$ such that $k(x, \alpha) \geq k_0$ for all x and α . This case is relatively straightforward, since the lower bound on the running cost provides an upper bound on the optimal stopping time, as in Section 10.6. This case is discussed briefly in Subsection 13.3.1. The remaining case is much more difficult. Here we have no a priori estimates on the optimal stopping time, and the convergence theorem for the case of strictly positive running cost no longer applies. In Subsection 13.3.2 we will describe a numerical method for this case that is significantly different from that of Subsection 13.3.1. We first prove that if $V^h(x)$ is the resulting approximation, then $\limsup_{h \rightarrow 0} V^h(x) \leq V(x)$ holds in a general setting. We then consider the reverse inequality $\liminf_{h \rightarrow 0} V^h(x) \geq V(x)$. It turns out that in order to apply the Markov chain method to prove this lower bound, we need additional assumptions on the “zero” sets $K(x) = \{\alpha : k(x, \alpha) = 0\}$. These conditions essentially take the form of a description of the large time behavior of any solutions to the differential inclusion $\dot{\phi} \in K(\phi)$. In Subsection 13.3.2 we consider first the case in which all solutions to the differential inclusion that start in a neighborhood of G are either attracted to a single point $x_0 \in G^0$ or else leave G . We then extend this result to cover the case in which the limit set is given by a finite collection of connected sets.

In Subsection 13.3.3 we derive an algorithm and present some numerical data for a particular example, namely, a calculus of variations problem that

occurs in shape-from-shading. Although this example is included in the problems of Subsection 13.3.2, particular properties of the problem allow an interesting alternative approach. After describing the numerical scheme we make several important observations concerning the implementation of all the schemes introduced in this section. Also included in Subsection 13.3.3 is a weakening of (A2.1) in which we allow the target set to be a point.

Except where explicitly stated otherwise we will use (A2.1), (A2.2), and the variation on (A2.3) given by:

A2.3'. *The function $g(\cdot)$ is uniformly continuous and bounded when restricted to either of the sets G^0 and $\mathbb{R}^k - G^0$.*

13.3.1 THE CASE $k(x, \alpha) \geq k_0 > 0$

Let $p^h(x, y|\alpha)$ and $\Delta t^h(\alpha)$ satisfy the local consistency conditions (2.1) and (2.2), and let $\{\xi_i^h, i < \infty\}$ denote the associated controlled Markov chain.

The Numerical Scheme. Our scheme for approximating the solution to (1.6) is given by

$$V^h(x) = \min \left[g(x), \min_{\alpha \in \mathbb{R}^k} \left[\sum_y p^h(x, y|\alpha) V^h(y) + k(x, \alpha) \Delta t^h(\alpha) \right] \right] \quad (3.1)$$

if $x \in G_h^0$ and $V^h(x) = g(x)$ for $x \notin G_h^0$. We then have the following result.

Theorem 3.1. *Assume (A2.1), (A2.2), and (A2.3'). Then for the scheme defined by (3.1) we have*

$$V^h(x) \rightarrow V(x).$$

The proof combines the ideas used in the proofs of Theorems 10.6.1 and 2.5 and is omitted.

13.3.2 THE CASE $k(x, \alpha) \geq 0$

We again assume $p^h(x, y|\alpha)$ and $\Delta t^h(\alpha)$ satisfy the local consistency conditions (2.1) and (2.2) and take $\{\xi_i^h, i < \infty\}$ to be the associated controlled Markov chain. Without a positive lower bound on $k(\cdot, \cdot)$, a naive discretization of the formally derived Bellman equation can yield incorrect results [although an examination of (3.1) shows that such an approach would work for the case of Subsection 13.3.1]. The basic difficulties can be illustrated by a simple example. Suppose $0 \in G^0$ and that $k(0, 0) = 0$. Then under (A2.3') the problem (1.6) is still well defined in the sense that $V(x)$ is bounded from below. However, there is a difficulty with the evaluation of $V(0)$ in that there is no guarantee of the existence of a minimizing path $\phi(\cdot)$ and associated controlled stopping time ρ which satisfy $\rho \wedge \inf\{t : \phi(t) \in \partial G\} < \infty$.

Furthermore, if $\phi(\cdot)$ is ϵ -optimal, then so is $(\phi^*(\cdot), \rho^*)$, where $\rho^* = \rho + c$,

$$\phi^*(t) = \begin{cases} \phi(t - c) & \text{for } t \geq c \\ 0 & \text{for } t \in [0, c], \end{cases}$$

and $c \geq 0$ is arbitrary. The facts that ρ can be arbitrarily large and $k(0, 0) = 0$ imply that a nearly minimizing path has no need to leave the origin in any finite amount of time.

Suppose one were tempted to use the scheme (3.1) with $p^h(x, y|\alpha)$ and $\Delta t^h(\alpha)$ as given in Example 2.1 to approximate $V(x)$. Recall that in many cases, equation (3.1) itself must ultimately be solved by one of the methods discussed in Chapter 6, such as the iteration in policy space method. It turns out that there are feedback controls $u^h(x)$ for which the associated cost is bounded and for which the matrix defined by

$$R = \{r(x, y), x, y \in \partial G_h \cup G_h^0\},$$

$$r(x, y) = \begin{cases} p^h(x, y|u^h(x)) & x \in G_h^0 \\ 0 & x \notin G_h^0, \end{cases}$$

$$\partial G_h = \{y \in G_h^0 : p^h(x, y|\alpha) > 0, \text{ for some } x \in G_h^0, \alpha \in \mathbb{R}^k\}$$

is *not* a contraction. (See the discussion in Chapter 6 on the role of the contraction condition.) This means that there will not, in general, be uniqueness for the solutions to (3.1). For a very simple example, assume $G_h^0 = \{0\}$ and let $p^h(x, y|\alpha)$ and $\Delta t^h(\alpha)$ be as in Example 2.1. Then $V^h(0) = c$ is a solution to (3.1) for any $c \leq \inf_{x \in G} g(x)$, with the associated feedback control given by $u^h(0) = 0$. More interesting examples are given later in this subsection and in Subsection 13.3.3.

In order to circumvent this difficulty, we return to the definition of $V(x)$. For each $\eta > 0$ let $k^\eta(x, \alpha)$ be a function that satisfies (A2.2), and also

$$k^\eta(x, \alpha) \geq \eta, \quad k^\eta(x, \alpha) \downarrow k(x, \alpha)$$

for all x and α . For example, we can take $k^\eta(x, \alpha) = k(x, \alpha) \vee \eta$. Define

$$V_\eta(x) = \inf \left[\int_0^{\rho \wedge \tau} k^\eta(\phi(s), \dot{\phi}(s)) ds + g(\phi(\rho \wedge \tau)) \right], \quad (3.2)$$

where the infimum is over the same functions and stopping times as in the definition of $V(x)$. Clearly, $k^\eta(\cdot, \cdot)$ satisfies the conditions of Subsection 13.3.1. Furthermore, it is quite easy to prove that $V_\eta(x) \downarrow V(x)$ as $\eta \rightarrow 0$.

Based on the strict positivity of $k^\eta(\cdot, \cdot)$, the methods of Subsection 13.3.1 can be applied to obtain approximations to $V_\eta(x)$. If $\eta > 0$ is sufficiently small, then $V_\eta(x)$ should be a good approximation to $V(x)$. Although this is the basic idea behind the algorithm presented below, it turns out that

we cannot send $\eta \rightarrow 0$ and $h \rightarrow 0$ in an arbitrary fashion. Rather, η and h must simultaneously be sent to their respective limits in a specific manner.

Numerical Schemes. Let $p^h(x, y|\alpha)$ and $\Delta t^h(\alpha)$ satisfy (2.1) and (2.2). For each $\eta > 0$, we define an approximation $V_\eta^h(x)$ of $V_\eta(x)$ by

$$V_\eta^h(x) = \min \left[g(x), \min_{\alpha \in \mathbb{R}^k} \left[\sum_y p^h(x, y|\alpha) V_\eta^h(y) + k^\eta(x, \alpha) \Delta t^h(\alpha) \right] \right] \quad (3.3)$$

if $x \in G_h^0$ and $V_\eta^h(x) = g(x)$ for $x \notin G_h^0$. It follows from the positivity of η that (3.3) uniquely defines $V_\eta^h(x)$. Our approximation to $V(x)$ will then be given by $V^h(x) = V_{\eta(h)}^h(x)$, where $\eta(h)$ tends to zero as $h \rightarrow 0$.

Remarks on Implementation. The strict positivity of $k^\eta(\cdot, \cdot)$ for $\eta > 0$ implies that (3.3) has a unique solution. Thus, the approximation $V_\eta^h(x)$ can be obtained by solving (3.3) by any of the methods of Chapter 6 (e.g., the approximation in value space method). However, it turns out that the performance of the algorithm used to solve (3.3) can depend heavily on the initial condition used to start the algorithm. In particular, choosing too low an initial condition can cause the algorithm to converge quite slowly. The effect of a low initial condition on the speed of convergence becomes more and more severe as the infimum of $k(\cdot, \cdot)$ tends to zero. On the other hand, convergence is typically quite fast when a large initial condition is used for the algorithm. This behavior is, in part, due to the fact that we are working with a minimization problem and is also due to the deterministic nature of the controlled process that the Markov chain is approximating [i.e., $\dot{\phi}(t) = u(t)$]. More extensive remarks on these points will be given in Subsection 13.3.3.

As remarked above, it will ultimately become necessary to impose a condition on the manner in which $\eta(h)$ tends to zero. However, this condition is not needed for the upper bound, and, therefore, we separate the statements for the different bounds into two theorems. The first deals with the upper bound. We will use the same notation and definitions as in the proof of Theorem 2.4 save that, for each $h > 0$, we replace $k(\cdot, \cdot)$ by $k^{\eta(h)}(\cdot, \cdot)$.

Theorem 3.2. *Assume (A2.1), (A2.2), and (A2.3') and that $\eta(h) > 0$ is any sequence satisfying $\eta(h) \rightarrow 0$ as $h \rightarrow 0$. Then*

$$\limsup_{h \rightarrow 0} V_{\eta(h)}^h(x) \leq V(x).$$

Proof. Let any $\epsilon > 0$ be given. Because $k^\eta(\cdot, \cdot) \downarrow k(\cdot, \cdot)$, there exists a relaxed control $m(\cdot)$ and an associated stopping time $\rho < \infty$ such that if

$$x(t) = x + \int_0^t \int_{\mathbb{R}^k} \alpha m(d\alpha ds), \quad \tau = \inf\{t : x(t) \in \partial G\},$$

then

$$\int_0^{\rho \wedge \tau} \int_{\mathbb{R}^k} k^\eta(x(s), \alpha) m_s(d\alpha) ds + g(x(\rho \wedge \tau)) \leq V(x) + \epsilon$$

for all sufficiently small $\eta > 0$. Using Theorems 2.2 and 2.3 in the same way as they were used in the proof of Theorem 2.5 we can extend the definition of $m(\cdot)$ beyond τ such that $m(\cdot)$, $x(\cdot)$, and ρ have the following properties. Either

$$x(\cdot) \in G^0 \text{ for all } t \in [0, \rho] \quad (3.4)$$

or $\rho > \tau$, and there is $\nu > 0$ such that

$$\begin{aligned} x(t) &\in G^0 \text{ for } t \in [0, \tau], \quad x(t) \notin G \text{ for } t \in (\tau, \tau + \nu], \\ \int_\tau^t k^\eta(x(s), \alpha) m_s(d\alpha) ds &\rightarrow 0 \text{ as } t \downarrow \tau \end{aligned} \quad (3.5)$$

uniformly in all sufficiently small $\eta > 0$. The remainder of the proof for these two possibilities is analogous to the proof of Theorem 2.5 for the two cases of (2.22) and (2.23). We will give the details only for the case of (3.4). The proof for (3.5) is a straightforward combination of the ideas used in the proofs for the cases (3.4) and (2.23).

By Theorem 2.2, we may assume the existence a finite set $\mathcal{U}_{\epsilon_1} \subset \mathbb{R}^k$, $\bar{\delta} > 0$, and an ordinary control $u^{\epsilon_1}(\cdot)$ with the following properties. $u^{\epsilon_1}(\cdot)$ takes values in \mathcal{U}_{ϵ_1} , is constant on intervals of the form $[j\bar{\delta}, j\bar{\delta} + \bar{\delta})$, and if $x^{\epsilon_1}(\cdot)$ is the associated solution, then

$$\sup_{0 \leq t \leq \rho} |x^{\epsilon_1}(t) - x(t)| \leq \epsilon_1$$

and

$$\sup_{0 \leq t \leq \rho} \left| \int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha) m_s(d\alpha) ds - \int_0^t k(x^{\epsilon_1}(s), u^{\epsilon_1}(s)) ds \right| \leq \epsilon_1.$$

In order to apply $u^{\epsilon_1}(\cdot)$ to the chain $\{\xi_i^h, i < \infty\}$, we recursively define the control applied at discrete time i by $u_i^h = u^{\epsilon_1}(t_i^h)$ until $\rho_h \equiv t_{i+1}^h \geq \rho$, where $t_{i+1}^h = t_i^h + \Delta t^h(u_i^h)$. By construction, the interpolated stopping times ρ_h are actually deterministic and satisfy

$$\rho_h \rightarrow \rho.$$

The proof from this point on is an easy consequence of the fact (due to Theorem 2.4) that $\sup_{0 \leq t \leq \rho_h} |\xi^h(t) - x^{\epsilon_1}(t)| \rightarrow 0$ in probability, and is the same as that of the finite time problem treated in Theorem 2.5. ■

It will turn out that much more is required for the lower bound, including a condition on the rate at which $\eta(h) \rightarrow 0$. In the next example we show that if $\eta(h) \rightarrow 0$ too quickly, then

$$\limsup_{h \rightarrow 0} V_{\eta(h)}^h(x) < V(x)$$

is possible.

Example. We consider a two dimensional problem in which $G = \{(x_1, x_2) : |x_1| \vee |x_2| \leq 1\}$ and $k(x, \alpha) = |\alpha_1 - \alpha_2|$ for $|\alpha| \leq 2$. We also assume that $k(\cdot, \cdot)$ is defined for $|\alpha| > 2$ in such a way that (A2.2) holds [e.g., $k(x, \alpha) = |\alpha_1 - \alpha_2| \vee (|\alpha|^2 - 2)$]. The zero cost controls therefore include the set $\{(\alpha_1, \alpha_2) : \alpha_1 = \alpha_2, |\alpha| \leq 2\}$. We consider a stopping cost $g(\cdot)$ that satisfies $g(x) = 1$ for $x \in G^0$. We will also assume that the restriction of $g(\cdot)$ to ∂G is continuous, with $g(x) \geq 0$ for all $x \in \partial G$, $g(x) = 0$ if $x \in \{x \in \partial G : |x_1 + x_2| \leq 1/2\}$, and $g(1, 1) = g(-1, -1) = 1$. Thus, the stopping cost on ∂G is strictly positive in a neighborhood of each of the northeast and southwest corners $(1, 1)$ and $(-1, -1)$, and identically equal to zero in a neighborhood of each of $(-1, 1)$ and $(1, -1)$. For S_h , G_h , $p^h(x, y|\alpha)$, and $\Delta t^h(\alpha)$ we will use the natural choices given in Example 2.1.

It is not hard to see that our assumptions on the running and stopping costs imply $V(0, 0) > 0$. However, as we now show, one can choose $\eta(h) > 0$ such that

$$\lim_{h \rightarrow 0} V_{\eta(h)}^h(0, 0) = 0 < V(0, 0).$$

Clearly, $\liminf_{h \rightarrow 0} V_{\eta(h)}^h(0, 0) \geq 0$. To prove $\limsup_{h \rightarrow 0} V_{\eta(h)}^h(0, 0) \leq 0$, we will exhibit a control scheme and $\eta(h) > 0$ that achieves cost arbitrarily close to zero. For each i , let

$$u_i^h = \begin{cases} (-1, -1) & \text{for } (\xi_i^h)_1 + (\xi_i^h)_2 > 0 \\ (1, 1) & \text{for } (\xi_i^h)_1 + (\xi_i^h)_2 \leq 0. \end{cases}$$

Note that the running cost for this control process is identically zero. With this definition of the control, the process $\{\xi_i^h, i < \infty\}$ always satisfies $|(\xi_i^h)_1 + (\xi_i^h)_2| < 1/2$ (for small enough h). Thus the control scheme accumulates no running cost and also prohibits the process from exiting near either the northeast or southwest corners (where the exit cost is high). However, it does allow exit near either of the other two corners, where the exit cost is zero. The process $\{(\xi_i^h)_1 - (\xi_i^h)_2, i < \infty\}$ is easily recognized as a symmetric random walk on $\{ih, i \in \mathbb{Z}\}$. As is well known, this implies that for any given value M , there exists $i < \infty$ such that $|(\xi_i^h)_1 - (\xi_i^h)_2| \geq M$ (w.p.1). Therefore, given any $\epsilon > 0$ and the initial condition $\xi_0^h = (0, 0)$, there exists $n < \infty$ such that the process ξ_i^h will enter ∂G_h at some point x [with $g(x) = 0$] by discrete time n and with probability greater than $1 - \epsilon$. Thus, $\limsup_{h \rightarrow 0} V_{\eta(h)}^h(0, 0) = 0$, which demonstrates the existence of $\eta(h) > 0$ such that $V_{\eta(h)}^h(0, 0) \rightarrow 0$. ■

A Condition On the Rate at Which $\eta(h) \rightarrow 0$. Let $p^h(x, y|\alpha)$ and $\Delta t^h(\alpha)$ satisfy (2.1) and (2.2). In particular, let $c(h)$ be such that

$$|o(|\alpha| \Delta t^h(\alpha))| \leq c(h) |\alpha| \Delta t^h(\alpha),$$

where $o(|\alpha| \Delta t^h(\alpha))$ represent the “error” terms in the local consistency equations (2.1) and (2.2) and $c(h) \rightarrow 0$. For example, if the transition probabilities of Example 2.1 are used then we can take $c(h) = h$.

We then impose the following condition on the rate at which $\eta = \eta(h)$ can tend to zero:

$$\eta(h)/c(h) \rightarrow \infty \quad (3.6)$$

as $h \rightarrow 0$.

The absence of an a priori bound on the optimal stopping times for nearly optimal approximating processes $\{\xi_i^h, i < \infty\}$ means that we will need yet another assumption besides (3.6) in order to prove convergence. As remarked at the beginning of this section, the assumption takes the form of conditions that are placed on the zero sets $K(x) = \{\alpha : k(x, \alpha) = 0\}$ in order to prove the lower bound. We begin by considering a very simple example of such an assumption. This assumption will be weakened and discussed further in (A3.2). For a set or point A , define $N_\gamma(A) = \{x : d(x, A) \leq \gamma\}$, where $d(x, A)$ is the Euclidean distance between x and A .

A3.1. *There exist a point $x_0 \in G^0$ and an open set G_1 containing G with the following properties. Given any $\gamma > 0$, there exists $T < \infty$ such that for all $t \geq T$,*

$$x(0) \in G^0, \dot{x}(s) = \int_{\mathbb{R}^k} \alpha m_s(d\alpha) \quad (\text{a.s.})$$

and

$$\int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha) m(d\alpha ds) = 0$$

imply

$$x(t) \in N_\gamma(x_0) \cup (\mathbb{R}^k - G_1).$$

Remarks. Thus, solutions with zero running cost either are attracted to the point x_0 or else leave on open neighborhood of G in finite time. The sets $K(x)$ often have a particular significance for the underlying problem which gave rise to the calculus of variations problem. See the book [48] and the example of the next subsection. Note that we may, of course, have points x where $K(x) = \emptyset$.

Theorem 3.3. *Assume (A2.1), (A2.2), (A2.3'), and (A3.1). Then for the scheme defined by (3.6) and (3.3), we have*

$$\liminf_{h \rightarrow 0} V^h(x) \geq V(x).$$

Before giving the proof, we state a lemma. The proof of a generalization of the lemma will be given later in this section.

Lemma 3.4. Assume (A2.2) and (A3.1). Then given any $\gamma > 0$ and $M < \infty$, there is $T < \infty$ such that if $\dot{x}(t) = \int_{\mathbb{R}^k} \alpha m_t(d\alpha)$ (a.s.) for an admissible relaxed control $m(\cdot)$, and if $x(t) \in G - N_\gamma(x_0)$ for $t \in [0, T]$, then $\int_0^T \int_{\mathbb{R}^k} k(x(s), \alpha) m(d\alpha ds) \geq M$.

As noted previously, a difficulty with applying weak convergence methods directly to this problem is the lack of information regarding the large time behavior of limits of the interpolated Markov chains. Consider the case when the stopping cost is so high that the controlled stopping option is never exercised. Suppose the interpolated processes corresponding to a given sequence of controls spend an ever larger amount of time in an ever smaller neighborhood of x_0 , before moving out of G^0 . In particular, assume that the limit paths of the process are all identically equal to x_0 . Then these limit paths *never* exit G^0 , and the substantial control effort needed to move the process from x_0 to ∂G will not appear in the limit of the running costs.

The technique we will use to avoid this problem is to consider only those sections of the sample path when the process is outside of a neighborhood of x_0 . Unfortunately, keeping track of such detailed information regarding the process will require some additional notation. A consequence of Lemma 3.4 will be that for small h the interpolated processes can only spend a finite amount of time in this region without building up an arbitrarily large running cost. It turns out that the uniformity given in the last statement of Theorem 2.4 allows us to concentrate on these sections of the sample paths (regardless of when they occur), and thus prove the lower bound.

Proof of Theorem 3.3. It is quite easy to show under (A2.1) and (A2.2) that given $\epsilon > 0$ there is $\gamma > 0$ such that $|x - y| \leq \gamma$ implies $V(x) \geq V(y) - \epsilon$ for all $x, y \in G$ [i.e., $V(\cdot)$ is uniformly continuous on G]. Thus, for any $\epsilon > 0$, there is $\gamma > 0$ such that $|V(x) - V(x_0)| \leq \epsilon$ for $x \in N_\gamma(x_0)$. For this γ and $M = V(x_0) + 1$, choose T according to Lemma 3.4.

We first consider the case when $x \in N_\gamma(x_0)$. Owing to the definition of $V^h(x)$ there is a controlled Markov chain $\{\xi_i^h, u_i^h, i < \infty\}$ that satisfies $\xi_0^h = x$ and a controlled stopping time N_h such that

$$V^h(x) \geq E_x^{u^h} \sum_{j=0}^{(N_h \wedge M_h)-1} k^{\eta(h)}(\xi_j^h, u_j^h) \Delta t^h(u_j^h) + E_x^{u^h} g(\xi_{N_h \wedge M_h}^h) - \epsilon, \quad (3.7)$$

where M_h is the time of first exit from G^0 . Let $\xi^h(\cdot)$ and $u^h(\cdot)$ be the continuous parameter interpolations of $\{\xi_i^h, i < \infty\}$ and $\{u_i^h, i < \infty\}$, respectively, and let

$$\rho_h = \sum_{i=0}^{N_h-1} \Delta t^h(u_i^h), \quad \tau_h = \sum_{i=0}^{M_h-1} \Delta t^h(u_i^h).$$

We can then rewrite (3.7) as

$$V^h(x) \geq E_x^{m^h} \int_0^{\rho_h \wedge \tau_h} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) + E_x^{m^h} g(\xi^h(\rho_h \wedge \tau_h)) - \epsilon, \quad (3.8)$$

where $m^h(\cdot)$ is the relaxed control representation of the ordinary control $u^h(\cdot)$. The superlinearity condition (1.1) implies

$$\limsup_{h \rightarrow 0} E_x^{m^h} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} I_{\{|\alpha| \geq b\}} |\alpha| m^h(d\alpha ds) < \infty$$

for some fixed $b < \infty$. Choose $s(h) \rightarrow \infty$ such that $c(h)s(h) \rightarrow 0$ and $\eta(h)s(h) \rightarrow \infty$. This is always possible under (3.6). When combined with $c(h)s(h) \rightarrow 0$, the last equation implies

$$\limsup_{h \rightarrow 0} c(h) E_x^{m^h} \int_0^{\rho_h \wedge \tau_h \wedge s(h)} \int_{\mathbb{R}^k} |\alpha| m^h(d\alpha ds) = 0.$$

By Theorems 2.4 and 9.1.7, we can assume $(\xi^h(\cdot), m^h(\cdot)) \rightarrow (x(\cdot), m(\cdot))$ w.p.1 and that

$$\sup_{0 \leq t \leq \rho_h \wedge \tau_h \wedge s(h)} \left| \xi^h(t) - \xi^h(0) - \int_0^t \int_{\mathbb{R}^k} \alpha m^h(d\alpha ds) \right| \rightarrow 0 \quad (3.9)$$

w.p.1.

Fix an $\omega \in \Omega$ such that these convergences hold. Define

$$i_h = \inf\{j : \xi_k^h \in G^0 - N_\gamma(x^0), \text{ for } k = j, \dots, (N_h \wedge M_h) - 1\}.$$

If no such j exists, we set $i_h = \infty$. Let

$$\sigma_h = \sum_{i=0}^{i_h-1} \Delta t^h(u_i^h),$$

and define

$$T_h = \begin{cases} (\rho_h \wedge \tau_h) - \sigma_h, & \sigma_h < \rho_h \wedge \tau_h \\ 0, & \sigma_h \geq \rho_h \wedge \tau_h. \end{cases}$$

Thus, T_h is the interpolated time between $\rho_h \wedge \tau_h$ and the time before that when $\xi^h(\cdot)$ last left $N_\gamma(x_0)$. We will now carefully examine the sample paths over the interval $[\sigma_h, (\rho_h \wedge \tau_h)]$.

Define $\bar{\xi}^h(\cdot) = \xi^h(\cdot + \sigma_h)$ and define $\bar{m}^h(\cdot)$ to be the relaxed control representation of the ordinary control $u^h(\cdot + \sigma_h)$. Finally, let $v^h = (\rho_h \wedge \tau_h)/s(h)$ and

$$r^h = \int_{\sigma_h}^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds).$$

For the fixed ω , consider the set

$$\{(\bar{m}^h(\cdot), T_h, \rho_h, \tau_h, \sigma_h, r^h, v^h), h > 0\}.$$

Since $\{\bar{m}^h(\cdot), h > 0\}$ is always precompact on the space $\mathcal{R}([0, \infty) \times \mathbb{R}^k)$, we can extract a convergent subsequence with limit denoted by $(\bar{m}(\cdot), T, \rho, \tau, \sigma, r, v)$.

The case $v > 0$. Since $v^h = (\rho_h \wedge \tau_h)/s(h)$, this case implies that $\rho_h \wedge \tau_h \rightarrow \infty$ quite fast. In fact, the condition $\eta(h)s(h) \rightarrow \infty$ implies

$$\liminf_{h \rightarrow 0} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) = \infty.$$

For the remainder of the proof we can assume $v = 0$. This implies that $\rho_h \wedge \tau_h < s(h)$ for sufficiently small $h > 0$ and, in particular, gives the estimate (3.9) without $s(h)$ appearing. We can also exclude the case $r = \infty$ because in this case the running costs again will tend to ∞ . An argument very similar to that used in Theorem 2.4 shows that if

$$\lim_{h \rightarrow 0} r^h = \lim_{h \rightarrow 0} \int_{\sigma_h}^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) < \infty, \quad (3.10)$$

then $\bar{m}(\mathbb{R}^k \times [0, T]) = T$ and

$$\int_0^t \int_{\mathbb{R}^k} \alpha \bar{m}^h(d\alpha ds) \rightarrow \int_0^t \int_{\mathbb{R}^k} \alpha \bar{m}(d\alpha ds)$$

as $h \rightarrow 0$ for $t \in [0, T]$. For the given ω it follows from (3.9) that

$$\begin{aligned} & (\bar{\xi}^h(\cdot), \bar{m}^h(\cdot), T_h, \xi^h(\cdot), m^h(\cdot), \rho_h, \tau_h, \sigma_h) \\ & \rightarrow (\bar{x}(\cdot), \bar{m}(\cdot), T, x(\cdot), m(\cdot), \rho, \tau, \sigma), \end{aligned}$$

where

$$\bar{x}(t) - \bar{x}(0) = \int_0^t \int_{\mathbb{R}^k} \alpha \bar{m}(d\alpha ds)$$

for $t \in [0, T]$. Note that by construction $\bar{x}(0) \in N_\gamma(x_0)$ and $\bar{x}(t) \notin N_{\gamma/2}(x_0)$ for $t \in [0, T]$. Thus, we are examining the limits of $\xi^h(\cdot)$ after the processes have left $N_{\gamma/2}(x_0)$ for the last time. Note also that it may be the case that $\sigma = \infty$, in which case $\bar{x}(\cdot)$ cannot be determined from $x(\cdot)$.

The case $\rho \wedge \tau < \infty$. For this case, we are essentially back in the setting of the finite time problem, and the proof of Theorem 2.5 shows that

$$\liminf_{h \rightarrow 0} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) + g(\xi^h(\rho_h \wedge \tau_h)) \geq V(x). \quad (3.11)$$

For the remainder of the proof we can assume $\rho = \tau = \infty$. Recall that, by construction, $\bar{x}(0) \in N_\gamma(x_0)$ and $\bar{x}(t) \in G - N_{\gamma/2}(x_0)$ for $t \in [0, T]$.

The case $T = \infty$. If $T = \infty$, then by Lemma 3.4

$$\begin{aligned} & \liminf_{h \rightarrow 0} \int_{\sigma_h}^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) \\ & \geq \int_0^\infty \int_{\mathbb{R}^k} k(\bar{x}(s), \alpha) \bar{m}(d\alpha ds) = \infty. \end{aligned} \tag{3.12}$$

The case $T < \infty$. By Fatou's lemma

$$\begin{aligned} & \liminf_{h \rightarrow 0} \int_{\sigma_h}^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) + g(\xi^h(\rho_h \wedge \tau_h)) \\ & \geq \int_0^T \int_{\mathbb{R}^k} k(\bar{x}(s), \alpha) \bar{m}(d\alpha ds) + g_*(\bar{x}(T)), \end{aligned} \tag{3.13}$$

where $g_*(x) = \lim_{\epsilon \rightarrow 0} \inf\{g(y) : |x - y| \leq \epsilon\}$. As noted in the remarks following (A2.1)–(A2.3) and also in the proof of Theorem 2.5, (A2.1) and (A2.2) imply that $V(\bar{x}(0))$, and therefore $V(x_0) - \epsilon$, give a lower bound for the right hand side of (3.13).

If we combine all the cases and use the assumption that $x \in N_\gamma(x_0)$, we obtain

$$\liminf_{h \rightarrow 0} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) + g(\xi^h(\rho_h \wedge \tau_h)) \geq V(x) - 3\epsilon$$

w.p.1. Therefore,

$$\liminf_{h \rightarrow 0} V^h(x) \geq V(x) - 3\epsilon \text{ uniformly for all } x \in N_\gamma(x_0) \tag{3.14}$$

for the original sequence h .

To complete the proof we must consider an arbitrary point in G^0 . Define the processes $\xi^h(\cdot)$ and $m^h(\cdot)$ in the usual fashion for any nonanticipating control sequence for which the associated running costs are bounded. Let $\tau_h = \inf\{t : \xi^h(t) \notin G^0\}$, $\bar{\tau}_h = \inf\{t : \xi^h(t) \in (\mathbb{R}^k - G^0) \cup N_\gamma(x_0)\}$, and let ρ_h be the continuous time interpolation of the stopping time N_h . Extract a convergent subsequence from the tight collection of random variables $\{(\xi^h(\cdot), m^h(\cdot), \rho_h, \tau_h, \bar{\tau}_h), h > 0\}$ and let $(x(\cdot), m(\cdot), \rho, \tau, \bar{\tau})$ denote the limit. We assume the convergence is w.p.1, and for each $\omega \in \Omega$ for which there is convergence we consider separately the possible cases: (1) $\rho \leq \tau, \rho < \infty$; (2) $\tau = \bar{\tau} < \rho$; (3) $\bar{\tau} < \tau \leq \rho$; (4) $\bar{\tau} = \tau = \rho = \infty$. For the cases (1) and (2) we are again dealing with a finite time problem (as far as the convergence

of the costs is concerned), and we have

$$\liminf_{h \rightarrow 0} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) + g(\xi^h(\rho_h \wedge \tau_h)) \geq V(x).$$

For case (3), the definition of $\bar{\tau}_h$ implies $\xi^h(\bar{\tau}_h) \in N_\gamma(x_0)$ for all small $h > 0$. Using the Markov property and the previous estimates for paths that start inside $N_\gamma(x_0)$, we have

$$\begin{aligned} & \liminf_{h \rightarrow 0} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) + g(\xi^h(\tau_h)) \\ & \geq \liminf_{h \rightarrow 0} \int_0^{\bar{\tau}_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) \\ & \quad + \liminf_{h \rightarrow 0} \int_{\bar{\tau}_h}^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) + g(\xi^h(\tau_h)) \\ & \geq \bar{V}(x) - 3\epsilon, \end{aligned}$$

where

$$\bar{V}(x) = \inf \left\{ \int_0^T k(\phi, \dot{\phi}) ds + V(\phi(T)) : \phi(0) = x, \phi(T) \in N_\gamma(x_0), T > 0 \right\}.$$

By the dynamic programming principle of optimality, $\bar{V}(x)$ is bounded below by $V(x)$. For the last case of $\bar{\tau} = \tau = \rho = \infty$, Lemma 3.4 gives

$$\liminf_{h \rightarrow 0} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^k} k^{\eta(h)}(\xi^h(s), \alpha) m^h(d\alpha ds) = \infty.$$

Combining the four cases and using an argument by contradiction gives

$$\liminf_{h \rightarrow 0} V^h(x) \geq V(x) - 3\epsilon.$$

The proof of the theorem is completed by sending $\epsilon \rightarrow 0$. ■

The proof of Theorem 3.3 can easily be extended. One such extension uses the following generalization of (A3.1).

A3.2. *There exist a finite collection of disjoint compact sets $\{K_i, i = 1, \dots, N\}$ and an open set G_1 containing G such that $\cup_{i=1}^N K_i \subset G^0$. Furthermore, given any $\gamma > 0$, there exists $T < \infty$ such that for all $t \geq T$,*

$$x(0) \in G^0, \quad \dot{x}(s) = \int_{\mathbb{R}^k} \alpha m_s(d\alpha) \text{ (a.s.)}$$

and

$$\int_0^t \int_{\mathbb{R}^k} k(x(s), \alpha) m(ds d\alpha) = 0$$

imply

$$x(t) \in \left(\bigcup_{i=1}^N N_\gamma(K_i) \right) \cup (\mathbb{R}^k - G_1).$$

Remark. Although the assumption has a somewhat complicated statement it is not very strong. The assumption is similar to and motivated by [48, assumption (A), p. 169].

We will also need the following extension of Lemma 3.4. The proof of this lemma is given at the end of the section.

Lemma 3.5. *Assume (A2.2) and (A3.2). Then given any $\gamma > 0$ and $M < \infty$, there is $T < \infty$ such that if $\dot{x}(s) = \int_{\mathbb{R}^k} \alpha m_s(d\alpha)$ (a.s.) for an admissible relaxed control $m(\cdot)$, and if $x(t) \in G - (\bigcup_{i=1}^N N_\gamma(K_i))$ for $t \in [0, T]$, then $\int_0^T \int_{\mathbb{R}^k} k(x(s), \alpha) m(d\alpha ds) \geq M$.*

Theorem 3.6. *Assume (A2.1), (A2.2), (A2.3'), (A3.2), and that $V(\cdot)$ is constant on each K_i . Then for the scheme defined by (3.6) and (3.3), we have*

$$\liminf_{h \rightarrow 0} V^h(x) \geq V(x).$$

Outline of the Proof. The proof is essentially the same as that of Theorem 3.3 but requires a good deal more notation. The constructions that would be used are very similar to those that are used in the proof of Theorem 3.8 below. An outline of the proof is as follows. Suppose that the quantities $\xi^h(\cdot), m^h(\cdot), \tau_h, \rho_h$, and so on, are defined as in Theorem 3.3. Assume the Skorokhod representation is used and that a convergent subsequence is under consideration. If the limit of $\rho_h \wedge \tau_h$ is finite, we are dealing with a finite time problem (at least as far as the convergence of the costs is concerned). We therefore assume that this is not the case. By suitably defining random times, we can keep track of the excursions of the sample paths taken between visits to small neighborhoods of the K_i and ∂G . Lemma 3.5 guarantees that if the running cost is to be bounded, then the interpolated time that the process is outside a neighborhood of the K_i and up until it exits G^0 is bounded. Thus, we can apply Theorem 2.4 as it was applied in Theorem 3.3 and get a lower bound on the running cost accumulated during any excursion between neighborhoods of the K_i (say from K_i to K_j) of the form

$$\inf \left\{ \int_0^T k(\phi, \dot{\phi}) ds, \phi(0) \in N_\gamma(K_i), \phi(T) \in N_\gamma(K_j), T > 0 \right\} - \epsilon.$$

A similar statement holds regarding excursions between a neighborhood of one of the K_i 's and a neighborhood of ∂G . Because the process $\xi^h(\cdot)$ eventually must either exit G^0 or be stopped at some finite time ρ_h , a dynamic programming argument gives the lower bound for points in $\bigcup_{i=1}^N N_\gamma(K_i)$.

For the points in G^0 that are not in $\cup_{i=1}^N N_\gamma(K_i)$, we use the same argument as in Theorem 3.3. By Lemma 3.5, the process $\xi^h(\cdot)$ must either exit G^0 or enter a small neighborhood of $\cup_{i=1}^N K_i$ in a uniformly bounded time, or else accumulate an infinite running cost (in the limit $h \rightarrow 0$). Thus, we again have the lower bound via a dynamic programming argument. ■

Proof of Lemma 3.5. The proof is very close to that of Lemma 2.2 in [48]. We first claim there exists $T < \infty$ and $c > 0$ such that $x(t) \in G - (\cup_{n=1}^N N_\gamma(K_n))$ for $t \in [0, T]$ implies $\int_0^T \int_{\mathbb{R}^k} k(x(s), \alpha) m(d\alpha ds) \geq c$. If not, we can find $T_i \rightarrow \infty$, $x_i(0) \in G - \cup_{i=1}^N N_\gamma(K_i)$, and $m_i(\cdot) \in \mathcal{R}([0, T_i] \times \mathbb{R}^k)$, such that for $x_i(\cdot)$ defined by

$$x_i(t) - x_i(0) = \int_0^t \int_{\mathbb{R}^k} \alpha m_i(d\alpha ds)$$

we have $x_i(t) \in G - (\cup_{n=1}^N N_\gamma(K_n))$ for all $t \in [0, T_i]$ and

$$\int_0^{T_i} \int_{\mathbb{R}^k} k(x_i(s), \alpha) m_i(d\alpha ds) \rightarrow 0.$$

We can choose a convergent subsequence such that

$$(x_i(0), x_i(\cdot), m_i(\cdot), T_i) \rightarrow (x(0), x(\cdot), m(\cdot), \infty),$$

where $m(\cdot) \in \mathcal{R}([0, \infty) \times \mathbb{R}^k)$,

$$x(t) - x(0) = \int_0^t \int_{\mathbb{R}^k} \alpha m(d\alpha ds),$$

and $x(t) \in G - (\cup_{n=1}^N N_\gamma(K_n))$ for all $t < \infty$. By Fatou's lemma we have

$$\int_0^\infty \int_{\mathbb{R}^k} k(x(s), \alpha) m(d\alpha ds) = 0,$$

which contradicts (A3.2). Thus, there are $T_1 < \infty$ and $c > 0$ satisfying the conditions of the claim. The conclusion of the lemma now follows if we take $T = T_1 M/c$. ■

13.3.3 A SHAPE FROM SHADING EXAMPLE

In this section we will consider numerical schemes for a shape from shading problem. For the sake of brevity we will consider only one particular setup and refer the reader to [39, 98] for generalizations. We also refer the reader to [58] for questions regarding terminology and modelling. In our approach to the shape-from-shading problem, we will use a calculus of variations problem whose minimal cost function describes the surface to be reconstructed. A related approach appears in [105], although the assumptions are quite different from those given here.

The formulation of our particular shape from shading problem is as follows. Consider a surface in \mathbb{R}^3 given in the explicit form $S = \{(x_1, x_2, x_3) : x_3 = f(x_1, x_2)\}$. The function $f(\cdot)$ appearing in the description of S is called the *height function*, and we will write x for (x_1, x_2) . Suppose that the surface is illuminated from the positive x_3 direction by a point light source that is assumed infinitely far away, and that the reflected light is recorded in an imaging plane that is parallel to the plane $\{(x_1, x_2, x_3) : x_3 = 0\}$. Assume that the “recorded light” is characterized in terms of a deterministic *intensity function* $I(x)$, where x identifies the (x_1, x_2) coordinates of a point on the imaging plane. Under a number of additional assumptions, including the assumption that the surface is “Lambertian,” the surface S and the intensity function $I(\cdot)$ are related by the equation

$$I(x) = \left(1 + |f_x(x)|^2\right)^{-1/2} \quad (3.15)$$

in regions where $f(\cdot)$ is continuously differentiable. Thus, $I(x)$ equals the absolute value of the x_3 -component of any unit vector orthogonal to S at $(x, f(x))$. We also have $I(x) \in (0, 1]$. We will refer the points where $I(x) = 1$, which obviously includes all local maximum and minimum points of $f(\cdot)$, as the *singular points* of $I(\cdot)$.

In the case of a single local minimum point, $I(\cdot)$ determines the surface up to a vertical translation. However, in the general case, (3.15) does not determine the function $f(\cdot)$ even with this sense of uniqueness. The assumption we will use is that the height function is known at all local minima. Because the minimum points are distinctive, it is likely that their heights and the local nature of the surface could be determined by some other method (e.g., stereo). An alternative to a priori knowledge of the heights of the local minima that is currently under investigation is based on reconstructing the surface with given (possibly incorrect) values assigned to the heights of the minima and then estimating the correct relative difference between the heights based on this reconstruction. We will also assume that an upper bound B is available for $f(\cdot)$ on G . The set G is given a priori, and represents the subset of the imaging plane on which data is recorded. G is often larger than the domain on which the reconstruction of $f(\cdot)$ is desired.

A3.3. Let $H \subset G$ be a compact set that is the closure of its interior, and assume H is of the form $H = \cap_{j=1}^J H_j$, $J < \infty$, where each H_j has a continuously differentiable boundary. Let L be the set of local minima of $f(\cdot)$ inside H , and assume that $I(\cdot)$ is continuous on H . Define $n_j(x)$ to be the inward (with respect to H) normal to ∂H_j at x . Assume that the value of $f(\cdot)$ is known at all points in L , and that $f_x(x)'n_j(x) < 0$ for all $x \in \partial H \cap \partial H_j$, $j = 1, \dots, J$.

Remarks. It turns out that the minimizing trajectories for the calculus of variations problem to be given in Theorem 3.7 below are essentially the

two dimensional projections of the paths of steepest descent on the surface represented by the height function. Thus, the assumptions that are placed on H in (A3.3) guarantee that any minimizing trajectory that starts in H stays in H . Theorem 3.7 shows that the height function has a representation as the minimal cost function of a calculus of variations problem that is correct for all points in the union of all sets H satisfying (A3.3). If we consider an initial point $x \in G$ such that the minimizing trajectory exits G , then we cannot construct $f(\cdot)$ at x by using the calculus of variations representation because this reconstruction would require $I(x)$ for values of x outside G . If we assume that the height function is specified at the local maximum points, then we can consider an analogous calculus of variations problem with a maximization.

The following theorem is proved in [98].

Theorem 3.7. *Assume (A3.3) and that $f(\cdot)$ is continuously differentiable. For $x \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}^2$, define*

$$g(x) = \begin{cases} f(x) & \text{for } x \in L \\ B & \text{for } x \notin L \end{cases}$$

and

$$k(x, \alpha) = \frac{1}{2}|\alpha|^2 + \frac{1}{2} \left(\frac{1}{I(x)^2} - 1 \right) = \frac{1}{2}|\alpha|^2 + \frac{1}{2}|f_x(x)|^2.$$

Define

$$V(x) = \inf \left[\int_0^{\rho \wedge \tau} k(\phi(s), \dot{\phi}(s)) ds + g(\phi(\rho \wedge \tau)) \right],$$

where $\tau = \inf\{t : \phi(t) \in \partial G \cup L\}$ and the infimum is over all $\rho > 0$ and absolutely continuous functions $\phi : [0, \rho] \rightarrow G$ that satisfy $\phi(0) = x$. Then

$$V(x) = f(x)$$

for all $x \in H$.

Remark. If $g(x)$ is set to a value that is less than $f(x)$ for some $x \in H$, then, in general, we will not have $V(x) = f(x)$ for $x \in H$. For example, if y is any point at which $I(y) = 1$ and if $g(y) < f(y)$, then $V(y) \leq g(y) < f(y)$.

We next present a numerical procedure for solving for $V(x)$. One feature of this problem that is quite different from those considered previously is the nature of the target set. For example, consider the case when $L = \{x_0\}$ for some $x_0 \in G^0$. The target set is then $\partial G \cup \{x_0\}$, and (A2.1) does not apply. The interior cone condition holds, but the exterior cone condition fails. The exterior cone condition was used in the proof of the upper bound for all convergence theorems that have been presented so far in this chapter. In those proofs, if an optimal (or ϵ -optimal) path $\phi(\cdot)$ terminated on a target set ∂G at time ρ , then the exterior cone condition was used to define $\phi(\cdot)$

on a small interval $(\rho, \rho + \nu], \nu > 0$, in such a way that the added cost was arbitrarily small and $\phi(t) \notin G$ for $t \in (\rho, \rho + \nu]$. This $\phi(\cdot)$ was then used to define a control scheme for the chain, and because $\phi(\cdot)$ had been constructed in this way, the exit times of the chain converged to ρ . See, for example, Subsection 13.2.3. If the target set does not satisfy an exterior cone condition, then this construction is no longer possible. Target sets such as an isolated point are typically difficult to deal with when proving the convergence of numerical schemes. A common technique is to replace the target set A by $N_\gamma(A), \gamma > 0$, prove an appropriate convergence property for the problem with this target set, and then send $\gamma \rightarrow 0$. We will show in this subsection that this “fattening” of the target set is not needed when a mild additional condition on the chain is assumed to hold.

Let $V_T(x)$ denote the optimal cost if the controlled stopping time is restricted to the range $[0, T]$. Our assumption that B is an upper bound for $f(\cdot)$ on G implies that it is never optimal to stop at a point in $G - L$ whenever T is sufficiently large. The stopping cost for points in $G - L$ was actually introduced in the definition of the calculus of variations problem of Theorem 3.7 solely for the purpose of forcing optimal trajectories to terminate in L . This use of a stopping cost could be avoided altogether if the minimization in the calculus of variations problem were only over paths that terminate in L at some finite time. However, this added constraint would be rather difficult to implement in the numerical approximations. We will see below that because $g(\cdot)$ is the proper stopping cost to introduce to force the trajectories to terminate in L , it also provides the proper initial condition for the numerical scheme.

Because the target set can possibly contain isolated points that may not be included in G_h , we really need to introduce a “discretized target set” $L_h \subset G_h$, and redefine $g(\cdot)$ in the obvious way. We would need that $L_h \rightarrow L$ in the Hausdorff metric [i.e., $d(x, L) \leq \epsilon_h$ for all $x \in L_h$, $d(x, L_h) \leq \epsilon_h$ for all $x \in L$, and $\epsilon_h \rightarrow 0$]. To simplify the notation we will just assume $L \subset G_h$.

Unlike the general situation of Subsection 13.3.2, for this problem we can approximate the finite time problem $V_T(x)$ and then send $T \rightarrow \infty$ and the discretization parameters to their limits in any way we choose. This allows the use of the cost $k(\cdot, \cdot)$, rather than $k^{\eta(h)}(\cdot, \cdot)$, in the following alternative to the method of Subsection 13.3.2.

A Numerical Scheme. Define $V_0^h(x) = g(x)$, and recursively define

$$V_{n+1}^h(x) = \min \left[g(x), \min_{\alpha \in \mathbb{R}^2} \left[\sum_y p^h(x, y | \alpha) V_n^h(y) + k(x, \alpha) \Delta t^h(\alpha) \right] \right] \quad (3.16)$$

if $x \in G_h^0$ and $V_{n+1}^h(x) = g(x)$ for $x \notin G_h^0$. Finally, define

$$V^h(x) = \lim_{n \rightarrow \infty} V_n^h(x). \quad (3.17)$$

The existence of the limit in the definition of $V^h(x)$ follows from the monotonicity $V_{n+1}^h(x) \leq V_n^h(x)$.

Remark. The iteration (3.16) is of “Jacobi” type (see Chapter 6). Convergence can also be demonstrated when (3.16) is replaced by a “Gauss-Seidel” type of iteration [39].

Remarks on the Implementation of the Algorithms. It is appropriate at this point to make a few remarks concerning the implementation of the algorithms described in this section. We have delayed our remarks until this point because it is useful to have an illustrative example at hand. The remarks are given for the approximation in value space method of solving equations such as (3.1). The particular examples that have been programmed by the authors have all involved cost functions for which the minimization in (3.1) could be done analytically. In all cases, the approximation in value space method was used (see Chapter 6). The approximation in policy space method was not used because the approximation in value space method worked well, and there seemed no need to use a more sophisticated method. The remarks are given in the context of the Jacobi iteration, but similar remarks apply for Gauss-Seidel type methods.

The most distinctive feature of the problems considered in this and the previous subsection is the lack of a positive lower bound on the running cost. One consequence of this feature has been that numerical schemes needed to be chosen carefully, and their convergence proofs have been more involved. This feature also has strong implications with regard to the actual implementation of the algorithms. In particular, the iterative schemes that we use to calculate our approximate solutions are either not contractions [as in (3.16) of this subsection], or else they what might be called “weak” contractions, whose contraction property rapidly disappears as $h \rightarrow 0$ [as in (3.3) of Subsection 13.3.2].

We begin by discussing the particular problem of this subsection. This problem is especially illuminating, since it forces us to deal with the lack of a contraction property right from the start. Consider the calculus of variations problem that uses the same running cost as in Theorem 3.7 but infimizes only over those paths that stay in G and reach L in finite time. This problem as stated is not exactly of the form considered in Subsection 13.1.2 because the stopping time is not controlled and since we infimize only over the paths which enter L in finite time. If the running cost were actually bounded from below away from zero, then the constraint that the trajectories enter L in finite time would obviously be satisfied by any trajectory with finite running cost. Thus, the constraint would not need to be explicitly accounted for in the numerical algorithm. However, without such a bound from below there can exist paths with zero running cost that stay inside $G - L$ for all $t > 0$. Hence, the constraint must be *explicitly* incorporated into any numerical algorithm. The constraint is actually difficult

to implement directly, and we have avoided this problem by introducing a controlled stopping time (as in Subsection 13.1.2) with a large stopping cost in $G - L$.

When this modified problem is implemented using the scheme (3.16), (3.17), we require that the stopping cost for the modified problem be the *initial condition* of the numerical algorithm. This is quite natural, since with such an initial condition $V_i^h(x)$ can be interpreted as the cost function of a controlled discrete time Markov chain that starts at x at time 0 and has transition probabilities $p^h(x, y|\alpha)$, running cost $k(x, \alpha)\Delta t^h(\alpha)$, stopping set ∂G_h^0 , stopping cost B , and terminal cost (assigned at time i if the chain has not yet entered ∂G_h^0 or L) of $g(\cdot)$. This interpretation is essential in the proof of convergence of the scheme. Suppose that one were to use (3.1) as the basis for an approximate solution. In particular, suppose that the approximation in value space method is used to compute a solution to (3.1). We would then end up with the algorithm (3.16), (3.17), and as will be proved below the “correct” solution to (3.1) is obtained only if we start with a sufficiently large initial condition. The lack of a contraction property means that our numerical approximation depends on the initial condition used in the algorithm. For example, if the initial condition $V_0^h(x) = 0$ were used in (3.16) for our example, then $V_i^h(x) = 0$ for all i at points where the running cost is zero. Owing to the form of (3.16), it is clear that any initial condition that is sufficiently large would work well.

We now discuss implementation of the algorithms of Subsections 13.3.1 and 13.3.2. We first focus on the approximation in value space method for solving (3.1) and (3.3). For the problem of this subsection, we have seen that if one were to attempt to use (3.1) to approximate the solution, and if the approximation in value space method were used to compute a solution to (3.1), then the correct solution to (3.1) is found only when we start with the proper sufficiently large initial condition. For the problems of Subsection 13.3.1, the lower bound on the running cost guarantees that the limit of any of the iterative algorithms that one would use to solve (3.1) is independent of the initial condition. In spite of this fact, the speed of convergence depends heavily on the initial condition, and, in particular, the algorithm can converge much more slowly if too small an initial condition is used. For the problems of Subsection 13.3.2 it is almost essential that a good (i.e., large) initial condition be used. This phenomenon has the following intuitive interpretation. For simplicity in discussing this point, let us consider the problem without a controlled stopping time, so that (3.1) is replaced by

$$V^h(x) = \min_{\alpha \in \mathbb{R}^k} \left[\sum_y p^h(x, y|\alpha) V^h(y) + k(x, \alpha) \Delta t^h(\alpha) \right] \quad (3.18)$$

if $x \in G_h^0$, $V^h(x) = g(x)$ if $x \notin G_h^0$, and $k(x, \alpha) \geq k_0$. Suppose one were to use the approximation in value space method to solve (3.18). Let the

iterates be denoted by $V_i^h(x)$, and for some given function $h(\cdot)$, let the initial condition be $V_0^h(x) = h(x)$. Then $V_i^h(x)$ can be interpreted as the cost function of a controlled discrete time Markov chain that starts at x at time 0 and has transition probabilities $p^h(x, y|\alpha)$, running cost $k(x, \alpha)\Delta t^h(\alpha)$, stopping set ∂G_h^0 , stopping cost $g(\cdot)$, and terminal cost (assigned at time i if the chain has not yet entered ∂G_h^0) of $h(\cdot)$.

For the calculus of variations problems of Subsections 13.3.1 and 13.3.2, the only given data are the values of $V(x)$ on ∂G , which must be propagated back into G^0 along the minimizing paths. The situation is similar with the discrete approximation $V_i^h(x)$. The only given data are the values of $V_i^h(x)$ at the points $x \in \partial G_h$. This information is “passed along” to the points $x \in G_h^0$ in the iterative scheme used to solve (3.18) by the controlled Markov chain. The information is passed along most quickly if the chain [while evolving under the optimal policy for which $V^h(x)$ is the value] reaches ∂G_h most quickly.

Now consider how the terminal cost $h(\cdot)$ (which is the initial condition for the algorithm) can affect the time it takes the chain to reach ∂G_h . If $h(\cdot)$ is large, then the optimal policy will most likely move the chain to the boundary in order to avoid the large terminal cost. If $h(\cdot)$ is low, then for small i there may be little advantage for the optimal policy to move the chain to the boundary. Regardless of the terminal cost for the chain (initial condition for the algorithm), the strict positivity of $k(\cdot, \cdot)$ implies the sum of the running costs will build up (as $i \rightarrow \infty$), eventually forcing the chain to enter the boundary set. This corresponds to the fact that we do in fact achieve convergence with any initial condition for the iterative algorithm. However, this is clearly inefficient, especially when k_0 is small. One can consider the extreme case of a point $x \in G_h^0$ that will eventually “learn” its value $V^h(x)$ from a single neighboring point $y \in \partial G_h$. If $h(x)$ is small compared with $g(y)$, then the approximation in value space method must be iterated until the accumulated running costs drive $V_i^h(x)$ near $g(y)$. It is usually better to use $h(x) > g(y)$, in which case the correct value at x is “learned” after the first iteration.

It has been the authors’ experience that the convergence of these iterative schemes to the solution of (3.2) is fast, at least if the initial condition is sufficiently large, especially when using the Gauss-Seidel algorithms and alternating the ordering of the states for each iteration. This is probably due to the deterministic nature of the controlled process $\dot{\phi}(t) = u(t)$. While the approximating processes $\xi^h(\cdot)$ are not deterministic, the fact that they are “nearly” deterministic means that information contained in the boundary or stopping cost is quickly passed to the interior points. This is especially true when transition probabilities such as those of Example 2.1 are used, which guarantee that the process moves away from its current grid point at each iteration.

A Simple Recursive Equation. We now return to the shape from shad-

ing example. It is frequently the case that the right hand side of (3.16) can be evaluated analytically, which yields a simple recursive formula for $V_n^h(x)$. As an example, we can use the transition probabilities of Example 2.1. For simplicity, we will omit the intermediate calculations and present only the conclusions. For $x \in G_h^0$, let v_1 and v_2 be the smallest values from the sets

$$\{V_n^h(x + h(1, 0)), V_n^h(x - h(1, 0))\} \text{ and } \{V_n^h(x + h(0, 1)), V_n^h(x - h(0, 1))\},$$

respectively. Define $m = (1/I^2(x)) - 1$. If $0 \leq h^2m < (v_1 - v_2)^2$, then we use the recursion

$$V_{n+1}^h(x) = \min \left[g(x), (v_1 \wedge v_2) + hm^{1/2} \right].$$

If $h^2m \geq (v_1 - v_2)^2$, then we use

$$V_{n+1}^h(x) = \min \left[g(x), \frac{1}{2} \left[(2h^2m - (v_1 - v_2)^2)^{1/2} + (v_1 + v_2) \right] \right].$$

On Weakening (A2.1). As mentioned previously, the target set need not satisfy an exterior cone condition. This means that in order to prove convergence, we need an additional assumption beyond local consistency. The additional condition is needed only in the proof of the upper bound $\limsup_{h \rightarrow 0} V^h(x) \leq V(x)$. Basically, all that is needed is a “controllability” condition on the Markov chain. Consider the particular case where the target set is a single point. The idea is to impose conditions on the chain which will guarantee that if we can get the chain “close” to the target set by a control scheme with nearly optimal running cost, then a control with small running cost can be found that will finish the job of driving the chain into the target set (at least with a high probability).

Consider the transition probabilities and interpolation interval of Example 2.1. Let $\epsilon > 0$ be given and assume (A2.2). Then there exists $h_0 > 0$ and $\gamma > 0$ with the following properties. For all $h < h_0$ and for any pair of points $x \in G_h$, $y \in G_h$ satisfying $|x - y| < \gamma$, there exists a nonanticipative control scheme $\{u_i^h, i < \infty\}$ with the following properties. If $\{\xi_i^h, i < \infty\}$ is the controlled chain that starts at x at time zero and if $N_h = \inf\{i : \xi_i^h = y\}$, then

1. $|\xi_i^h - y| \leq \epsilon$ for all $i = 0, \dots, N_h - 1$,
2. $\sum_{i=0}^{N_h-1} \Delta t^h(u_i^h) \leq \epsilon$,
3. $\sum_{i=0}^{N_h-1} k(\xi_i^h, u_i^h) \Delta t^h(u_i^h) \leq \epsilon$

w.p.1. Hence, if the proof of the upper bound can be given when the target set is of the form $N_\gamma(x_0)$ (i.e., a set that satisfies the exterior cone

condition), then it can be given for target sets of the form $\{x_0\}$ as well. The details will be given below. A formulation of conditions on the Markov chain that includes Example 2.1 and the chains derived from this example via (2.3) and (2.4) is as follows.

A3.4. *Given $\epsilon > 0$ there exist $h^0 > 0$, $\gamma > 0$, and $M < \infty$ with the following properties. Given any $h < h^0$ and any $x, y \in G_h$ such that $|x - y| < \gamma$, there exists a nonanticipative control scheme $\{u_i^h, i < \infty\}$ satisfying $|u_i^h| \leq M, i < \infty$, with the following properties. If $\{\xi_i^h, i < \infty\}$ is the resulting controlled chain that starts at x and if $N_h = \inf\{i : \xi_i^h = y\}$, then*

$$P \left\{ \sup_{0 \leq i \leq N_h - 1} |\xi_i^h - y| \geq \epsilon \text{ or } \sum_0^{N_h - 1} \Delta t^h(u_i^h) \geq \epsilon \right\} < \epsilon.$$

Our convergence theorem can finally be stated.

Theorem 3.8. *Assume (A3.3) and (A3.4), and that $K = \{x : I(x) = 1\}$ consists of a finite number of disjoint connected compact sets. Define $V^h(x)$ by (3.16) and (3.17), where the running cost and terminal cost are as given in Theorem 3.7. Then*

$$V^h(x) \rightarrow V(x).$$

Remark. Recall that L was defined in (A3.3) as the set of local minima inside H . The points $\{x : I(x) = 1, x \in G - L\}$ are the singular points that are not local minima inside H . On these points we have $k(x, 0) = 0$, and thus the assumptions used in Subsection 13.3.1 do not apply.

Proof of the Upper Bound. The proof of the upper bound follows the lines of the finite time problem of Theorem 2.5, except for the difficulties related to the nature of the target set. Fix $x \in G - L$. If $V(x) = B$, there is nothing to prove. Assume $V(x) < B$, and let $\epsilon > 0$ be given such that $V(x) + 2\epsilon < B$. Recall that $V_T(x)$ is the minimal cost subject to the restriction $\rho \in [0, T]$. Since $V_T(x) \downarrow V(x)$, there exists $T < \infty$ such that $V_T(x) \leq V(x) + \epsilon < B$. If $m(\cdot)$ is an ϵ -optimal relaxed control for $V_T(x)$ and if $x(\cdot)$ is the associated solution, then $x(\tau) \in L$ for some $\tau \leq T$, and

$$\int_0^\tau \int_{\mathbb{R}^2} k(x(s), \alpha) m(d\alpha ds) + g(x(\tau)) \leq V_T(x) + \epsilon.$$

[Note that it is not optimal to stop with this given control before $x(\cdot)$ enters L .]

Select $\gamma > 0$ for the given ϵ according to (A3.4). By Theorem 2.2, we may assume the existence a finite set $\mathcal{U}_{\gamma/2} \subset \mathbb{R}^k$, $\bar{\delta} > 0$, and an ordinary control $u^{\gamma/2}(\cdot)$ with the following properties. $u^{\gamma/2}(\cdot)$ takes values in $\mathcal{U}_{\gamma/2}$, is constant on intervals of the form $[j\bar{\delta}, j\bar{\delta} + \bar{\delta})$, and if $x^{\gamma/2}(\cdot)$ is the associated solution, then

$$\sup_{0 \leq t \leq \tau} |x^{\gamma/2}(t) - x(t)| \leq \gamma/2$$

and

$$\sup_{0 \leq t \leq \tau} \left| \int_0^t \int_{\mathbb{R}^2} k(x(s), \alpha) m_s(d\alpha) ds - \int_0^t k(x^{\gamma/2}(s), u^{\gamma/2}(s)) ds \right| \leq \epsilon.$$

We now define a control scheme for the Markov chain. We will use $u^{\gamma/2}(\cdot)$ to define the scheme until the interpolated time reaches τ . If the chain has not yet entered the target set by this time, then we may have to extend the definition for times after τ via (A3.4).

In order to apply $u^{\gamma/2}(\cdot)$ to the chain $\{\xi_i^h, i < \infty\}$, we recursively define the control applied at discrete time i by $u_i^h = u^{\gamma/2}(t_i^h)$ and $t_{i+1}^h = t_i^h + \Delta t^h(u_i^h)$. This defines a control until i such that $t_{i+1}^h \geq \tau$. Let $\{\xi_i^h, i < \infty\}$ be the chain that starts at x and uses this control.

Define

$$N_h = \inf\{i : t_i^h \geq \tau \text{ or } \xi_i^h \in L \text{ or } \xi_i^h \notin G^0\},$$

and let $\tau_h = t_{N_h}^h$. By Theorem 2.4, we have $\sup_{0 \leq t \leq \tau_h} |\xi^h(t) - x^{\gamma/2}(t)| \rightarrow 0$ in probability, and $P\{|\xi_{N_h}^h - x(\tau)| \geq \gamma\} \rightarrow 0$. Assume $h > 0$ is small enough that $P\{|\xi_{N_h}^h - x(\tau)| \geq \gamma\} \leq \epsilon$.

On the set where $|\xi_{N_h}^h - x(\tau)| < \gamma$, we extend the definition of the control sequence for discrete times larger than N_h in such a way that (A3.4) is satisfied. We then stop the process at the discrete time

$$M_h = \inf\{i \geq N_h : t_i^h - t_{N_h}^h \geq \epsilon \text{ or } \xi_i^h \in L \text{ or } \xi_i^h \notin G^0\}.$$

Note that if $\xi_{M_h}^h \in L$, then $|\xi_{M_h}^h - x(\tau)| \leq \epsilon$ by (A3.4), at least with probability close to one. Recall that $g(x) = f(x)$ for $x \in L$. On the set where $|\xi_{N_h}^h - x(\tau)| \geq \gamma$, we stop at N_h and pay the stopping cost. The total cost is then bounded above by

$$\begin{aligned} & E_x^{m^h} \int_0^{\tau_h} \int_{\mathbb{R}^2} k(\xi^h(s), \alpha) m^h(d\alpha ds) \\ & + P\{|\xi_{N_h}^h - x(\tau)| \geq \gamma\} B + \epsilon \sup_{y \in G, |\alpha| \leq M} k(y, \alpha) \\ & + P\{|\xi_{N_h}^h - x(\tau)| < \gamma, \xi_{M_h}^h \in L, \text{ and } |\xi_{M_h}^h - x(\tau)| \leq \epsilon\} \sup_{|z - x(\tau)| \leq \epsilon} f(z) \\ & + P\{|\xi_{N_h}^h - x(\tau)| < \gamma, \xi_{M_h}^h \notin L, \text{ or } |\xi_{M_h}^h - x(\tau)| > \epsilon\} B. \end{aligned}$$

This last sum is itself bounded above by

$$\begin{aligned} & E_x^{m^h} \int_0^{\tau_h} \int_{\mathbb{R}^2} k(\xi^h(s), \alpha) m^h(d\alpha ds) + \epsilon \sup_{y \in G, |\alpha| \leq M} k(y, \alpha) \\ & + \sup_{|z - x(\tau)| \leq \epsilon} f(z) + 2\epsilon B. \end{aligned}$$

Sending $h \rightarrow 0$, we obtain

$$\begin{aligned} \limsup_{h \rightarrow 0} V^h(x) &\leq \int_0^\tau \int_{\mathbb{R}^2} k(x(s), u^{\gamma/2}(s)) ds \\ &+ \epsilon \sup_{y \in G, |\alpha| \leq M} k(y, \alpha) + \sup_{|z - x(\tau)| \leq \epsilon} f(z) + 2\epsilon B. \end{aligned}$$

Sending $\epsilon \rightarrow 0$ gives

$$\limsup_{h \rightarrow 0} V^h(x) \leq V_T(x).$$

Since $T < \infty$ is arbitrary,

$$\limsup_{h \rightarrow 0} V^h(x) \leq V(x).$$

Proof of the Lower Bound. Fix $x \in G - L$ and $\epsilon > 0$. According to (3.16) and (3.17), for each $h > 0$ there is $n < \infty$ such that

$$V^h(x) \geq V_n^h(x) - \epsilon. \quad (3.19)$$

For the rest of the proof we will assume that n has been chosen such that (3.19) holds. Owing to the definition of $V_n^h(x)$, there is a controlled Markov chain $\{\xi_i^h, i < \infty\}$ with control sequence $\{u_i^h, i < \infty\}$ that satisfies $\xi_0^h = x$, and a finite stopping time N_h such that

$$V_n^h(x) \geq E_x^{u^h} \sum_{j=0}^{(N_h \wedge M_h)-1} k(\xi_j^h, u_j^h) \Delta t^h(u_j^h) + E_x^{u^h} g(\xi_{N_h \wedge M_h}^h) - \epsilon, \quad (3.20)$$

where M_h is the time of first exit from G^0 or entrance into the set L . The stopping time N_h is the minimum of n and the controlled stopping time. Although the chain, control sequence, and stopping times depend on n , we will not indicate the dependence in the notation. Let $\xi^h(\cdot)$ and $u^h(\cdot)$ be the continuous parameter interpolations of $\{\xi_i^h, i < \infty\}$ and $\{u_i^h, i < \infty\}$, respectively, and let

$$\rho_h = \sum_{i=0}^{N_h-1} \Delta t^h(u_i^h), \quad \tau_h = \sum_{i=0}^{M_h-1} \Delta t^h(u_i^h).$$

We can then rewrite (3.20) as

$$V_n^h(x) \geq E_x^{m^h} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^2} k(\xi^h(s), \alpha) m^h(d\alpha ds) + E_x^{m^h} g(\xi^h(\rho_h \wedge \tau_h)) - \epsilon, \quad (3.21)$$

where $m^h(\cdot)$ is the relaxed control representation of the ordinary control $u^h(\cdot)$.

Let $K_q, q = 1, \dots, Q$ be disjoint compact connected sets such that $K = \cup_{q=1}^Q K_q$. The existence of such a decomposition has been assumed in the statement of Theorem 3.8. Now $V(x)$ is constant on each K_q , so there exists $\gamma > 0$ such that

$$x \in K_q, y \in N_\gamma(K_q) \Rightarrow |V(x) - V(y)| \leq \epsilon \quad (3.22)$$

and such that the sets $N_\gamma(K_q)$ are separated by a distance greater than γ for distinct q . Because the reflected light intensity $I(\cdot)$ is continuous, there is $c > 0$ such that

$$k(x, \alpha) \geq c \text{ for all } \alpha \in \mathbb{R}^2, x \in G - \cup_{q=1}^Q N_{\gamma/2}(K_q). \quad (3.23)$$

For simplicity, we will consider the proof of the lower bound for the case when the initial condition satisfies $x \in N_{\gamma/2}(K_q)$ for some q . The general case follows easily using the same arguments. We define a sequence of stopping times by

$$\begin{aligned} \tau_0^h &= 0, \\ \sigma_j^h &= \inf\{t \geq \tau_j^h : \xi^h(t) \notin \cup_{q=1}^Q N_\gamma(K_q)\}, \\ \tau_j^h &= \inf\{t \geq \sigma_{j-1}^h : \xi^h(t) \in \cup_{q=1}^Q N_{\gamma/2}(K_q) \text{ or } \xi^h(t) \notin G^0\}. \end{aligned}$$

Consider the processes

$$\Xi^h(\cdot) = (\xi_0^h(\cdot), \xi_1^h(\cdot), \dots), \quad M^h(\cdot) = (m_0^h(\cdot), m_1^h(\cdot), \dots),$$

where $\xi_j^h(\cdot) = \xi^h(\cdot + \sigma_j^h)$ and where $m_j^h(\cdot)$ is the relaxed control representation of the ordinary control $u^h(\cdot + \sigma_j^h)$. We consider $(\Xi^h(\cdot), M^h(\cdot))$ as taking values in the space

$$(D^2[0, \infty) \times \mathcal{R}([0, \infty) \times \mathbb{R}^2))^{\mathbb{Z}},$$

endowed with the usual product space topology. Owing to its definition, $V_n^h(x)$ is uniformly bounded from above. Thus Theorem 2.4 shows that given any subsequence of $\{(\Xi^h(\cdot), M^h(\cdot)), h > 0\}$, we can extract a further subsequence that converges weakly, and that any limit point

$$(X(\cdot), M(\cdot)) = ((x_0(\cdot), x_1(\cdot), \dots), (m_0(\cdot), m_1(\cdot), \dots))$$

of such a convergent subsequence satisfies

$$x_j(t) - x_j(0) = \int_0^t \int_{\mathbb{R}^2} \alpha m_j(d\alpha ds),$$

where each $m_j(\cdot)$ is an admissible relaxed control. In addition, the definition of the stopping times $\{\sigma_j^h\}$ guarantees that $x_0(0) \in \partial N_\gamma(x)$ and that for all $j > 0$, either $x_j(0) \in \partial N_\gamma(K_q)$ for some q or $x_j(0) \notin G^0$.

Let $J^h = \min\{j : \tau_j^h \geq \tau_h\}$, where τ_h has been defined to be the interpolated time at which $\xi^h(\cdot)$ first exited G^0 or entered L . It then follows from the uniform bound from below given in (3.23) and (3.21) that

$$\limsup_{h \rightarrow 0} E_x^{m^h} \sum_{0 \leq j < J^h} (\tau_{j+1}^h - \sigma_j^h) < \infty. \quad (3.24)$$

Define $s_j^h = \tau_{j+1}^h - \sigma_j^h$ and $S^h = (s_0^h, s_1^h, \dots)$. It also follows from (3.23) that there exists $\bar{c} > 0$ such that for all q_1 and q_2 ,

$$\begin{aligned} & \inf \left\{ \int_0^T k(\phi, \dot{\phi}) ds : \phi(0) \in \partial N_\gamma(K_{q_1}), \phi(T) \in N_{\gamma/2}(K_{q_2}), T > 0 \right\} \\ & \geq \bar{c} \end{aligned} \quad (3.25)$$

[e.g. $\bar{c} = (2c)^{1/2}\gamma$].

We now prove the lower bound $\liminf_{h \rightarrow 0} V^h(x) \geq V(x)$. Extract a subsequence along which

$$(\xi^h(\cdot), m^h(\cdot), \rho_h, \tau_h, \Xi^h(\cdot), M^h(\cdot), J^h, S^h)$$

converges to a limit

$$(x(\cdot), m(\cdot), \rho, \tau, \Xi(\cdot), M(\cdot), J, S).$$

We assume via the Skorokhod representation that the convergence is w.p.1, and consider any ω for which there is convergence. If $\rho < \tau$, we have

$$\liminf_{h \rightarrow 0} g(\xi^h(\rho_h \wedge \tau_h)) \geq B.$$

Next assume $\tau \leq \rho$. We have

$$\begin{aligned} & \liminf_{h \rightarrow 0} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^2} k(\xi^h(s), \alpha) m^h(d\alpha ds) \\ & \geq \liminf_{h \rightarrow 0} \sum_{0 \leq j < J^h} \int_{\sigma_j^h}^{\tau_{j+1}^h} \int_{\mathbb{R}^2} k(\xi^h(s), \alpha) m^h(d\alpha ds) \\ & \geq \sum_{0 \leq j < J} \int_0^{s_j} \int_{\mathbb{R}^2} k(x_j(s), \alpha) m_j(d\alpha ds). \end{aligned}$$

By using the definition of $V(\cdot)$ and an elementary dynamic programming argument, for each $j < J$ we have

$$V(x_j(0)) \leq \int_{\mathbb{R}^2} k(x_j(s), \alpha) m_j(d\alpha ds) + V(x_j(s_j)). \quad (3.26)$$

By (3.24), the s_j for $j < J$ are finite w.p.1, and by (3.25) we can assume that J is uniformly bounded from above. By construction, if $j < J-1$ and if

$x_j(s_j) \in N_{\gamma/2}(K_q)$ (which it must be for some q), then $x_{j+1}(0) \in \partial N_\gamma(K_q)$. Recall that $L \subset \cup_{q=1}^Q K_q$ and that L is the set of local minimum points. It follows from the definitions of τ_j^h and σ_j^h that if $\tau_h < \rho_h$ and $\xi^h(\tau_h) \in K_q$ for some q , then $\xi^h(\tau_{J^h-1}) \in N_{\gamma/2}(K_q)$ for that same q . On the other hand, if $\tau_h < \rho_h$ and $\xi^h(\tau_h) \notin K_q$ for any q , then $g(\xi^h(\rho_h \wedge \tau_h)) = B$. Therefore, in general,

$$\liminf_{h \rightarrow 0} g(\xi^h(\rho_h \wedge \tau_h)) \geq g(x_{J-1}(s_{J-1})) - \epsilon \quad (3.27)$$

w.p.1. Now consider the paths $x_j(\cdot), j < J < \infty$. Clearly, $x_0(0) \in N_\gamma(x)$, and by (3.22) we have $|V(x_{j-1}(s_{j-1})) - V(x_j(0))| \leq 2\epsilon$. By combining this with (3.26) and (3.27), we obtain

$$\begin{aligned} & \liminf_{h \rightarrow 0} \int_0^{\rho_h \wedge \tau_h} \int_{\mathbb{R}^2} k(\xi^h(s), \alpha) m^h(d\alpha ds) + g(\xi^h(\rho_h \wedge \tau_h)) \\ & \geq \sum_{0 \leq j < J} \int_0^{s_j} \int_{\mathbb{R}^2} k(x_j(s), \alpha) m_j(d\alpha ds) + g(x_{J-1}(s_{J-1})) - \epsilon \\ & \geq \sum_{0 \leq j < J} [V(x_j(0)) - V(x_j(s_j))] + V(x_{J-1}(s_{J-1})) - \epsilon \\ & \geq V(x) + \sum_{0 \leq j < J} [V(x_j(0)) - V(x_{j-1}(s_{j-1}))] - 2\epsilon \\ & \geq V(x) - 2J\epsilon \end{aligned}$$

w.p.1. Using (3.19) and (3.20), we have

$$\liminf_{h \rightarrow 0} V^h(x) \geq V(x) - 2[J+1]\epsilon.$$

We conclude by sending $\epsilon \rightarrow 0$ and using the fact that J is uniformly bounded from above. ■

Computational Results and Examples. Figure 13.1 displays a surface that has been sampled at a grid of 128×128 points. Figure 13.2 shows the reconstruction provided by the Jacobi algorithm after 150 iterations. The maximum value of the height function is 20, and the reconstructed image is within 0.5 of the true image. In producing Figure 13.2, we have used an analytic expression for the surface to compute $I(\cdot)$ at the grid points. In order to estimate the number of iterations required for the algorithm itself to converge, we computed discrete approximations to the derivatives on the basis of the surface sampled only at the grid points, and then used these values to compute $I(\cdot)$ at the grid points. With these values of $I(\cdot)$, the reconstruction should be equal to the original surface modulo only errors due to machine accuracy. Thus, we are able to separate errors due to the numerical method from those due to stopping the algorithm (3.16) after too few iterations. Given these values for $I(\cdot)$ and starting with a large initial condition, the algorithm converged to a solution correct to within

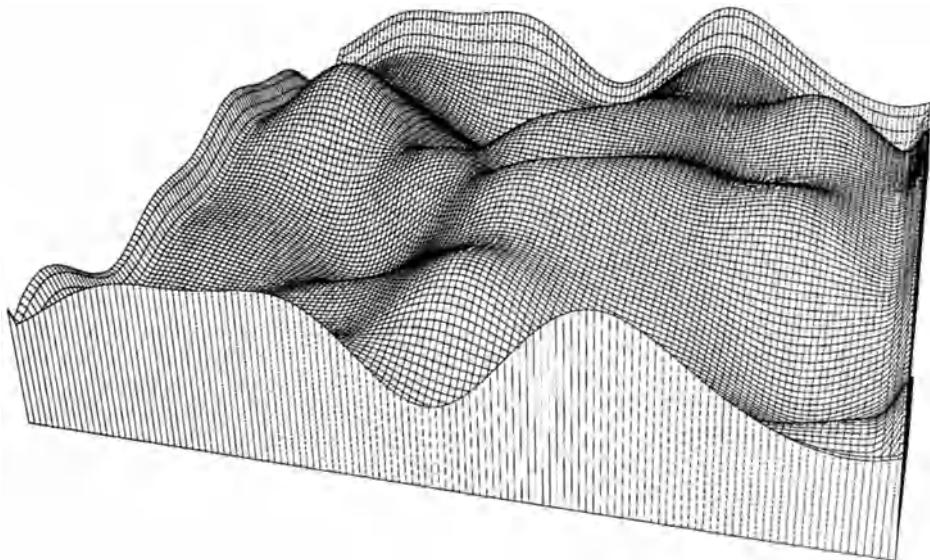


Figure 13.1. Original surface.

10^{-6} after 150 iterations of Jacobi type. By using the Gauss-Seidel version of the algorithm presented here and varying the ordering of the states with each iteration, convergence was observed after 11 iterations.

In Figures 13.3-13.5, we consider the application of the algorithm to a real 200 by 200 image. Figure 13.1 gives the original image, which is a picture of the head of a mannequin. The head was actually illuminated from an oblique direction, and the algorithm used in the reconstruction was the oblique direction version of the algorithm presented in this subsection [98]. For the reconstruction, the set L was taken to be the position of the tip of the nose, and the version of the algorithm that allows data to be given at local maxima was used. Figure 13.4 shows the surface reconstruction obtained after using 6 Gauss-Seidel iterations, and Figure 13.5 shows how this reconstruction would appear if illuminated by light coming from the same direction as in Figure 13.3. The Jacobi algorithm required 160 iterations to converge.

For further examples and detail as well as a comparison of the algorithm with other proposed schemes, we refer the reader to [39, 98].

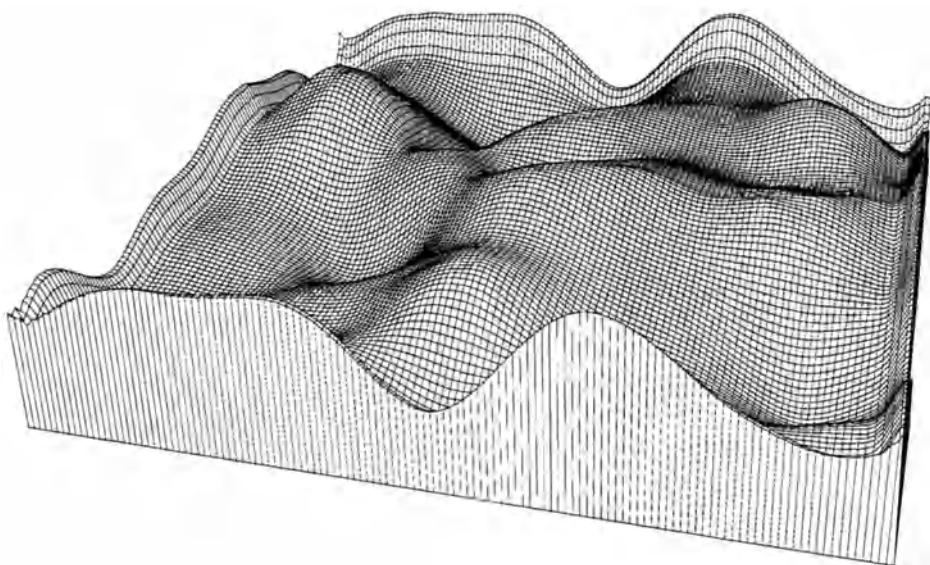


Figure 13.2. Reconstructed surface.

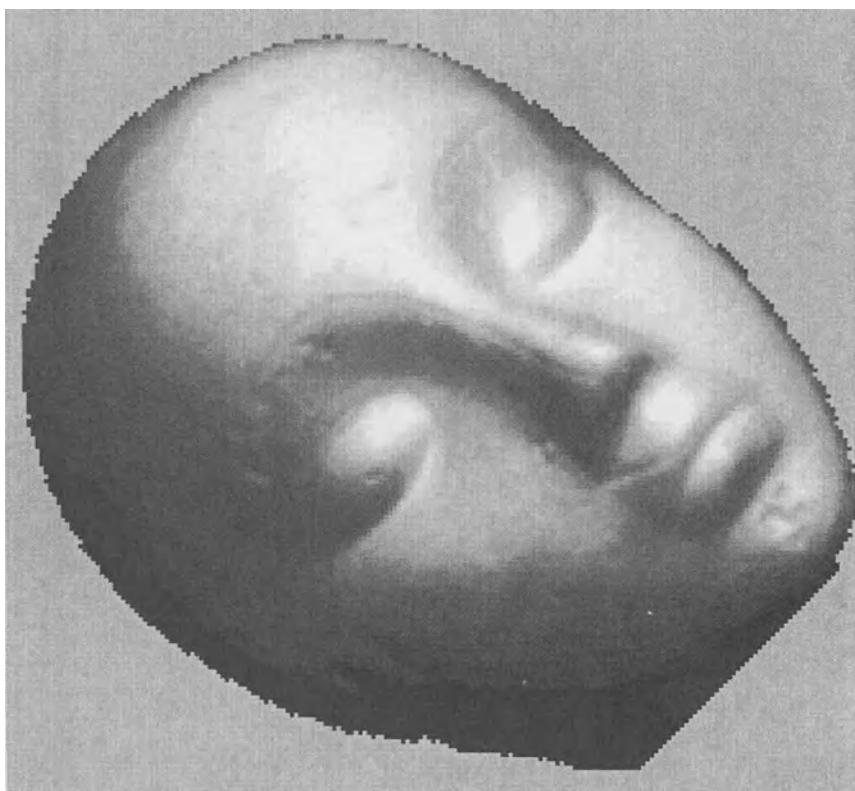


Figure 13.3. Mannequin image.

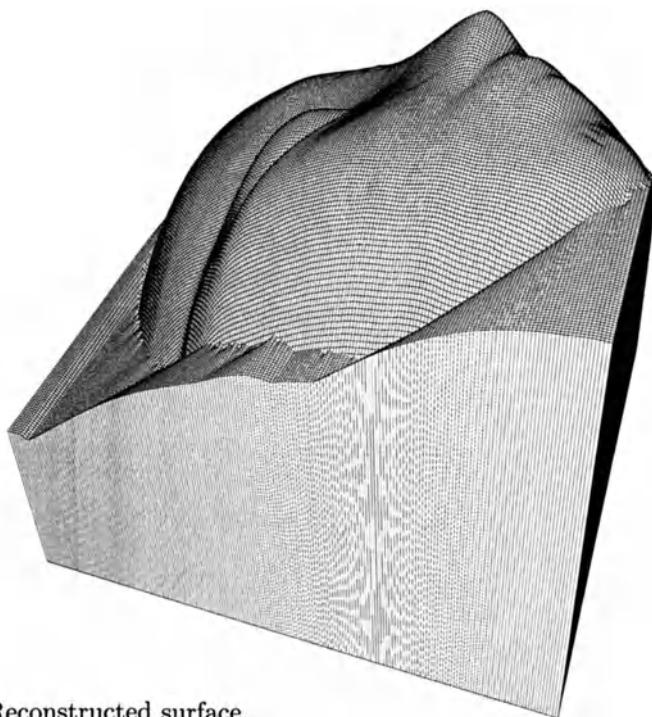


Figure 13.4. Reconstructed surface.

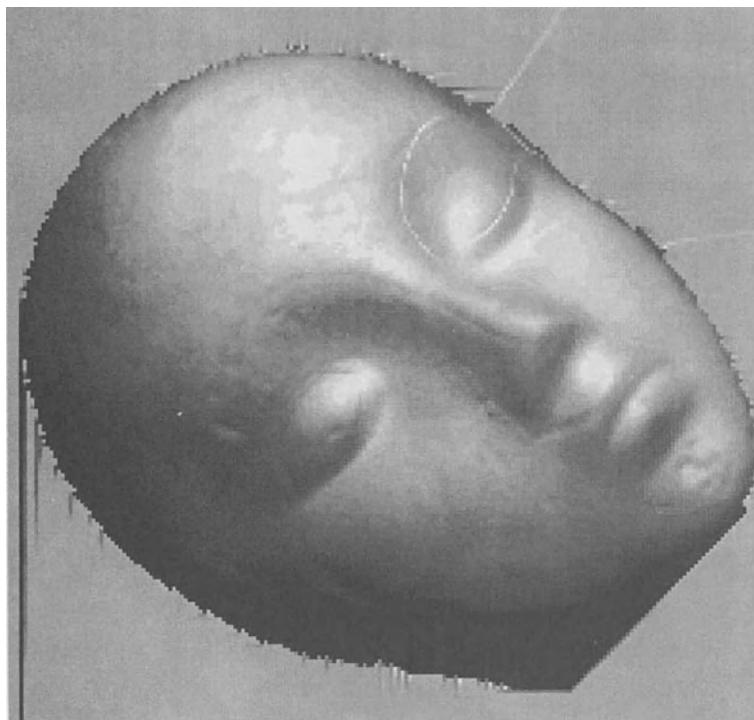


Figure 13.5. Illuminated reconstruction.

13.4 Problems with a Discontinuous Running Cost

Consider once again the finite time problem first introduced in (1.4) in Subsection 13.1.1.

$$V(x) = \inf \left[\int_0^{T \wedge \tau} k(\phi(s), \dot{\phi}(s)) ds + g(\phi(T \wedge \tau), T \wedge \tau) \right]. \quad (4.1)$$

Recall that τ is the first time that $\phi(\cdot)$ exits the interior of a given set G and that $T > 0$ is fixed.

In many calculus of variations problems, there is some underlying dynamical model for a “physical” system which determines the function $k(\cdot, \cdot)$, and minimizing (or nearly minimizing) paths $\phi(\cdot)$ have important physical interpretations. For example, in the case of classical Hamiltonian mechanics, the relevant laws of motion define $k(\cdot, \cdot)$, and the path $\phi(\cdot)$ describes the trajectory followed by a particle subject to those laws. In the case of geometric optics, $k(\cdot, \cdot)$ is defined in terms of the local speed of light. Thus, in the typical formulation of a calculus of variations problem appropriate to some given applied problem, the function $k(\cdot, \cdot)$ is obtained during the modelling stage.

Clearly, the function $k(\cdot, \cdot)$ reflects properties of the “medium” in which the underlying dynamical system evolves. Typically, the definition is local in the sense that $k(x, \cdot)$ reflects the properties of the medium at x . If $k(x, \alpha)$ possesses some kind of continuity in x , then it seems reasonable to believe this reflects a type of spatial continuity in the properties of the medium; e.g., the speed of propagation varies continuously in x in the geometrical optics problem. However, in many problems of interest such continuity properties may be violated. For example, it may be the case that the space is divided into two (or more) disjoint regions $R^{(1)}$ and $R^{(2)}$ interior to each of which the physical properties of the media vary continuously, with a smooth boundary or interface separating the regions, but with the physical properties being quite different from one region to the other. It is simple to produce examples of this type from classical mechanics or geometrical optics. More recent examples come from large deviation theory [34] (and, in particular, the application of large deviation theory to queueing systems [38]).

In such a case, one must rethink the modelling of the original physical problem as a calculus of variations problem. Clearly, the modelling appropriate for a single region of continuous behavior should be appropriate for defining or identifying $k(x, \cdot)$ when x is in the relative interior of either of the regions $R^{(1)}$ or $R^{(2)}$. However, there is still the question of the proper definition of $k(x, \cdot)$ for points x on the interface. This is quite important because in many cases the optimal paths will spend some time on the interface. The mathematical problem (4.1) will be well posed under

just appropriate measurability assumptions on $k(\cdot, \cdot)$. But from the point of view of modelling, certain additional properties can be expected (or perhaps should even be demanded). For example, regardless of how $k(x, \cdot)$ is defined on the interface it should lead to a cost function $V(x)$ which has desirable stability properties under approximations, discretizations, small perturbations, etc. This turns out to impose restrictions on the form of $k(x, \cdot)$.

In this section we will present what appears to be a “natural” definition of the integrand on the interface, and describe an associated numerical procedure. By natural, what is meant is that the definition occurs in applications, leads to a value that is stable under discretizations, and can be shown to be the *only* definition on the boundary that is stable under a wide class of approximations. We will also show that this particular definition of the cost on the interface, in spite of its complicated appearance, allows the use of relatively simple numerical approximation procedures. To simplify the notation we will assume that the interface is “flat”, in which case we can take the interface to be $\{x \in \mathbb{R}^k : x_1 = 0\}$, $R^{(1)} = \{x : x_1 \leq 0\}$, and $R^{(2)} = \{x : x_1 > 0\}$, where the subscript denotes the first component of x . Generalizing the results contained in this section to cover a smooth curved boundary is not difficult. The case of several intersecting boundaries, which we do not consider, is more subtle and not yet fully understood.

Remark on the Notation. We will need to use subscripts and superscripts in several different ways, such as to denote discrete time indices, regions [i.e., $R^{(1)}$ or $R^{(2)}$], components of a vector, and so on. To simplify, we will use the following conventions. We will denote the first component of a vector α by $(\alpha)_1$. Thus, the symbol α_1 will not be used for the first component of α . Furthermore, if a quantity is intrinsically associated to the region $R^{(i)}$, then it will bear a superscript of the form (i) [e.g., $\alpha^{(1)}$ and $\alpha^{(2)}$].

13.4.1 DEFINITION AND INTERPRETATION OF THE COST ON THE INTERFACE

For each $i = 1, 2$, let $k^{(i)} : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$ satisfy (1.1) and (A2.2). For simplicity we will assume that $k^{(i)}(x, \cdot)$ is convex for each x and $i = 1, 2$. The convexity assumption can be dropped at the expense of complicating the formula for $k^{(0)}(\cdot, \cdot)$ given in the next two paragraphs.

We define

$$k^{(0)}(x, \alpha) = \inf \left\{ \rho^{(1)} k^{(1)}(x, \alpha^{(1)}) + \rho^{(2)} k^{(2)}(x, \alpha^{(2)}) \right\}, \quad (4.2)$$

where the infimum is over $(\rho^{(1)}, \rho^{(2)}) \in \mathbb{R}^2$ and $(\alpha^{(1)}, \alpha^{(2)}) \in \mathbb{R}^{2k}$ satisfying

$$\rho^{(1)} + \rho^{(2)} = 1, \quad \rho^{(1)} \geq 0, \quad \rho^{(2)} \geq 0, \quad (4.3)$$

$$(\alpha^{(1)})_1 \geq 0, (\alpha^{(2)})_1 \leq 0, \quad (4.4)$$

$$\rho^{(1)}\alpha^{(1)} + \rho^{(2)}\alpha^{(2)} = \alpha. \quad (4.5)$$

We then define

$$k(x, \alpha) = \begin{cases} k^{(1)}(x, \alpha) & \text{if } (x)_1 < 0 \\ k^{(0)}(x, \alpha) & \text{if } (x)_1 = 0 \\ k^{(2)}(x, \alpha) & \text{if } (x)_1 > 0. \end{cases} \quad (4.6)$$

Clearly, the function $k(x, \alpha)$ may be discontinuous in x , and, therefore, (A2.2) may not be satisfied. It follows directly from the definition that $k^{(0)}(x, \alpha) \leq k^{(1)}(x, \alpha)$ for α satisfying $(\alpha)_1 \geq 0$ and $k^{(0)}(x, \alpha) \leq k^{(2)}(x, \alpha)$ for α satisfying $(\alpha)_1 \leq 0$. It also follows that $k^{(0)}(\cdot, \cdot)$ satisfies the superlinearity condition (1.1). Note that only those values of $k^{(0)}(x, \alpha)$ for x and α satisfying $(x)_1 = 0$ and $(\alpha)_1 = 0$ affect the value of $V(x)$.

Remark. An analogous definition can be given for more general classes of deterministic optimal control problems under a mild “controllability” condition. The appropriate definition and controllability condition can be found by applying the heuristic argument given below.

Interpretation of the Cost. Perhaps the simplest way to motivate the form of the cost $k^{(0)}(\cdot, \cdot)$ on the interface is to imagine the continuous time problem (4.1) as arising as the limit of a sequence of discrete time problems. To simplify the notation, we will take $G = \mathbb{R}^k$, and use $g(x)$ instead of $g(x, t)$. Thus, let $\Delta > 0$ and consider discrete time dynamics of the form

$$\phi_{i+1}^\Delta = \phi_i^\Delta + \Delta u_i, \quad \phi_0^\Delta = x,$$

and suppose that the cost to be minimized over all $\{u_i, i = 0, \dots, (T/\Delta) - 1\}$ is

$$\sum_{i=0}^{(T/\Delta)-1} \Delta \left[I_{\{(\phi_i^\Delta)_1 \leq 0\}} k^{(1)}(\phi_i^\Delta, u_i) + I_{\{(\phi_i^\Delta)_1 > 0\}} k^{(2)}(\phi_i^\Delta, u_i) \right] + g(\phi_{(T/\Delta)}^\Delta).$$

For specificity, we have assigned the cost $k^{(1)}(\cdot, \cdot)$ to points on the interface for the discrete time problem, but this does not actually matter in the limit $\Delta \rightarrow 0$.

One can then ask if there is a continuous time calculus of variations problem for which the minimal cost is the limit of the minimal costs for these discrete time problems, and if so, what is the running cost. It is not hard to show that there is a limiting continuous time problem and that it is given by (4.1) with the running cost defined by (4.2) to (4.6). It is obvious that the correct running cost away from the interface is either $k^{(1)}(\cdot, \cdot)$ or $k^{(2)}(\cdot, \cdot)$. That $k^{(0)}(\cdot, \cdot)$ gives the proper definition on the interface can be heuristically argued as follows. Consider a section $i = i_1, i_1 + 1, \dots, i_2$ of the

discrete trajectory $\{\phi_i^\Delta, i = 0, \dots, T/\Delta - 1\}$ when ϕ_i^Δ is close to the interface, and suppose that the average (in time) “velocity” (i.e., the average of the corresponding u_i) for $i \in \{i_1, \dots, i_2\}$ is α . Because $(\phi_i^\Delta)_1 \approx 0$ at both the beginning and end of the section, we have $(\alpha)_1 \approx 0$. Given that the process will remain in a small neighborhood of the boundary for a given period of time, it has the option of “selecting” between the two costs $k^{(1)}(\cdot, \cdot)$ and $k^{(2)}(\cdot, \cdot)$, since the process can move quickly and “cheaply” from one side of the interface to the other. Thus, for a certain fraction of time, the running cost $k^{(1)}(\cdot, \cdot)$ will be used and, for the remaining fraction of time, the cost $k^{(2)}(\cdot, \cdot)$ will apply. In order to learn the exact form of the cost on the interface, we must determine the optimal way in which the discrete time process can exploit the two available costs $k^{(1)}(\cdot, \cdot)$ and $k^{(2)}(\cdot, \cdot)$ while simultaneously maintaining an average velocity α . The definition of $k^{(0)}(\phi_{i_1}^\Delta, \alpha)$ suggests the answer. Let $\rho^{(1)}, \rho^{(2)}, \alpha^{(1)}$, and $\alpha^{(2)}$ minimize (4.2) subject to (4.3)–(4.6). Then a nearly optimal way for the discrete time process to behave is to use $u_i = \alpha^{(1)}$ when the state ϕ_i^Δ is in the set $\{x : (x)_1 \leq 0\}$ and $u_i = \alpha^{(2)}$ when the state is in the set $\{x : (x)_1 > 0\}$. The $\rho^{(i)}$ turn out to represent respective fractions of time spent in the two sets, and the interpretations of (4.3) and (4.5) are clear. The constraint (4.4) is essentially a feasibility condition. If this condition is violated, then the discrete time process will actually move away from the interface, contradicting our assumption that it remain near the interface for $i \in \{i_1, i_1 + 1, \dots, i_2\}$.

13.4.2 NUMERICAL SCHEMES AND THE PROOF OF CONVERGENCE

We next consider the construction of numerical schemes for the problem (4.1). It turns out that we do not need to have access to or to compute $k(\cdot, \cdot)$ for the points on the interface $\{x : (x)_1 = 0\}$. Instead, a convergent scheme can be constructed using only $k^{(1)}(\cdot, \cdot)$ and $k^{(2)}(\cdot, \cdot)$. This is due to the form that the running cost takes on the interface and is another manifestation of the robustness of the definition of the cost given in (4.2)–(4.6). The numerical scheme to be defined below, which uses only the costs $k^{(1)}(\cdot, \cdot)$ and $k^{(2)}(\cdot, \cdot)$, will nonetheless converge to the value function with the cost $k^{(0)}(\cdot, \cdot)$ on the interface. This fact is of some significance because it means that $k^{(0)}(\cdot, \cdot)$ *never actually needs to be computed* when computing the approximations to $V(x)$. This is especially useful for problems where the functions $k^{(1)}(\cdot, \cdot)$ and $k^{(2)}(\cdot, \cdot)$ take a relatively simple form in α , such as quadratic or exponential. For these cases, the minimizations that must be done in order to compute an approximation to $V(x)$ often can be done analytically rather than numerically. This would not be the case if $k^{(0)}(\cdot, \cdot)$ appeared in the definition of the schemes.

We can use either the implicit or explicit schemes described in Subsection 13.2.1. For concreteness, we will consider the explicit method given by (2.7).

However, as remarked above, we will not use the running cost $k(\cdot, \cdot)$ in the formulation of the scheme. If $x \in G_h^0$ and $(x)_1 < 0$, then we will use the running cost $k^{(1)}(\cdot, \cdot)$. If $x \in G_h^0$ and $(x)_1 > 0$, we will use $k^{(2)}(\cdot, \cdot)$. For any points in G_h^0 that satisfy $(x)_1 = 0$, we can use either $k^{(1)}(\cdot, \cdot)$ or $k^{(2)}(\cdot, \cdot)$ (it will not actually matter in the limit as the discretization parameters tend to zero), and for specificity we will use $k^{(1)}(\cdot, \cdot)$ for these points. Thus, if we define

$$\bar{k}(x, \alpha) = \begin{cases} k^{(1)}(x, \alpha) & \text{if } (x)_1 \leq 0 \\ k^{(2)}(x, \alpha) & \text{if } (x)_1 > 0, \end{cases} \quad (4.7)$$

then the numerical scheme becomes

$$V^{h,\delta}(x, n\delta) = \min_{\alpha \in \mathcal{U}^{h,\delta}} \left[\sum_y \tilde{p}^{h,\delta}(x, y|\alpha) V^{h,\delta}(y, n\delta + \delta) + \bar{k}(x, \alpha)\delta \right] \quad (4.8)$$

for $x \in G_h^0$ and $n\delta < T$, together with the boundary and terminal condition $V^{h,\delta}(x, n\delta) = g(x, n\delta)$ for $x \notin G_h^0$ and $n\delta < T$ or $x \in G_h^0$ and $n\delta = T$. Recall the $\mathcal{U}^{h,\delta}$ was defined in (2.5).

Because our main interest in this section is in examining the new features that are associated with the discontinuous running cost, we will replace (A2.1) and (A2.3) by the following somewhat stronger conditions.

A4.1. *The set G is compact and satisfies interior and exterior cone conditions: There exist $\epsilon > 0$ and continuous functions $v(\cdot)$ and $w(\cdot)$ such that given any $x \in \partial G$, $\cup_{0 < a < \epsilon} B_{\epsilon a}(x + av(x)) \subset G$ and $\cup_{0 < a < \epsilon} B_{\epsilon a}(x + aw(x)) \cap G = \emptyset$. The set G also satisfies the additional condition that $(v(x))_1 = 0$ if $x \in \partial G$ and $(x)_1 = 0$.*

A4.2. *The function $g(\cdot, \cdot)$ is continuous and bounded on $\mathbb{R}^k \times [0, T]$.*

Condition (A4.1) essentially states that there are no points in ∂G at which ∂G is “tangent” to the interface $\{x : (x)_1 = 0\}$. We can then state the following result.

Theorem 4.1. *Assume (A4.1), (A4.2), and that both $k^{(1)}(\cdot, \cdot)$ and $k^{(2)}(\cdot, \cdot)$ satisfy (1.1) and (A2.2). Let $k(\cdot, \cdot)$ be defined by (4.2)–(4.6), assume that (2.6) holds, and let $V(x)$ be defined by (4.1). Then for the explicit scheme defined by (4.8), we have*

$$V^{h,\delta}(x) \rightarrow V(x).$$

Proof. We begin by recalling the notation used in the proof of Theorem 2.5. Thus, $\{\xi_i^{h,\delta}, i < \infty\}$ is a controlled Markov chain as described in Subsection 13.2.1 with transition probabilities $\tilde{p}^{h,\delta}(x, y|\alpha)$, interpolation interval δ , and initial condition $x \in G^0$. If $\{u_i^{h,\delta}, i < \infty\}$ is a sequence of controls applied to the chain, then we define the interpolated process and control by setting

$$\xi^{h,\delta}(t) = \xi_i^{h,\delta}, \quad u^{h,\delta}(t) = u_i^{h,\delta}, \quad t \in [i\delta, i\delta + \delta].$$

We define $\tau_{h,\delta}$ to be the first time $\xi^{h,\delta}(\cdot)$ leaves G_h^0 , and let $m^{h,\delta}(\cdot)$ denote the relaxed control representation of $u^{h,\delta}(\cdot)$.

Before proving the lower bound, we will state the following lemma. The lemma is analogous to and is used in the same way as Theorem 2.3.

Lemma 4.2. *Assume the conditions of Theorem 4.1, and let $\phi(\cdot)$ be an absolutely continuous function that satisfies $\phi(0) \in G^0$, $\phi(t) \in G$ for $t \in [0, \sigma]$, and*

$$\int_0^\sigma k(\phi(s), \dot{\phi}(s))ds < \infty.$$

Then given $\epsilon > 0$, there exists an absolutely continuous function $\phi^\epsilon(\cdot)$ such that $\phi^\epsilon(t) \in G^0$ for $t \in [0, \sigma]$, $\phi^\epsilon(0) = \phi(0)$, $\phi^\epsilon(\sigma) = \phi(\sigma)$, and

$$\left| \int_0^\sigma k(\phi(s), \dot{\phi}(s))ds - \int_0^\sigma k(\phi^\epsilon(s), \dot{\phi}^\epsilon(s))ds \right| < \epsilon.$$

The proof is essentially the same as that of Theorem 2.3 and is therefore omitted. We will only note that under (A4.1) we can assume the existence of $\gamma > 0$ and $v^*(\cdot)$ such that $x \in \partial G$ and $|x_1| \leq \gamma$ imply $(v^*(x))_1 = 0$, where $\hat{v}(\cdot)$ satisfies the conditions on $v(\cdot)$ given in the statement of the interior cone condition. This means that when perturbing $\phi(\cdot)$ as in the proof of Theorem 2.3, the applied perturbation will be parallel to the interface when $\phi(\cdot)$ is on the interface. This implies a continuity property for the running cost in spite of the discontinuity across the interface.

Proof of the Lower Bound. Consider any sequence of ordinary admissible controls $\{u_i^{h,\delta}, i < \infty\}$ that are applied to the chain and let $m^{h,\delta}(\cdot)$ be the associated relaxed control representation of the interpolation $u^{h,\delta}(\cdot)$. In proving the lower bound, we may assume that the associated running costs are bounded from above. By Theorem 2.4, $\{(\xi^{h,\delta}(\cdot), m^{h,\delta}(\cdot)), h > 0, \delta > 0\}$ is tight.

We next define random measures $\nu^{h,\delta}(\cdot)$ on $\mathbb{R}^k \times D^k [0, T] \times [0, T]$ by

$$\nu^{h,\delta}(A \times B \times C) = \int_C I_{\{u^{h,\delta}(s) \in A, \xi^{h,\delta}(s) \in B\}} ds.$$

Note that $\nu^{h,\delta}(A \times D^k [0, T] \times C) = m^{h,\delta}(A \times C)$. Therefore, the tightness of $\{(\xi^{h,\delta}(\cdot), m^{h,\delta}(\cdot)), h > 0, \delta > 0\}$ implies the tightness of $\{\nu^{h,\delta}(\cdot), h > 0, \delta > 0\}$. We also define

$$S^{(1)} = \{(\phi(\cdot), s) \in D^k [0, T] \times [0, T] : (\phi(s))_1 \leq 0\},$$

$$S^{(2)} = \{(\phi(\cdot), s) \in D^k [0, T] \times [0, T] : (\phi(s))_1 > 0\},$$

$$\nu^{(i),h,\delta}(Q) = \nu^{h,\delta}\left(Q \cap (\mathbb{R}^k \times S^{(i)})\right), \quad i = 1, 2.$$

The measures $\nu^{(1),h,\delta}(\cdot)$ and $\nu^{(2),h,\delta}(\cdot)$ record the control effort that is applied, when it is applied, and also distinguish between when the state $\xi^{h,\delta}(s)$ is in $\{x : (x)_1 \leq 0\}$ and $\{x : (x)_1 > 0\}$. Clearly, $\nu^{h,\delta}(\cdot) = \nu^{(1),h,\delta}(\cdot) + \nu^{(2),h,\delta}(\cdot)$.

We now apply the Skorohod representation and extract a weakly converging subsequence from

$$\left\{ \left(\xi^{h,\delta}(\cdot), m^{h,\delta}(\cdot), \nu^{h,\delta}(\cdot), \nu^{(1),h,\delta}(\cdot), \nu^{(2),h,\delta}(\cdot), \tau_{h,\delta} \right), h > 0, \delta > 0 \right\}$$

with limit

$$(x(\cdot), m(\cdot), \nu(\cdot), \nu^{(1)}(\cdot), \nu^{(2)}(\cdot), \tilde{\tau}).$$

It follows easily from $\xi^{h,\delta}(\cdot) \rightarrow x(\cdot)$ and the weak convergence that

$$\nu^{(1)}(\mathbb{R}^k \times \{x(\cdot)\}^c \times [0, T]) = \nu^{(2)}(\mathbb{R}^k \times \{x(\cdot)\}^c \times [0, T]) = 0$$

w.p.1. We may therefore conclude the existence of measures $\nu_s^{(1)}(\cdot)$ and $\nu_s^{(2)}(\cdot)$, $0 \leq s \leq T$, such that $\nu^{(i)}(A \times \{x(\cdot)\} \times [0, t]) = \int_0^t \nu_s^{(i)}(A) ds$, $i = 1, 2$, for all Borel sets A and $t \in [0, T]$.

The measures $\nu_s^{(1)}(\cdot)$ and $\nu_s^{(2)}(\cdot)$ possess the following properties. Almost surely in s , and w.p.1,

$$(x(s))_1 < 0 \Rightarrow \int_{\mathbb{R}^k} \nu_s^{(1)}(d\alpha) = 1, \quad (4.9)$$

$$(x(s))_1 > 0 \Rightarrow \int_{\mathbb{R}^k} \nu_s^{(2)}(d\alpha) = 1,$$

$$\int_{\mathbb{R}^k} \nu_s^{(1)}(d\alpha) + \int_{\mathbb{R}^k} \nu_s^{(2)}(d\alpha) = 1, \quad (4.10)$$

$$(x(s))_1 = 0 \Rightarrow \int_{\mathbb{R}^k} (\alpha)_1 \nu_s^{(1)}(d\alpha) \geq 0 \text{ and } \int_{\mathbb{R}^k} (\alpha)_1 \nu_s^{(2)}(d\alpha) \leq 0, \quad (4.11)$$

$$\int_{\mathbb{R}^k} \alpha \nu_s^{(1)}(d\alpha) + \int_{\mathbb{R}^k} \alpha \nu_s^{(2)}(d\alpha) = \int_{\mathbb{R}^k} \alpha m_s(d\alpha) = \dot{x}(s). \quad (4.12)$$

Equation (4.9) follows easily from the definitions of the $\nu^{(i),h,\delta}(\cdot)$ and the weak convergence, whereas (4.10) and (4.12) follow from the relationship $\nu^{h,\delta}(A \times D^k [0, T] \times C) = m^{h,\delta}(A \times C)$. The only property that is not obvious is (4.11). We will first prove the lower bound assuming (4.11) and then show (4.11).

Now fix an ω for which there is convergence via the Skorokhod representation. As usual we must consider separate cases.

The case $\tilde{\tau} > T$. In this case,

$$\begin{aligned}
& \liminf_{(h,\delta) \rightarrow 0} \int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbb{R}^k} \bar{k}(\xi^{h,\delta}(s), \alpha) m^{h,\delta}(d\alpha ds) \\
&= \liminf_{(h,\delta) \rightarrow 0} \int_0^T \int_{\mathbb{R}^k} k^{(1)}(\xi^{h,\delta}(s), \alpha) \nu^{(1),h,\delta}(d\alpha \times D^k [0, T] \times ds) \\
&\quad + \int_0^T \int_{\mathbb{R}^k} k^{(2)}(\xi^{h,\delta}(s), \alpha) \nu^{(2),h,\delta}(d\alpha \times D^k [0, T] \times ds) \\
&\geq \int_0^T \int_{\mathbb{R}^k} k^{(1)}(x(s), \alpha) \nu^{(1)}(d\alpha \times D^k [0, T] \times ds) \\
&\quad + \int_0^T \int_{\mathbb{R}^k} k^{(2)}(x(s), \alpha) \nu^{(2)}(d\alpha \times D^k [0, T] \times ds) \\
&= \int_0^T \int_{\mathbb{R}^k} \left[k^{(1)}(x(s), \alpha) \nu_s^{(1)}(d\alpha) + k^{(2)}(x(s), \alpha) \nu_s^{(2)}(d\alpha) \right] ds.
\end{aligned}$$

The set $\{s : (x(s))_1 = 0, (\dot{x}(s))_1 \neq 0\}$ is a set of measure zero. Therefore, the definition of $k(\cdot, \cdot)$, the convexity of the $k^{(i)}(x, \cdot)$, and the properties of the $\nu_s^{(i)}(\cdot)$ given in (4.9)-(4.12) imply

$$\int_{\mathbb{R}^k} \left[k^{(1)}(x(s), \alpha) \nu_s^{(1)}(d\alpha) + k^{(2)}(x(s), \alpha) \nu_s^{(2)}(d\alpha) \right] \geq k(x(s), \dot{x}(s))$$

a.s. in s . We also have

$$g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}) \rightarrow g(x(T), T).$$

Assembling the inequalities and using the definition of $V(x)$, we conclude that

$$\begin{aligned}
& \liminf_{(h,\delta) \rightarrow 0} \left[\int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbb{R}^k} \bar{k}(\xi^{h,\delta}(s), \alpha) m^{h,\delta}(d\alpha ds) \right. \\
& \quad \left. + g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}) \right] \geq V(x).
\end{aligned} \tag{4.13}$$

The case $\tilde{\tau} \leq T$. In this case, $x(\cdot)$ exits G^0 by time T . As always, there is the difficulty due to the fact that $\tau = \inf\{t : x(t) \in \partial G\}$ might be smaller than $\tilde{\tau}$. The same argument as in the case $\tilde{\tau} > T$ shows

$$\begin{aligned}
& \liminf_{(h,\delta) \rightarrow 0} \int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbb{R}^k} \bar{k}(\xi^{h,\delta}(s), \alpha) m^{h,\delta}(d\alpha ds) + g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}) \\
& \geq \int_0^{\tilde{\tau}} k(x(s), \dot{x}(s)) ds + g(x(\tilde{\tau}), \tilde{\tau}).
\end{aligned}$$

Using Lemma 4.2 in precisely the same way that Theorem 2.3 was used in the proof of Theorem 2.5, we again conclude that (4.13) holds.

It follows from the conclusions for the two cases and Fatou's lemma that

$$\liminf_{(h,\delta) \rightarrow 0} V^{h,\delta}(x) \geq V(x)$$

for the convergent subsequence. The proof of the lower bound is completed by using an argument by contradiction.

We now prove (4.11). Consider any subsequence from

$$\left\{ \left(\xi^{h,\delta}(\cdot), m^{h,\delta}(\cdot), \nu^{h,\delta}(\cdot), \nu^{(1),h,\delta}(\cdot), \nu^{(2),h,\delta}(\cdot) \right), h > 0, \delta > 0 \right\}$$

that converges weakly to a limit

$$(x(\cdot), m(\cdot), \nu(\cdot), \nu^{(1)}(\cdot), \nu^{(2)}(\cdot)).$$

By the Skorokhod representation, we can assume that the convergence is w.p.1. Fix $\gamma > 0$. In order to proceed with the proof of (4.11) we must first build an approximation to the function $F : I\!\!R \rightarrow I\!\!R$ given by

$$F(z) = \begin{cases} |z| & \text{if } |z| \leq \gamma \\ \gamma & \text{if } |z| > \gamma. \end{cases}$$

For each $\eta > 0$, let $F^\eta(\cdot)$ be a function such that

$$\begin{aligned} |F^\eta(z)| &\leq 2\gamma \quad \text{for all } z, \\ F^\eta(z) &= F(z) \quad \text{for } z \in [-\gamma, \gamma], \\ |F_{zz}^\eta(z)| &< B \quad \text{for } z \notin [-\gamma/2, \gamma/2], \end{aligned}$$

where $B < \infty$ depends on $\eta > 0$. Let

$$f^\eta(z) = \begin{cases} F_z^\eta(z) & \text{if } z \neq 0 \\ -1 & \text{if } z = 0, \end{cases}$$

and assume that $f^\eta(z) \rightarrow 0$ for $z \notin [-\gamma, \gamma]$. For all small $h > 0$, local consistency implies that

$$\begin{aligned} 4\gamma &\geq E_x^{m^{h,\delta}} \sum_{j=0}^{T/\delta-1} F^\eta((\xi_{j+1}^{h,\delta})_1) - F^\eta((\xi_j^{h,\delta})_1) \\ &\geq E_x^{m^{h,\delta}} \sum_{j=0}^{T/\delta-1} \delta f^\eta((\xi_j^{h,\delta})_1) \left[(u_j^{h,\delta})_1 + o(|u_j^{h,\delta}|) \right] + B\delta o(|u_j^{h,\delta}|). \end{aligned}$$

As noted many times so far, the boundedness of the running costs and the superlinearity condition (1.1) imply

$$\limsup_{(h,\delta) \rightarrow 0} E_x^{m^{h,\delta}} \sum_{j=0}^{T/\delta-1} \delta |u_j^{h,\delta}| = \limsup_{(h,\delta) \rightarrow 0} E_x^{m^{h,\delta}} \int_0^T \int_{I\!\!R^k} |\alpha| m^{h,\delta}(d\alpha ds) < \infty.$$

Sending $(h, \delta) \rightarrow 0$, we have

$$4\gamma \geq E_x^m \left[\int_0^T \int_{\mathbb{R}^k} (\alpha)_1 f^{(2),\eta}(x(s)) \nu^{(2)}(d\alpha \times D^k [0, T] \times ds) \right. \\ \left. + \int_0^T \int_{\mathbb{R}^k} (\alpha)_1 f^{(1),\eta}(x(s)) \nu^{(1)}(d\alpha \times D^k [0, T] \times ds) \right],$$

where $f^{(1),\eta}(\cdot)$ [respectively $f^{(2),\eta}(\cdot)$] is a continuous extension of $f^\eta(\cdot)$ from $(-\infty, 0)$ to $(-\infty, 0]$ [respectively $(0, \infty)$ to $[0, \infty)$].

Sending $\eta \rightarrow 0$, and using the definition of $f^\eta(\cdot)$, we have

$$E_x^m \left[\int_{\mathbb{R}^k} (\alpha)_1 \nu^{(2)}(d\alpha \times S^\gamma) - \int_{\mathbb{R}^k} (\alpha)_1 \nu^{(1)}(d\alpha \times S^\gamma) \right] \leq 4\gamma,$$

where $S^\gamma = \{(\psi(\cdot), s) \in D^k [0, T] \times [0, T] : |(\psi(s))_1| \leq \gamma\}$. A very similar proof that is based on approximating the function

$$F(z) = \begin{cases} z & \text{if } |z| \leq \gamma \\ \gamma & \text{if } z > \gamma \\ -\gamma & \text{if } z < -\gamma \end{cases}$$

shows that

$$\left| E_x^m \left[\int_{\mathbb{R}^k} (\alpha)_1 \nu^{(2)}(d\alpha \times S^\gamma) + \int_{\mathbb{R}^k} (\alpha)_1 \nu^{(1)}(d\alpha \times S^\gamma) \right] \right| \leq 4\gamma.$$

Adding and subtracting these equations gives

$$E_x^m \int_{\mathbb{R}^k} (\alpha)_1 \nu^{(1)}(d\alpha \times S^\gamma) \geq -8\gamma, \quad E_x^m \int_{\mathbb{R}^k} (\alpha)_1 \nu^{(2)}(d\alpha \times S^\gamma) \leq 8\gamma.$$

Thus, for each $\theta > 0$, both the quantities

$$P_x \left\{ \int_{\mathbb{R}^k} (\alpha)_1 \nu^{(1)}(d\alpha \times S^\gamma) \leq -\theta \right\}, \quad P_x \left\{ \int_{\mathbb{R}^k} (\alpha)_1 \nu^{(2)}(d\alpha \times S^\gamma) \geq \theta \right\}$$

tend to zero as $\gamma \rightarrow 0$. From the definitions of the sets S^γ and the measures $\nu^{(1)}(\cdot)$ and $\nu^{(2)}(\cdot)$,

$$\int_{\mathbb{R}^k} (\alpha)_1 \nu^{(1)}(d\alpha \times \{(\psi(\cdot), s) \in D^k [0, T] \times [0, T] : (\psi(s))_1 = 0\}) \geq 0, \\ \int_{\mathbb{R}^k} (\alpha)_1 \nu^{(2)}(d\alpha \times \{(\psi(\cdot), s) \in D^k [0, T] \times [0, T] : (\psi(s))_1 = 0\}) \leq 0 \tag{4.14}$$

w.p.1.

The argument that led to (4.14) can be repeated with s restricted to any interval $[a, b] \subset [0, T]$ with the same conclusion. Thus we can assume it

holds simultaneously for all such intervals with rational endpoints. Using the definitions of $\nu_s^{(1)}(\cdot)$ and $\nu_s^{(2)}(\cdot)$, this implies

$$\int_{\mathbb{R}^k} (\alpha)_1 \nu_s^{(1)}(d\alpha) \geq 0 \quad \text{and} \quad \int_{\mathbb{R}^k} (\alpha)_1 \nu_s^{(2)}(d\alpha) \leq 0$$

whenever $(x(s))_1 = 0$ a.s. in s and with probability one. This proves (4.11).

Before proving the upper bound we present the following lemma. The lemma implies the existence of an ϵ -optimal piecewise linear path for the calculus of variations problem and is analogous to Theorem 2.2. The proof of the lemma is given at the end of the section.

Lemma 4.3. *Assume the conditions of Theorem 4.1. Then given any $\epsilon > 0$ and any absolutely continuous path $\phi : [0, T] \rightarrow \mathbb{R}^k$ satisfying $\phi(0) = x$ and $\int_0^T k(\phi(s), \dot{\phi}(s))ds < \infty$, there exist $N < \infty$, $t_0 < t_1 < \dots < t_N = T$, and a function $u^\epsilon : [0, T] \rightarrow \mathbb{R}^k$ which is constant on the intervals $[t_n, t_{n+1})$, $n < N$, such that if*

$$\phi^\epsilon(t) = x + \int_0^t u^\epsilon(s)ds,$$

then

$$\sup_{0 \leq t \leq T} |\phi^\epsilon(t) - \phi(t)| \leq \epsilon$$

and

$$\sup_{0 \leq t \leq T} \left[\int_0^t k(\phi^\epsilon(s), u^\epsilon(s))ds - \int_0^t k(\phi(s), \dot{\phi}(s))ds \right] \leq \epsilon.$$

Furthermore, we can assume that for any $n < N$ that either $(\phi^\epsilon(t))_1 \neq 0$ for all $t \in (t_n, t_{n+1})$ or $(\phi^\epsilon(t))_1 = 0$ for all $t \in (t_n, t_{n+1})$.

Proof of the Upper Bound. Fix $\epsilon > 0$, and choose $\phi(\cdot)$ with $\phi(0) = x$ such that if $\tau = \{t : \phi(t) \in \partial G\}$, then

$$\int_0^{T \wedge \tau} k(\phi(s), \dot{\phi}(s))ds + g(\phi(T \wedge \tau), T \wedge \tau) \leq V(x) + \epsilon. \quad (4.15)$$

The case $\tau > T$. By Lemma 4.3 and (A4.2), there exists $\phi^\epsilon(\cdot)$ satisfying the conditions of the lemma and also

$$\int_0^T k(\phi^\epsilon(s), \dot{\phi}^\epsilon(s))ds + g(\phi^\epsilon(T), T) \leq \int_0^T k(\phi(s), \dot{\phi}(s))ds + g(\phi(T), T) + \epsilon. \quad (4.16)$$

For each $n < N$, let $\alpha_n = \dot{\phi}^\epsilon(t)$, where t is any point in (t_n, t_{n+1}) . If $(\phi^\epsilon(t))_1 \neq 0$ for $t \in (t_n, t_{n+1})$, then we define $\alpha_n^{(1)} = \alpha_n^{(2)} = \alpha_n$. If $(\phi^\epsilon(t))_1 = 0$ for $t \in (t_n, t_{n+1})$, then we must prescribe a control that will yield a

running cost close to $k^{(0)}(\phi^\epsilon(t), \alpha_n)$. By exploiting the continuity of the $k^{(j)}(\cdot, \cdot)$, $j = 1, 2$, there exist $\rho_n^{(1)}, \rho_n^{(2)}, \alpha_n^{(1)}$ and $\alpha_n^{(2)}$ which satisfy

$$\rho_n^{(1)} + \rho_n^{(2)} = 1, \quad \rho_n^{(1)} > 0, \quad \rho_n^{(2)} > 0, \quad (4.17)$$

$$(\alpha_n^{(1)})_1 > 0, \quad (\alpha_n^{(2)})_1 < 0, \quad (4.18)$$

$$\rho_n^{(1)}\alpha_n^{(1)} + \rho_n^{(2)}\alpha_n^{(2)} = \alpha_n \quad (4.19)$$

and

$$\rho_n^{(1)}k^{(1)}(\phi^\epsilon(t), \alpha_n^{(1)}) + \rho_n^{(2)}k^{(2)}(\phi^\epsilon(t), \alpha_n^{(2)}) \leq k^{(0)}(\phi^\epsilon(t), \alpha_n) + \epsilon \quad (4.20)$$

for all $t \in (t_n, t_n + \delta)$, where δ is some positive number. Because $\phi^\epsilon(t)$ is not constant, it may not actually be the case that (4.20) can be guaranteed for all $t \in (t_n, t_{n+1})$ simultaneously. However, the continuity properties of $k^{(i)}(\cdot, \cdot)$ and $\phi^\epsilon(\cdot)$ imply that we can replace the original partition by a finer partition $0 = \tilde{t}_0 < \dots < \tilde{t}_{\tilde{N}} = T$ (if necessary) such that if $(\phi^\epsilon(t))_1 = 0$ for $t \in (\tilde{t}_n, \tilde{t}_{n+1})$, then (4.20) holds for all $t \in (\tilde{t}_n, \tilde{t}_{n+1})$. For simplicity, we will retain the same notation for the original and refined partition.

For the remainder of the proof we will assume that h and δ are small enough that

$$\cup_{n=0}^{N-1}\{\alpha_n^{(1)}, \alpha_n^{(2)}\} \subset \mathcal{U}^{h, \delta}.$$

We can then define a nonanticipative control scheme for $\{\xi_i^{h, \delta}, i < \infty\}$ in terms of the $\alpha_n^{(j)}$'s. For $i\delta \in [t_n, t_{n+1})$, we set

$$u_i^{h, \delta} = \begin{cases} \alpha_n^{(1)} & \text{if } (\xi_i^{h, \delta})_1 \leq 0 \\ \alpha_n^{(2)} & \text{if } (\xi_i^{h, \delta})_1 > 0. \end{cases}$$

Define $\xi^{h, \delta}(\cdot), m^{h, \delta}(\cdot), \nu^{h, \delta}(\cdot), \nu^{(1), h, \delta}(\cdot), \nu^{(2), h, \delta}(\cdot)$ and $\tau_{h, \delta}$ as in the proof of the lower bound, but for this new control sequence. Because the $\alpha_n^{(j)}$'s are all bounded, the collection

$$\left\{ \left(\xi^{h, \delta}(\cdot), m^{h, \delta}(\cdot), \nu^{h, \delta}(\cdot), \nu^{(1), h, \delta}(\cdot), \nu^{(2), h, \delta}(\cdot), \tau_{h, \delta} \right), h > 0, \delta > 0 \right\}$$

is clearly tight. Extract a convergent subsequence, with limit

$$(x(\cdot), m(\cdot), \nu(\cdot), \nu^{(1)}(\cdot), \nu^{(2)}(\cdot), \tilde{\tau}).$$

Let the Skorokhod representation be used, and fix any ω for which there is convergence. Let $\nu_s(\cdot), \nu_s^{(1)}(\cdot)$, and $\nu_s^{(2)}(\cdot)$ be the derivatives of $\nu(\cdot), \nu^{(1)}(\cdot)$, and $\nu^{(2)}(\cdot)$, respectively.

Fix $n < N$, and assume for now that $x(t_n) = \phi^\epsilon(t_n)$. It will be proved below that this implies $x(t_{n+1}) = \phi^\epsilon(t_{n+1})$, and therefore this assumption

will be justified by induction. First consider the case $(\phi^\epsilon(t))_1 \neq 0$ for all $t \in (t_n, t_{n+1})$. Then by the definition of the control scheme

$$\nu_s(\{\alpha_n\}) = 1 \text{ a.s. for } s \in (t_n, t_{n+1}).$$

Therefore, $x(t_{n+1}) = \phi^\epsilon(t_{n+1})$ in this case. Next consider the case $(\phi^\epsilon(t))_1 = 0$ for all $t \in (t_n, t_{n+1})$. The definition of the control scheme $u^{h,\delta}(\cdot)$ over the interpolated time (t_n, t_{n+1}) implies

$$\begin{aligned} \int_{\mathbf{R}^k} \nu_s^{(1)}(d\alpha) > 0 &\Rightarrow \int_{\mathbf{R}^k} \alpha \nu_s^{(1)}(d\alpha) / \int_{\mathbf{R}^k} \nu_s^{(1)}(d\alpha) = \alpha_n^{(1)}, \\ \int_{\mathbf{R}^k} \nu_s^{(2)}(d\alpha) > 0 &\Rightarrow \int_{\mathbf{R}^k} \alpha \nu_s^{(2)}(d\alpha) / \int_{\mathbf{R}^k} \nu_s^{(2)}(d\alpha) = \alpha_n^{(2)} \end{aligned} \quad (4.21)$$

a.s. for $s \in (t_n, t_{n+1})$. An elementary stability argument that uses the Lyapunov function $f(x) = |(x)_1|$ and (4.9), (4.12), (4.18), and (4.21) shows that $(x(s))_1 = 0$ for $s \in (t_n, t_{n+1})$. Therefore,

$$(\dot{x}(s))_1 = \left(\int_{\mathbf{R}^k} \alpha \nu_s^{(1)}(d\alpha) + \int_{\mathbf{R}^k} \alpha \nu_s^{(2)}(d\alpha) \right)_1 = 0. \quad (4.22)$$

Combining (4.21) and (4.22), we see that

$$\left(\int_{\mathbf{R}^k} \nu_s^{(1)}(d\alpha) \right) (\alpha_n^{(1)})_1 + \left(\int_{\mathbf{R}^k} \nu_s^{(2)}(d\alpha) \right) (\alpha_n^{(2)})_1 = 0.$$

Note that under (4.18), $\rho_n^{(1)}$ and $\rho_n^{(2)}$ are uniquely determined by $\rho_n^{(1)} + \rho_n^{(2)} = 1$ and

$$\rho_n^{(1)}(\alpha_n^{(1)})_1 + \rho_n^{(2)}(\alpha_n^{(2)})_1 = 0.$$

This uniqueness implies $\int_{\mathbf{R}^k} \nu_s^{(j)}(d\alpha) = \rho_n^{(j)}$, and, therefore, $\dot{x}(s) = \rho_n^{(1)} \alpha_n^{(1)} + \rho_n^{(2)} \alpha_n^{(2)} = \alpha_n$ a.s. for $s \in (t_n, t_{n+1})$. Since we have assumed $x(t_n) = \phi^\epsilon(t_n)$, this implies $x(t_{n+1}) = \phi^\epsilon(t_{n+1})$. By induction, $x(s) = \phi^\epsilon(s)$ for $s \in [0, T]$. Together with $\phi^\epsilon(t) \in G^0$ for $t \in [0, T]$, this implies $\tilde{\tau} > T$.

We now use the properties of the $\nu_s^{(j)}(\cdot)$ shown in the previous paragraph to prove the upper bound. For the given control scheme, we have

$$\begin{aligned} &\int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbf{R}^k} \bar{k}(\xi^{h,\delta}(s), \alpha) m^{h,\delta}(d\alpha ds) + g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}) \\ &= \int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbf{R}^k} k^{(1)}(\xi^{h,\delta}(s), \alpha) \nu^{(1),h,\delta}(d\alpha \times D^k [0, T] \times ds) \\ &\quad + \int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbf{R}^k} k^{(2)}(\xi^{h,\delta}(s), \alpha) \nu^{(2),h,\delta}(d\alpha \times D^k [0, T] \times ds) \\ &\quad + g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}). \end{aligned}$$

The boundedness of $\cup_{n=0}^{N-1} \{\alpha_n^{(1)}, \alpha_n^{(2)}\}$ and the dominated convergence theorem then give

$$\begin{aligned} & \lim_{(h,\delta) \rightarrow 0} \left[\int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbb{R}^k} \bar{k}(\xi^{h,\delta}(s), \alpha) m^{h,\delta}(d\alpha ds) \right. \\ & \quad \left. + g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}) \right] \\ &= \int_0^T \int_{\mathbb{R}^k} \left[k^{(1)}(x(s), \alpha) \nu_s^{(1)}(d\alpha) + k^{(2)}(x(s), \alpha) \nu_s^{(2)}(d\alpha) \right] \\ & \quad + g(x(T), T). \end{aligned} \quad (4.23)$$

Using the properties of the $\nu_s^{(j)}(\cdot)$ shown in the previous paragraph and the equality $x(\cdot) = \phi^\epsilon(\cdot)$, the right hand side of (4.23) is equal to

$$\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} f_n(t) dt + g(\phi^\epsilon(T), T), \quad (4.24)$$

where

$$f_n(t) = \begin{cases} k^{(1)}(\phi^\epsilon(t), \alpha_n) & \text{if } (\phi^\epsilon(t))_1 < 0 \\ k^{(2)}(\phi^\epsilon(t), \alpha_n) & \text{if } (\phi^\epsilon(t))_1 > 0 \\ \rho^{(1)} k^{(1)}(\phi^\epsilon(t), \alpha_n^{(1)}) + \rho^{(2)} k^{(2)}(\phi^\epsilon(t), \alpha_n^{(2)}) & \text{if } (\phi^\epsilon(t))_1 = 0. \end{cases}$$

The definition of the α_n and the $\alpha_n^{(i)}$ imply that (4.24) is bounded above by

$$\int_0^T k(\phi^\epsilon(s), \dot{\phi}^\epsilon(s)) ds + g(\phi^\epsilon(T), T) + T\epsilon \leq V(x) + (2 + T)\epsilon.$$

Note that (4.23) and (4.24) hold w.p.1. By again using the boundedness of $\cup_{n=0}^{N-1} \{\alpha_n^{(1)}, \alpha_n^{(2)}\}$ and the dominated convergence theorem, we have

$$\begin{aligned} & \limsup_{(h,\delta) \rightarrow 0} E_x^{m^{h,\delta}} \left[\int_0^{T \wedge \tau_{h,\delta}} \int_{\mathbb{R}^k} \bar{k}(\xi^{h,\delta}(s), \alpha) m^{h,\delta}(d\alpha ds) \right. \\ & \quad \left. + g(\xi^{h,\delta}(T \wedge \tau_{h,\delta}), T \wedge \tau_{h,\delta}) \right] \leq V(x) + (2 + T)\epsilon, \end{aligned}$$

and, therefore,

$$\limsup_{(h,\delta) \rightarrow 0} V^{h,\delta}(x) \leq V(x) + (2 + T)\epsilon.$$

This proves the upper bound for the case $\tau > T$.

The case $\tau \leq T$. The only difference between this case and the one treated previously concerns the limiting behavior of the exit times $\tau_{h,\delta}$. By redefining the ϵ -optimal path $\phi(\cdot)$ if need be, we can assume the existence

of $c > 0$ such that $\phi(t) \notin G$ for $t \in (\tau, \tau + c)$ and $\int_{\tau}^{\tau+c} k(\phi(s), \dot{\phi}(s))ds \leq \epsilon$. For arbitrary $\epsilon > 0$ and $\epsilon_1 > 0$, we can find a path $\phi^\epsilon(\cdot)$ according to Lemma 4.3 such that $\phi^\epsilon(t) \in G^0$ for $t \in [0, \tau - \epsilon_1]$ and $\phi^\epsilon(t) \notin G$ for $t \in [\tau, \tau + \epsilon_1]$. This allows us to control the time and location of exit. The proof now follows the same lines as for the case $\tau > T$. ■

Proof of Lemma 4.3. Let $\phi(\cdot)$ be absolutely continuous with $\phi(0) = x$ and $\int_0^T k(\phi(s), \dot{\phi}(s))ds < \infty$. Given $\epsilon > 0$, there exists $\gamma > 0$ such that $|s - t| \leq \gamma$ implies $|\phi(s) - \phi(t)| \leq \epsilon$. By assumption, $|x - y| \leq \epsilon$ implies $|k^{(i)}(x, \alpha) - k^{(i)}(y, \alpha)| \leq f(\epsilon)(M + k^{(i)}(x, \alpha))$ for all $\alpha \in \mathbb{R}^k$ and $i = 1, 2$. It is easy to show that the definition of $k^{(0)}(x, \alpha)$ in terms of $k^{(1)}(x, \alpha)$ and $k^{(2)}(x, \alpha)$ implies that $|k^{(0)}(x, \alpha) - k^{(0)}(y, \alpha)| \leq f(\epsilon)(M + k^{(0)}(x, \alpha))$ for all $\alpha \in \mathbb{R}^k$ and $|x - y| \leq \epsilon$. It is convenient to recall at this point that the definition also implies $k^{(0)}(x, \alpha) \leq k^{(1)}(x, \alpha) \wedge k^{(2)}(x, \alpha)$ whenever $(\alpha)_1 = 0$.

Define

$$\begin{aligned} A^{(0)} &= \{t : (\phi(t))_1 = 0\}, \\ A^{(1)} &= \{t : (\phi(t))_1 < 0\}, \\ A^{(2)} &= \{t : (\phi(t))_1 > 0\}. \end{aligned}$$

Let $B^{(0)}$ be a finite union of intervals $[c_j, d_j]$, $1 \leq j \leq J$, such that $d_i \leq c_{i+1}$, and such that $A^{(0)} \subset B^{(0)}$. Without loss of generality, we can assume

$$(\phi(c_i))_1 = (\phi(d_i))_1 = 0$$

for $1 \leq j \leq J$ and that $\max_j(d_j - c_j) < \gamma$. For simplicity, we assume $c_1 = 0$ and $d_J = T$. The required changes when this is not the case are slight. Define

$$u_j^0 = \frac{1}{d_j - c_j} \int_{c_j}^{d_j} \dot{\phi}(s)ds.$$

If $c_{j+1} > d_j$, let $e_j^k, k \leq K_j$, be such that $d_j = e_j^1 < e_j^2 < \dots < e_j^{K_j} = c_{j+1}$ and $\max_j(e_j^{k+1} - e_j^k) < \gamma$. Note that $(\phi(t))_1 \neq 0$ for $t \in (d_j, c_{j+1})$. Define

$$u_j^k = \frac{1}{e_j^{k+1} - e_j^k} \int_{e_j^k}^{e_j^{k+1}} \dot{\phi}(s)ds.$$

Finally, we set

$$\begin{aligned} u^\epsilon(t) &= \sum_{j=1}^J \left(I_{\{t \in [c_j, d_j]\}} u_j^0 + \sum_{k=1}^{K_j-1} I_{\{t \in [e_j^k, e_j^{k+1}]\}} u_j^k \right), \\ \phi^\epsilon(t) &= x + \int_0^t u^\epsilon(s)ds. \end{aligned}$$

With these definitions, $\phi(t) = \phi^\epsilon(t)$ whenever $t = c_j$, $t = d_j$, or $t = e_j^k$. Clearly, $\phi^\epsilon(\cdot) \rightarrow \phi(\cdot)$ as $\epsilon \rightarrow 0$.

Since $|\phi(s) - \phi(c_j)| \leq \epsilon$ for $s \in [c_j, d_j]$,

$$\begin{aligned} & \int_{c_j}^{d_j} k^{(0)}(\phi(c_j), \dot{\phi}(s)) ds \\ & \leq \int_{c_j}^{d_j} k(\phi(c_j), \dot{\phi}(s)) ds \\ & \leq \int_{c_j}^{d_j} \left(k(\phi(s), \dot{\phi}(s)) + f(\epsilon) [M + k(\phi(s), \dot{\phi}(s))] \right) ds. \end{aligned}$$

The first inequality in the last display is due to the convexity of each of the $k^{(i)}(x, \cdot)$, $i = 1, 2, 3$, and the fact that $k^{(0)}(x, \alpha) \leq k^{(1)}(x, \alpha) \wedge k^{(2)}(x, \alpha)$ whenever $(\alpha)_1 = 0$. Convexity and the definition of the u_j^0 imply that

$$k^{(0)}(\phi(c_j), u_j^0)(d_j - c_j) \leq \int_{c_j}^{d_j} k^{(0)}(\phi(c_j), \dot{\phi}(s)) ds.$$

Because $\phi(c_j) = \phi^\epsilon(c_j)$ and $|\phi^\epsilon(s) - \phi^\epsilon(c_j)| \leq \epsilon$ for $s \in [c_j, d_j]$,

$$\begin{aligned} & \int_{c_j}^{d_j} k^{(0)}(\phi^\epsilon(s), u^\epsilon(s)) ds \\ & \leq \left(k^{(0)}(\phi(c_j), u_j^0) + f(\epsilon) [M + k^{(0)}(\phi(c_j), u_j^0)] \right) (d_j - c_j). \end{aligned}$$

Combining these last three inequalities, we obtain

$$\begin{aligned} & \int_{c_j}^{d_j} k^{(0)}(\phi^\epsilon(s), u^\epsilon(s)) ds \\ & \leq (1 + (2 + f(\epsilon))f(\epsilon)) \int_{c_j}^{d_j} \left(k(\phi(s), \dot{\phi}(s)) + (2 + f(\epsilon))f(\epsilon)M \right) ds. \end{aligned}$$

A similar estimate applies for each interval of the form $[e_j^k, e_j^{k+1})$. Combining the estimates over the different intervals gives the last inequality in the statement of the lemma. Finally, the last sentence of the lemma is a consequence of the definitions of the c_j , d_j , and e_j^k . ■

14

The Viscosity Solution Approach to Proving Convergence of Numerical Schemes

In Chapters 10 to 13, we have shown the convergence of properly designed numerical approximations for a wide range of stochastic and deterministic optimal control problems. The approach to proving the convergence has been based on demonstrating the convergence of a sequence of controlled Markov chains to a controlled process (diffusion, jump diffusion, etc.) appropriate to the given stochastic or deterministic optimal control problem.

In this chapter, we will very briefly describe an alternative approach for proving the convergence of numerical schemes. The approach is based on what is referred to as the “viscosity solution” method (due to Crandall and Lions [28]) of defining and characterizing solutions to a wide class of partial differential equations. In particular, this method is applicable to many equations for which there are no classical sense solutions, a situation that is common for the PDE that are associated to stochastic and deterministic optimal control problems. The notion of solution allows an alternative method for proving the convergence of schemes for certain types of problems. In general, all the development of Chapters 4 to 8 that is concerned with deriving and solving numerical schemes applies here as well.

The application of viscosity solution methods to proving convergence of numerical approximations is currently a very active research area. Consequently, we will not give a detailed exposition of the methodology in its most general form. Rather, we will try to describe some of the basic ideas involved in using the method, and indicate the appropriate literature for the reader interested in learning more. The approach we describe follows Souganidis [112] and Barles and Souganidis [7].

For illustrative purposes, we will examine a control problem that involves a reflected diffusion process with a discounted running cost. Our example includes and gives an alternative approach to some of the “heavy traffic” problems discussed in Chapter 8. In Section 14.1, we give the definition of viscosity solutions for the associated Bellman equation. This equation was formally derived in Section 3.4. After stating the definition of the solution, we discuss the existence and uniqueness of solutions, as well as the relation-

ship of the solution to the value function of the stochastic control problem. Following this in Section 14.2 is a discussion of the key assumptions that are required, both of the solution to the equation and the numerical scheme itself, in order that a convergence proof be applicable. These assumptions are in some rough sense analogous to the main assumptions used by the Markov chain method, and some remarks along these lines are included. We conclude the chapter in Section 14.3 by exhibiting the proof of convergence for our illustrative example.

14.1 Definitions and Some Properties of Viscosity Solutions

We first describe the illustrative example. We consider a compact domain G , which for simplicity we take to be a rectangle of the form $\{x : c_i \leq x_i \leq d_i, i = 1, \dots, k\}$, where x_i denotes the i th component of $x \in \mathbb{R}^k$ and $c_i < d_i$. As our controlled process, we consider a controlled reflected diffusion of the type described in Section 1.4. Let $G_j, j = 1, \dots, 2k$, be the sets of the form $\{x : x_i \geq c_i\}$ and $\{x : x_i \leq d_i\}$ for $i = 1, \dots, k$. For each $j = 1, \dots, 2k$, we let $r_j : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be a Lipschitz continuous function such that $r_j(x) \neq 0$ for $x \in \partial G_j$. Let n_j be the inward normal to ∂G_j . Then we also assume $r_j(x)'n_j \geq 0$ for all $x \in \partial G_j$. We define the set of directions $r(\cdot)$ by

$$r(x) = \left\{ \sum_{j=1}^{2k} a_j r_j(x) : a_j \geq 0, a_j = 0 \text{ if } x \notin \partial G_j, \text{ and } \left| \sum_{j=1}^{2k} a_j r_j(x) \right| = 1 \right\}.$$

In other words, $r(x)$ is the intersection of the unit sphere with the closed convex cone generated by $r_j(x)$ for all j such that $x \in \partial G_j$. Our model then becomes

$$x(t) = x + \int_0^t b(x(s), u(s))ds + \int_0^t \sigma(x(s))dw(s) + z(t),$$

where $z(\cdot)$ satisfies the conditions of Definition 1.4.2.

For a cost criteria we use a discounted cost. For an admissible ordinary control $u(\cdot)$, the cost is

$$W(x, u) = E_x^u \left[\int_0^\infty e^{-\beta t} k(x(t), u(t))dt \right],$$

where $\beta > 0$. The value function is then

$$V(x) = \inf W(x, u), \quad (1.1)$$

where the infimum is over all admissible ordinary controls. Of course, conditions must be given that will guarantee that this problem is well defined. Precise conditions on $b(\cdot)$, $\sigma(\cdot)$, $r_j(\cdot)$, and $k(\cdot, \cdot)$ will be given below.

In Chapter 3 it was formally demonstrated that $V(x)$ satisfies the following Bellman equation with reflecting boundary condition:

$$\begin{cases} \inf_{\alpha \in \mathcal{U}} [\mathcal{L}^\alpha V(x) - \beta V(x) + k(x, \alpha)] = 0, \\ V_x(x)'r = 0 \text{ for } r \in r(x), x \in \partial G, \end{cases} \quad (1.2)$$

where for a twice continuously differentiable function $f(\cdot)$, we define

$$(\mathcal{L}^\alpha f)(x) = f'_x(x)b(x, \alpha) + \frac{1}{2}\text{tr}[f_{xx}(x)a(x)],$$

and where $a(\cdot) = \sigma(\cdot)\sigma'(\cdot)$. In general, equations of this type need not have any classical sense solutions (i.e., solutions which satisfy the equation and boundary condition at all points in G). Thus, one is tempted to pose a weaker notion of solution, e.g., to require satisfaction of the equation and boundary condition save on a subset of G having Lebesgue measure zero. One can show that the minimal cost function defined by (1.1) is a solution of this type (under the conditions given below). However, with such a definition, equations such as (1.2) may have many solutions. One of the successes of the theory of viscosity solutions is that for a large class of such equations it gives a notion of solution that is weak enough that solutions will exist, and strong enough that uniqueness can also be guaranteed. It also turns out [47, 91] that for many problems one can prove that the value function for an optimal control problem is a viscosity solution for the appropriate Bellman equation. In the presence of a uniqueness result, this gives a useful alternative characterization of the value function.

The theory of viscosity solutions was first developed in the context of first order nonlinear PDE [26, 28] and has since been extended in many directions, including fully nonlinear second order partial differential equations [66, 91]. For a recent overview of the theory as well as a more complete list of references than is given here, the reader may consult the survey paper of Crandall, Ishii, and Lions [27].

Definition of Viscosity Solutions. There are several equivalent definitions of viscosity solutions to (1.2), and the particular definition that is used may depend on the intended application. We will use the following. Let S^k denote the set of real valued symmetric $k \times k$ matrices. For $v \in I\!\!R$, $x \in G$, $p \in I\!\!R^k$, and $X \in S^k$, we define the function

$$F(x, v, p, X) = \sup_{\alpha \in \mathcal{U}} \left[-\frac{1}{2}\text{tr}[Xa(x)] - p'b(x, \alpha) + \beta v - k(x, \alpha) \right].$$

[We have reversed the sign in (1.2) to follow a standard convention in the literature.] Owing to the possible degeneracy of $a(\cdot)$, it may be the case that the boundary conditions are not satisfied in any normal sense for parts of ∂G . In part, this motivates the following form in which the boundary

conditions are incorporated into the definition. We set

$$\begin{aligned} F^*(x, v, p, X) &= F_*(x, v, p, X) = F(x, v, p, X) \text{ for } x \in G^0, \\ F^*(x, v, p, X) &= F(x, v, p, X) \vee \max\{-r'p : r \in r(x)\} \\ F_*(x, v, p, X) &= F(x, v, p, X) \wedge \min\{-r'p : r \in r(x)\} \end{aligned} \quad \left. \right\} \text{ for } x \in \partial G.$$

Definition. An upper semicontinuous function $V(\cdot)$ on G is called a viscosity subsolution if the following condition holds. If $\phi(\cdot) \in C^2(G)$ and if $x_0 \in G$ is a local maximum point of $V(\cdot) - \phi(\cdot)$, then

$$F_*(x_0, V(x_0), \phi_x(x_0), \phi_{xx}(x_0)) \leq 0. \quad (1.3)$$

Similarly, a lower semicontinuous function $V(\cdot)$ on G is called a viscosity supersolution if whenever $\phi(\cdot) \in C^2(G)$ and $x_0 \in G$ is a local minimum point of $V(\cdot) - \phi(\cdot)$, then

$$F^*(x_0, V(x_0), \phi_x(x_0), \phi_{xx}(x_0)) \geq 0. \quad (1.4)$$

A continuous function $V(\cdot)$ is a viscosity solution if it is both a subsolution and a supersolution.

Remarks. Note that in the definition given above subsolutions and supersolutions are required to be only semicontinuous and not continuous. This turns out to allow a considerable simplification in the proof of convergence of schemes. The technique we will use originates in the papers of Barles and Perthame [5, 6] and Ishii [63, 62].

Because in the remainder of the chapter we consider only viscosity solutions (respectively, viscosity subsolutions and viscosity supersolutions), we will drop the viscosity term and refer to such functions simply as solutions (respectively, subsolutions and supersolutions).

The key properties we will require of solutions to (1.2) are summarized as the conclusion to Theorem 1.1 below. Before stating the theorem, we list the properties of $b(\cdot, \cdot)$, $\sigma(\cdot)$, $r_j(\cdot)$, and $k(\cdot, \cdot)$ that will be used. These conditions are more than sufficient to guarantee the existence of a weak sense solution to our stochastic differential equation with reflection that is unique in the weak sense [35].

A1.1. *The set \mathcal{U} is compact. The functions $b(\cdot, \cdot)$ and $k(\cdot, \cdot)$ are continuous on $G \times \mathcal{U}$, and $b(\cdot, \alpha)$ is Lipschitz continuous, uniformly in $\alpha \in \mathcal{U}$. The function $\sigma(\cdot)$ is Lipschitz continuous on G .*

A1.2. *For each $x \in \partial G$, the convex hull of the set $r(x)$ does not contain the origin. For each $x \in \partial G$, let $J(x) = \{j : x \in \partial G_j\}$. We may assume without loss of generality that $J(x) = \{1, \dots, m\}$ for some $m \leq k$. Set*

$v_{ij} = |n'_i r_j(x)| - \delta_{ij}$, where $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$. Then for each $x \in \partial G$, we assume that the spectral radius of the $m \times m$ matrix $V = (v_{ij})$ is strictly less than one.

Remark. The spectral radius assumption (A1.2) is common in the literature related to heavy traffic problems [57, 86]. However, it is significantly stronger than the “completely– S ” condition discussed in Chapter 5.

Theorem 1.1. *Assume (A1.1) and (A1.2). Then the following comparison result holds. If $V^*(\cdot)$ and $V_*(\cdot)$ are respectively a subsolution and supersolution to (1.2), then*

$$V^*(x) \leq V_*(x) \quad (1.5)$$

for all $x \in G$. Define $V(x)$ by (1.1). If $V(\cdot)$ is continuous, then it is a solution to (1.2).

Remarks. Note that if a solution to (1.2) exists, then, by (1.5), it must be unique. If $V_1(\cdot)$ and $V_2(\cdot)$ are both solutions, then since they are also subsolutions and supersolutions, we have

$$V_1(x) \leq V_2(x) \leq V_1(x)$$

for all $x \in G$. There are many sets of conditions that will guarantee that $V(\cdot)$ defined by (1.1) is continuous. Although we will not prove it, the continuity of $V(\cdot)$ follows under (A1.1) and (A1.2). A proof can be based on the methods of weak convergence and the uniqueness results given in [35]. We note that the proof of continuity of $V(\cdot)$ must be based on the representation (1.1). The continuity of $V(\cdot)$ is related to the continuity of the total cost under the measure on the path space induced by an optimally (or ϵ -optimally) controlled process. This continuity has been an important consideration for the probabilistic approach described previously.

In keeping with the objectives of this chapter, we will not give a proof of Theorem 1.1. Instead, we will simply piece together some existing results in the literature. The comments that are given are intended to outline what is needed to apply uniqueness results for a class of PDE to prove convergence of numerical schemes for a related control problem. The assumptions we have made on G and $r(\cdot)$ make the setup used here a special case of the general result given in [37]. These conditions relate directly to the reflected diffusion and need no further discussion.

In the proof of an inequality such as (1.5), it is most natural to place conditions on $F(\cdot, \cdot, \cdot, \cdot)$. It is proved in [37] that the comparison principle of Theorem 1.1 holds under the following conditions.

1. For all $x \in G$, $v \in \mathbb{R}$, $p \in \mathbb{R}^k$, and $X, Y \in S^k$ with $X \geq Y$,

$$F(x, v, p, X) \leq F(x, v, p, Y). \quad (1.6)$$

2.

$$F(\cdot, \cdot, \cdot, \cdot) \text{ is continuous on } G \times \mathbb{R} \times \mathbb{R}^k \times S^k. \quad (1.7)$$

3. There is a continuous function $m_1 : [0, \infty) \rightarrow \mathbb{R}$ satisfying $m_1(0) = 0$ such that for all $\theta \geq 1$, $x, y \in G$, $v \in \mathbb{R}$, $p \in \mathbb{R}^k$, $X, Y \in S^k$,

$$F(y, v, p, -Y) - F(x, v, p, X) \leq m_1(|x - y|(|p| + 1) + \theta|x - y|^2) \quad (1.8)$$

whenever

$$-\theta \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq \theta \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}$$

in the sense of symmetric matrices (here I is the $k \times k$ unit matrix).

4. There is an open neighborhood U of ∂G (relative to G) and a continuous function $m_2 : [0, \infty) \rightarrow \mathbb{R}$ satisfying $m_2(0) = 0$ such that for all $x \in U$, $v \in \mathbb{R}$, $p, q \in \mathbb{R}^k$, $X, Y \in S^k$,

$$|F(x, v, p, Y) - F(x, v, q, X)| \leq m_2(|p - q| + \|X - Y\|). \quad (1.9)$$

(As noted previously there are many equivalent definitions of viscosity solution. In particular, the definition used in [37] is equivalent to the one used here.)

Thus, in order to make use of the results of [37], we must describe conditions on the components appearing in the statement of the control problem ($b(\cdot)$, $\sigma(\cdot)$ and $k(\cdot, \cdot)$) that are sufficient to guarantee (1.6)–(1.9). Property 1 is called “degenerate ellipticity” and follows easily from the definition of $F(\cdot, \cdot, \cdot, \cdot)$ (it also is a consequence of property 3, as remarked in [27]). Properties 2 and 4 are rather simple consequences of (A1.1). Property 3 is the most difficult to verify, and, in particular, requires the Lipschitz continuity conditions given in (A1.1). We omit the proof, and instead refer the reader to [27].

The last issue to be resolved is whether or not $V(\cdot)$ as defined by (1.1) is indeed a solution to (1.2). It turns out that the main difficulty here is in verifying that $V(\cdot)$ satisfies the dynamic programming principle: for $\Delta > 0$ and $x \in G$,

$$V(x) = \inf E_x^u \left[\int_0^\Delta e^{-\beta t} k(x(t), u(t)) dt + e^{-\beta \Delta} V(x(\Delta)) \right] \quad (1.10)$$

(where the infimum is over all admissible controls).

If we assume that $V(\cdot)$ is continuous and that (1.10) holds, then minor modifications of the proof of Theorem 3.1 in [47, Chapter 5] show that $V(\cdot)$ is indeed a solution of (1.2).

Thus, all that needs to be verified is (1.10). For our particular problem the process is never stopped, and the dynamic programming equation follows from the Markov property and the definition of $V(\cdot)$.

14.2 Numerical Schemes

We retain the notation of the previous chapters. Thus S_h is a grid in \mathbb{R}^k , $G_h = S_h \cap G$, and $G_h^0 = S_h \cap G^0$. We let ∂G_h^+ (the “discretized reflecting boundary”) be a subset of $S_h - G_h^0$ that satisfies $\sup_{x \in \partial G_h^+} \inf_{y \in G} |x - y| \rightarrow 0$ as $h \rightarrow 0$. In this chapter we will consider numerical approximations that are defined by a collection of relations of the form

$$S^h(x, V^h(x), V^h(\cdot)) = 0, \quad x \in G_h^0 \cup \partial G_h^+. \quad (2.1)$$

Of course, most of the numerical schemes of the previous chapters are of this form. In previous chapters, the assumptions on the scheme have been phrased in terms of the local consistency of an associated Markov chain. Here, we will impose assumptions directly on $S^h(\cdot, \cdot, \cdot)$. In the assumptions that are given below, $u_1(\cdot)$ and $u_2(\cdot)$ will denote real valued functions on $G_h^0 \cup \partial G_h^+$. In order to make sense of $S^h(x, \cdot, \cdot)$ for $x \in \partial G_h^+$, it will be necessary to extend the domain on which $F^*(\cdot, \cdot, \cdot, \cdot)$ and $F_*(\cdot, \cdot, \cdot, \cdot)$ are defined. To do this, we will assume (as in Chapter 5) the existence of an extension of $r(\cdot)$ [again denoted by $r(\cdot)$] from ∂G to an open neighborhood G_1 of ∂G , such that $r(\cdot)$ is upper semicontinuous: if $x_n \rightarrow x$ and $r_n \rightarrow r$, with $r_n \in r(x_n)$, then $r \in r(x)$. Such an extension can easily be constructed given the special form of G and the definition of $r(\cdot)$. We then define

$$\left. \begin{aligned} F^*(x, v, p, X) &= \max\{-r'p : r \in r(x)\} \\ F_*(x, v, p, X) &= \min\{-r'p : r \in r(x)\} \end{aligned} \right\} \text{ for } x \in G_1 - G.$$

The extensions $F^*(\cdot, \cdot, \cdot, \cdot)$ and $F_*(\cdot, \cdot, \cdot, \cdot)$ retain the upper semicontinuity and lower semicontinuity properties, respectively.

We can now state the assumptions on $S^h(\cdot, \cdot, \cdot)$. We assume $h > 0$ is small enough that $\partial G_h^+ \subset G_1$.

A2.1. If $u_1(\cdot) \geq u_2(\cdot)$, then for all $h > 0$, $x \in G_h^0 \cup \partial G_h^+$, and $v \in \mathbb{R}$, we have

$$S^h(x, v, u_1(\cdot)) \leq S^h(x, v, u_2(\cdot)).$$

A2.2. For all $h > 0$, there is a unique solution $V^h(\cdot)$ to (2.1), and this solution has a bound that is independent of h for small $h > 0$:

$$\limsup_{h \rightarrow 0} \sup_{x \in G_h^0 \cup \partial G_h^+} |V^h(x)| < \infty.$$

A2.3. There is consistency in the following sense. For any $\delta \in \mathbb{R}$ and $u(\cdot)$ defined on $G_h^0 \cup \partial G_h^+$, we define $f(\cdot) = u(\cdot) + \delta$ by $f(x) = u(x) + \delta$. Suppose we are given any $\phi(\cdot) \in C^2(G_1)$, any $x \in G$, any sequence $\{y^h, h > 0\}$ satisfying $y^h \in G_h^0 \cup \partial G_h^+$ and $y^h \rightarrow x$, and any sequence $\delta^h \rightarrow 0 \in \mathbb{R}$. Then

$$\limsup_{h \rightarrow 0} S^h(y^h, \phi(y^h) + \delta^h, \phi(\cdot) + \delta^h) \leq F^*(x, \phi(x), \phi_x(x), \phi_{xx}(x))$$

and

$$\liminf_{h \rightarrow 0} S^h(y^h, \phi(y^h) + \delta^h, \phi(\cdot) + \delta^h) \geq F_*(x, \phi(x), \phi_x(x), \phi_{xx}(x)).$$

Remark. Given x, y^h , and δ^h as in the statement of (A2.3), but with $\delta^h \rightarrow \delta$, it follows that

$$\limsup_{h \rightarrow 0} S^h(y^h, \phi(y^h) + \delta^h, \phi(\cdot) + \delta^h) \leq F^*(x, \phi(x) + \delta, \phi_x(x), \phi_{xx}(x))$$

and

$$\liminf_{h \rightarrow 0} S^h(y^h, \phi(y^h) + \delta^h, \phi(\cdot) + \delta^h) \geq F_*(x, \phi(x) + \delta, \phi_x(x), \phi_{xx}(x)).$$

This follows from (A2.3) by replacing $\phi(\cdot)$ by $\tilde{\phi}(\cdot) = \phi(\cdot) + \delta$.

Remarks on the Assumptions. In both the viscosity solution and Markov chain approaches to the convergence of schemes, two different types of assumptions appear. The first category may be labeled as “conditions on the limit,” where limit refers to a limit controlled process in the Markov chain case and a limit PDE in the viscosity solution case. In the Markov chain approach, we usually require weak sense uniqueness of an ϵ -optimal controlled process. For the viscosity solution approach we need a comparison principle as in Theorem 1.1, which implies the uniqueness of solutions.

The second category of assumptions are conditions that are placed directly on the scheme itself, i.e. (A2.1)–(A2.3). Here we will make a few remarks on this type of assumption. An obvious question to ask is how might functions $S^h(\cdot, \cdot, \cdot)$ satisfying (A2.1)–(A2.3) be found. One possibility is to examine the numerical schemes constructed via the Markov chain approach in Chapter 5. In the context of our example, we have transition probabilities $p^h(x, y|\alpha)$ and interpolation times $\Delta t^h(x, \alpha)$ that satisfy the proper local consistency properties: if $\{\xi_i^h, i < \infty\}$ is the controlled Markov chain that uses these probabilities, then

$$E_{x,n}^{h,\alpha} \Delta \xi_n^h = b(x, \alpha) \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)), \quad (2.2)$$

$$\text{cov}_{x,n}^{h,\alpha} \Delta \xi_n^h = \sigma(x) \Delta t^h(x, \alpha) + o(\Delta t^h(x, \alpha)) \quad (2.3)$$

for $x \in G_h^0$; there are $\epsilon_1 > 0$, $c_1 > 0$ and $c_2(h) \rightarrow 0$ as $h \rightarrow 0$ such that

$$E_{x,n}^{h,\alpha} \Delta \xi_n^h \in \{\theta \gamma + o(h) : c_2(h) \geq \theta \geq c_1 h, \gamma \in r(x)\}, \quad (2.4)$$

$$\text{cov}_{x,n}^{h,\alpha} \Delta \xi_n^h = O(h^2), \quad (2.5)$$

$$p^h(x, G_h^0) \geq \epsilon_1, \quad \Delta t^h(x, \alpha) = 0 \quad (2.6)$$

for all $h > 0$ and $x \in \partial G_h^+$. We can then use an approximation $V^h(\cdot)$ defined by

$$V^h(x) = \inf_{\alpha \in \mathcal{U}} \left[e^{-\beta \Delta t^h(x, \alpha)} \sum_y p^h(x, y | \alpha) V^h(y) + \Delta t^h(x, \alpha) k(x, \alpha) \right].$$

Approximating $e^{-\beta \Delta t^h(x, \alpha)}$ by $1/(1 + \beta \Delta t^h(x, \alpha))$ (as discussed in Chapter 5), we can also use the scheme

$$V^h(x) = \inf_{\alpha \in \mathcal{U}} \left[\sum_y p^h(x, y | \alpha) V^h(y) + \Delta t^h(x, \alpha) (-\beta V^h(x) + k(x, \alpha)) \right].$$

We can rewrite this last equation as

$$\sup_{\alpha \in \mathcal{U}} \left[\frac{-1}{\Delta t^h(x, \alpha)} \sum_y p^h(x, y | \alpha) [V^h(y) - V^h(x)] + \beta V^h(x) - k(x, \alpha) \right] = 0. \quad (2.7)$$

This suggests the definition

$$S^h(x, v, u(\cdot)) = \sup_{\alpha \in \mathcal{U}} \left[\frac{-1}{\Delta t^h(x, \alpha)} \sum_y p^h(x, y | \alpha) [u(y) - v] + \beta v - k(x, \alpha) \right].$$

The property (A2.1) is clearly a consequence of the non-negativity of the $p^h(x, y | \alpha)$ and $\Delta t^h(x, \alpha)$. Using the interpretation of $V^h(x)$ as the value function for a controlled Markov chain, one can show (A2.2). Finally (A2.3) follows from the local consistency conditions (2.2)–(2.6).

Thus, the schemes one is naturally led to via the Markov chain approach satisfy (A2.1)–(A2.3) (at least for the given example). However, a method that is perhaps more natural from the PDE point of view is to substitute “finite difference” approximations into (1.2) that are based on evaluating $V(\cdot)$ at the grid points $G_h^0 \cup \partial G_h^+$, and then try to draw out the restrictions placed on the approximations by (A2.1)–(A2.3). For example, consider a point $x \in G^0$. Suppose that finite difference approximations based on the values $V^h(y), y \in G_h^0 \cup \partial G_h^+$, are substituted into

$$F(x, V(x), V_x(x), V_{xx}(x)) = 0.$$

In particular, let us assume that the finite difference approximations are linear functions of differences of the form $V(y) - V(x)$. We thereby obtain a relation of the form

$$\sup_{\alpha \in \mathcal{U}} \left[- \sum_y q^h(x, y | \alpha) [V^h(y) - V^h(x)] - \beta V^h(x) + k(x, \alpha) \right] = 0, \quad (2.8)$$

where the $q^h(x, y | \alpha)$ depend on $b(\cdot, \cdot)$ and $\sigma(\cdot)$. From (A2.1), we deduce that $q^h(x, y | \alpha) \geq 0$. Define $Q^h(x, \alpha) = \sum_y q^h(x, y | \alpha)$. If (2.8) is to be

meaningful, then we would expect $Q^h(x, \alpha) \neq 0$. By defining $\Delta t^h(x, \alpha) = 1/Q^h(x, \alpha)$ and $p^h(x, y|\alpha) = q^h(x, y|\alpha)/Q^h(x, \alpha)$, we put (2.8) into the form of (2.7). Note that the $p^h(x, y|\alpha)$ so defined are the transition probabilities of a controlled Markov chain.

Owing to the presence of the supremum operation in (2.7) it is difficult to directly draw a conclusion on the properties that the $p^h(x, y|\alpha)$ must satisfy. However, by choosing the linear function $\phi(x) = p'x$, we see that (A2.3) implies

$$\begin{aligned} & \lim_{h \rightarrow 0} \sup_{\alpha \in \mathcal{U}} \left[\frac{-1}{\Delta t^h(x, \alpha)} \sum_y p^h(x, y|\alpha)(y - x)'p - k(x, \alpha) \right] \\ &= \sup_{\alpha \in \mathcal{U}} [-b(x, \alpha)'p - k(x, \alpha)]. \end{aligned} \quad (2.9)$$

If we make the plausible assumption that in order to achieve (2.9) we will need

$$\lim_{h \rightarrow 0} \frac{-1}{\Delta t^h(x, \alpha)} \sum_y p^h(x, y|\alpha)(y - x)'p = b(x, \alpha)'p$$

uniformly in $\alpha \in \mathcal{U}$, then (A2.3) would require (2.2). By using quadratic functions, we see that (A2.3) would also require (2.3).

This suggests that although it may not be necessary in every case we would still expect, in general, that a finite difference scheme chosen to satisfy (A2.1)–(A2.3) would also produce a Markov chain that is locally consistent, at least for the points in G_h^0 . Analogous remarks hold for points in ∂G_h^+ .

14.3 Proof of Convergence

Theorem 3.1. *Assume (A1.1), (A1.2), (A2.1), (A2.2), (A2.3), and that $V(\cdot)$ defined by (1.1) is continuous. Then for $V^h(\cdot)$ defined by (2.1), we have*

$$V^h(x) \rightarrow V(x).$$

Proof. For $x \in G$, we define

$$V^*(x) = \limsup_{\delta \rightarrow 0} \limsup_{h \rightarrow 0} \sup \{V^h(y) : |x - y| \leq \delta, y \in G_h^0 \cup \partial G_h^+ \},$$

$$V_*(x) = \liminf_{\delta \rightarrow 0} \liminf_{h \rightarrow 0} \inf \{V^h(y) : |x - y| \leq \delta, y \in G_h^0 \cup \partial G_h^+ \}.$$

These definitions imply that $V^*(\cdot)$ is upper semicontinuous, $V_*(\cdot)$ is lower semicontinuous, and $V^*(\cdot) \geq V_*(\cdot)$. We will prove that $V^*(\cdot)$ is a subsolution and that $V_*(\cdot)$ is a supersolution. Assuming for now that this is true, we can apply Theorem 1.1 and conclude

$$V^*(x) \leq V(x) \leq V_*(x) \leq V^*(x).$$

Thus, we obtain the conclusion of the theorem.

We now prove that $V^*(\cdot)$ is a subsolution. The proof that $V_*(\cdot)$ is a supersolution is essentially the same and omitted. Let $\phi(\cdot) \in C^2(G_1)$, and assume that $x_0 \in G$ is a maximum point of $V^*(\cdot) - \phi(\cdot)$. We wish to show that

$$F_*(x_0, V^*(x_0), \phi_x(x_0), \phi_{xx}(x_0)) \leq 0. \quad (3.1)$$

We first note that by the continuity given in (1.7), (3.1) follows if it can be proved for functions of the form

$$\phi_\delta(x) = \phi(x) + \delta|x - x_0|^2, \delta > 0.$$

This allows us to assume without loss that x_0 is a strict local maximum of $V^*(\cdot) - \phi(\cdot)$. By (A2.2) we can further assume that x_0 is a strict maximum on G by suitably redefining $\phi(\cdot)$ outside an appropriate open neighborhood of x_0 . Let x^h be a point in $G_h^0 \cup \partial G_h^+$ at which $V^h(\cdot) - \phi(\cdot)$ attains its maximum, relative to $G_h^0 \cup \partial G_h^+$. The definition of $V^*(\cdot)$ and the assumption that x_0 is a strict global maximum of $V^*(\cdot) - \phi(\cdot)$ imply the existence of a subsequence of h (again denoted by h) and $x^h \in G_1$ such that $x^h \rightarrow x_0$ and also $V^h(x^h) \rightarrow V^*(x_0)$. By the definition of $V^h(\cdot)$,

$$S^h(x^h, V^h(x^h), V^h(\cdot)) = 0. \quad (3.2)$$

Owing to the definition of x^h ,

$$\phi(\cdot) + [V^h(x^h) - \phi(x^h)] \geq V^h(\cdot).$$

This inequality, (A2.1), and (3.2) then yield

$$S^h(x^h, \phi(x^h) + [V^h(x^h) - \phi(x^h)], \phi(\cdot) + [V^h(x^h) - \phi(x^h)]) \leq 0,$$

where we interpret $\phi(\cdot)$ here as the restriction of $\phi(\cdot)$ to $G_h^0 \cup \partial G_h^+$. Using (A2.3) we have

$$\begin{aligned} & F_*(x_0, V^*(x_0), \phi_x(x_0), \phi_{xx}(x_0)) \\ & \leq \liminf_{h \rightarrow 0} S^h(x^h, \phi(x^h) + [V^h(x^h) - \phi(x^h)], \phi(\cdot) + [V^h(x^h) - \phi(x^h)]) \\ & \leq 0, \end{aligned}$$

which is (3.1). ■

References

- [1] M. Akian. Resolution numerique d'équations d'Hamilton-Jacobi-Bellman au moyen d'algorithmes multigrilles et d'itérations sur les politiques. In *Eighth Conference on Analysis and Optimization of Systems*, Antibes, France, 1988. INRIA.
- [2] M. Akian. *Méthodes Multigrilles en Contrôle Stochastique*. PhD thesis, University of Paris, 1990.
- [3] R. Anderson and S. Orey. Small random perturbations of dynamical systems with reflecting boundary. *Nagoya Math. J.*, 60:189–216, 1976.
- [4] J. Baras. Real time architecture for the Zakai equation and applications. In E. Mertzbach E. Mayer-Wolfe and A. Shwartz, editors, *Stochastic Analysis*, pages 15–38. Academic Press, New York, 1991.
- [5] G. Barles and B. Perthame. Discontinuous solutions of deterministic optimal stopping time problems. *Modél. Math. Anal. Num.*, 21:557–579, 1987.
- [6] G. Barles and B. Perthame. Exit time problems in optimal control and the vanishing viscosity method. *SIAM J. Control Optim.*, 26:1133–1148, 1988.
- [7] G. Barles and P. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *J. Asymptotic Analysis*, 4:271–283, 1991.
- [8] M.S. Bazaraa and J.J. Jarvis. *Linear Programming and Network Flows*. Wiley, New York, 1977.
- [9] L. Berkovitz. *Optimal Control Theory*, volume 12 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 1974.
- [10] D.P. Bertsekas. *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall, Englewood Cliffs, NJ, 1987.
- [11] D.P. Bertsekas and D.A. Castanon. Adaptive aggregation methods for infinite horizon dynamic programming. *IEEE Trans. Auto. Control*, 34:589–598, 1989.

- [12] P. Billingsley. *Convergence of Probability Measures*. John Wiley, New York, 1968.
- [13] E.-K. Boukas and A. Haurie. Manufacturing flow control and preventative maintenance: A stochastic control approach. *IEEE Trans. Auto. Control*, 35:1024–1031, 1990.
- [14] W.L. Briggs. *A Multigrid Tutorial*. SIAM, Philadelphia, 1987.
- [15] F. Campillo. Optimal ergodic control of nonlinear stochastic systems. Technical Report 1257, INRIA, Sophia Antipolis, France, 1990.
- [16] P. Chancelier, C. Gomez, J.-P. Quadrat, and A. Sulem. Automatic study in stochastic control. In W. Fleming and P.-L. Lions, editors, *Vol. 10, IMA Volumes in Math. and Its Applications*, Berlin, 1988. Springer-Verlag.
- [17] P. Chancelier et al. An expert system for control and signal processing with automatic FORTRAN program generation. In *Math. Systems Symp.*, Stockholm, 1986. Royal Inst of Technology.
- [18] F. Chatelin and W.L. Miranker. Acceleration by aggregation of successive approximation methods. *Linear Alg. Appl.*, 43:17–48, 1982.
- [19] P.L. Chow, J.-L. Menaldi, and M. Robin. Additive control of stochastic linear systems with finite horizons. *SIAM J. Control Optimiz.*, 23:858–899, 1985.
- [20] K.-L. Chung. *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, Berlin, 1960.
- [21] S.-L. Chung and F.B. Hanson. Optimization techniques for stochastic dynamic programming. In *Proceedings of the 29th IEEE Conference on Decision and Control*, Honolulu, New York, 1990. IEEE Publishers.
- [22] S.-L. Chung and F.B. Hanson. Parallel optimization for computational stochastic dynamic programming. In P.C. Yew, editor, *Proceedings of the 30th IEEE Conference on Parallel Processing*, University Park, PA, 1990. Penn State University Press.
- [23] S.-L. Chung, F.B. Hanson, and H.H. Xu. Finite element method for parallel stochastic dynamic programming. In *Proceedings of the 30th IEEE Conference on Decision and Control*, Bristol, England, New York, 1991. IEEE Publishers.
- [24] C. Costantini. *The Skorokhod oblique reflection problem and a diffusion approximation for a class of transport processes*. PhD thesis, University of Wisconsin, Madison, 1987.

- [25] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience, New York, 1937. (first english edition).
- [26] M.G. Crandall, L.C. Evans, and P.-L. Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 282:487–501, 1984.
- [27] M.G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, page to appear, 1992.
- [28] M.G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277:1–42, 1983.
- [29] G. Dahlquist and A. Bjorck. *Numerical Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1974.
- [30] J. Dai. *Steady state analysis of reflected Brownian motions: characterization, numerical methods and queueing applications*. PhD thesis, Operations Research Dept., Stanford University, 1990.
- [31] M.H.A. Davis and A.R. Norman. Portfolio selection with transactions costs. *Math. Oper. Res.*, 15:676–713, 1990.
- [32] C. Derman. Denumerable state Markov decision processes—average criterion. *Ann. Math. Statist.*, 37:1545–1554, 1966.
- [33] G.B. DiMasi and W.J. Runggaldier. Approximations and bounds for discrete time nonlinear filtering. In *Lecture Notes in Control and Information Sciences*, volume 44, pages 191–202. Springer-Verlag, Berlin, 1982.
- [34] P. Dupuis and R.S. Ellis. Large deviations for Markov processes with discontinuous statistics, II: Random walks. *Probab. Theory Rel. Fields*, 91:153–194, 1992.
- [35] P. Dupuis and H. Ishii. SDE’s with oblique reflection on nonsmooth domains. Technical report, University of Massachusetts, 1990. To appear in *The Annals of Probability*.
- [36] P. Dupuis and H. Ishii. On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stochastics*, 35:31–62, 1991.
- [37] P. Dupuis and H. Ishii. On oblique derivative problems for fully nonlinear second-order elliptic PDE’s on domains with corners. *Hokkaido Math. J.*, 20:135–164, 1991.

- [38] P. Dupuis, H. Ishii, and H.M. Soner. A viscosity solution approach to the asymptotic analysis of queueing systems. *Annals Probab.*, 18:226–255, 1990.
- [39] P. Dupuis and J. Oliensis. An optimal control formulation and related numerical methods for a problem in shape reconstruction. Technical report, Brown University, LCDS, Div. of Applied Math., 1992.
- [40] E.B. Dynkin. *Markov Processes*. Springer-Verlag, Berlin, 1965.
- [41] R. Elliot. *Stochastic Calculus and Applications*. Springer-Verlag, Berlin, 1982.
- [42] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.
- [43] A. Federgruen and P.J. Schweitzer. Discounted and nondiscounted value iteration Markov decision processes. In M.L. Puterman, editor, *Dynamic Programming and Its Applications*. Academic Press, New York, 1978.
- [44] W. Feller. *An Introduction to Probability Theory and its Applications, Volume 2*. Wiley, New York, 1966.
- [45] W.H. Fleming and M. Nisio. On stochastic relaxed control for partially observed diffusions. *Nagoya Math. J.*, 93:71–108, 1984.
- [46] W.H. Fleming and R. Rishel. *Deterministic and Stochastic Optimal Control*. Springer-Verlag, Berlin, 1975.
- [47] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, New York, 1992.
- [48] M.I. Freidlin and A.D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag, New York, 1984.
- [49] A. Friedman. *Stochastic Differential Equations and Applications*. Academic Press, New York, 1975.
- [50] S. J. Gass. *Linear Programming*. McGraw Hill, New York, NY, fifth edition, 1985.
- [51] H. Goldstein. *Classical Mechanics*. Addison-Wesley, Reading, MA, 1950.
- [52] G. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins Press, Baltimore, second edition, 1989.
- [53] W. Hackbusch. *Multigrid Methods and Applications*. Springer-Verlag, Berlin, 1985.

- [54] J. M. Harrison and V. Nguyen. The QNET method for two-moment analysis of open queueing systems. *Queueing Systems*, 6:1–32, 1990.
- [55] J.M. Harrison. *Brownian Motion and Stochastic Flow Systems*. Wiley, New York, 1985.
- [56] J.M. Harrison. Brownian models of queueing networks with heterogeneous customer populations. In W. Fleming and P.-L. Lions, editors, *Vol. 10, IMA Volumes in Math. and Its Applications*, pages 147–186, Berlin, 1988. Springer-Verlag.
- [57] J.M. Harrison and M.I. Reiman. Reflected Brownian motion on an orthant. *Annals Probab.*, 9:302–308, 1981.
- [58] B.K.P. Horn and Brooks. *Shape From Shading*. M.I.T. Press, Cambridge, MA, 1989.
- [59] R. Howard. *Dynamic Programming and Markov Processes*. M.I.T. Press, Cambridge, MA, 1960.
- [60] D.L. Iglehart and W. Whitt. Multiple channel queues in heavy traffic. *Adv. Appl. Probab.*, 2:150–177, 1970.
- [61] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam, 1981.
- [62] H. Ishii. Perron’s method for Hamilton-Jacobi equations. *Duke Math. J.*, 55:369–384, 1987.
- [63] H. Ishii. A boundary value problem of the Dirichlet type for Hamilton-Jacobi equations. *Ann. Sc. Norm. Sup. Pisa, (IV)*, 16:105–135, 1989.
- [64] J. Jacod. *Calcul Stochastique et Problèmes de Martingales*. Springer-Verlag, New York, 1979.
- [65] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, 1987.
- [66] R. Jensen. The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. *Arch. Rat. Mech. Anal.*, 101:1–27, 1988.
- [67] L.C.M. Kallenberg. Linear programming and finite markov control problems. Technical report, Mathematical Center, Amsterdam, 1983. Mathematical Center Tract no. 148.
- [68] I. Karatzas and S.E. Shreve. Equivalent models for finite fuel stochastic control. *Stochastics*, 18:245–276, 1986.

- [69] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, 1988.
- [70] S. Karlin and H.M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, New York, 1981.
- [71] T.G. Kurtz. *Approximation of Population Processes*, volume 36 of *CBMS-NSF Regional Conf. Series in Appl. Math.* SIAM, Philadelphia, 1981.
- [72] H. J. Kushner. Control of trunk line systems in heavy traffic. Technical report, Brown University, LCDS, Div. of Applied Math., 1992.
- [73] H.J. Kushner. Dynamical equations for nonlinear filtering. *J. Diff. Equations*, 3:179–190, 1967.
- [74] H.J. Kushner. *Introduction to Stochastic Control Theory*. Holt, Rinehart and Winston, New York, 1972.
- [75] H.J. Kushner. Probabilistic methods for finite difference approximation to degenerate elliptic and parabolic equations with Neumann and Dirichlet boundary conditions. *J. Math. Appl.*, 53:644–668, 1976.
- [76] H.J. Kushner. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. Academic Press, New York, 1977.
- [77] H.J. Kushner. Optimality conditions for the average cost per unit time problem with a diffusion model. *SIAM J. Control Optimization*, 16:330–346, 1978.
- [78] H.J. Kushner. A robust computable approximation to the optimal nonlinear filter. *Stochastics*, 3:75–83, 1979.
- [79] H.J. Kushner. *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic System Theory*. MIT Press, Cambridge, MA, 1984.
- [80] H.J. Kushner. Numerical methods for stochastic control problems in continuous time. *SIAM J. Control Optimization*, 28:999–1048, 1990.
- [81] H.J. Kushner. *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, volume 3 of *Systems and Control*. Birkhaeuser, Boston, 1990.
- [82] H.J. Kushner and G.B. DiMasi. Approximations for functionals and optimal control problems on jump-diffusion processes. *J. Math. Anal. Appl.*, 63:772–800, 1978.
- [83] H.J. Kushner and A.J. Kleinman. Accelerated procedures for the solution of discrete Markov control problems. *IEEE Trans. Auto. Control*, AC-16:147–152, 1971.

- [84] H.J. Kushner and A.J. Kleinman. Mathematical programming and the control of Markov chains. *Int. J. Control.*, 13:801–820, 1971.
- [85] H.J. Kushner and L.F. Martins. Numerical methods for stochastic singular control problems. *SIAM J. Control Optimization*, 29:1443–1475, 1991.
- [86] H.J. Kushner and K.M. Ramachandran. Optimal and approximately optimal control policies for queues in heavy traffic. *SIAM J. Control Optimization*, 27:1293–1318, 1989.
- [87] H.J. Kushner and W. Rungaldier. Nearly optimal state feedback controls for stochastic systems with wideband noise disturbances. *SIAM J. Control Optimization*, 25:289–315, 1987.
- [88] L.S. Lasdon. *Optimization Theory for Large Systems*. MacMillan, New York, 1960.
- [89] J.P. Lehoczky and S.E. Shreve. Absolutely continuous and singular stochastic control. *Stochastics*, 17:91–110, 1986.
- [90] A.J. Lemoine. Networks of queues: A survey of weak convergence results. *Management Science*, 24:1175–1193, 1978.
- [91] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 2: Viscosity solutions and uniqueness. *Comm. P. D. E.*, 8:1220–1276, 1983.
- [92] P.-L. Lions and A.-S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.*, 37:511–553, 1984.
- [93] R. Liptser and A.N. Shiryaev. *Statistics of Random Processes*. Springer-Verlag, Berlin, 1977.
- [94] J. Mandel, editor. *Proceedings of the Fourth Copper Mountain Conference on Multigrid Methods*, Philadelphia, 1989. SIAM.
- [95] L.F. Martins and H.J. Kushner. Routing and singular control for queueing networks in heavy traffic. *SIAM J. Control Optimization*, 28:1209–1233, 1990.
- [96] S. F. McCormack. *Multigrid Methods*. SIAM, Philadelphia, PA, 1987.
- [97] S.F. McCormack. *Multilevel Adaptive Methods for Partial Differential Equations*. SIAM, Philadelphia, 1989.
- [98] J. Oliensis and P. Dupuis. Direct method for reconstructing shape from shading. In *Proc. SPIE Conf. on Geometric Methods in Computer Vision*, pages 116–128, 1991.

- [99] M.J. Puterman and M.C. Shin. Modified policy algorithms for discounted Markov decision problems. *Management Science*, 24:1127–1137, 1978.
- [100] M.L. Puterman. Markov decision processes. In D.P. Heyman and M.J. Sobel, editors, *Stochastic Models, Volume 2*, chapter 8. North-Holland, Amsterdam, 1991.
- [101] J.-P. Quadrat. The discrete state linear quadratic problem. In *Proceedings of the 28th IEEE Conference on Decision and Control, Austin*, New York, 1989. IEEE Publishers.
- [102] M.I. Reiman. Open queueing networks in heavy traffic. *Math. Oper. Res.*, 9:441–458, 1984.
- [103] S. Ross. *Introduction to Stochastic Dynamic Programming*. Academic Press, New York, 1983.
- [104] S.M. Ross. Arbitrary state Markovian decision problems. *Ann. Math. Stat.*, 39:412–423, 1968.
- [105] E. Rouy and A. Tourin. A viscosity solutions approach to shape-from-shading. *SIAM J. Num. Analysis*, page to appear, 1992.
- [106] Y. Saisho. Stochastic differential equations for multi-dimensional domain with reflecting boundary. *Probab. Theory Rel. Fields*, 74:455–477, 1987.
- [107] J.N.T. Schult. Numerical solution of the hamilton-jacobi-bellman equation for a freeway traffic control problem. Technical report, Center for Mathematics and Computer Science, Amsterdam, 1989. Note BS N8901.
- [108] A.N. Shiryaev. Stochastic equations of nonlinear filtering of jump Markov processes. *Problemy Peredachi Informatsii*, 3:3–22, 1966.
- [109] A.N. Shiryaev. *Optimal Stopping Rules*. Springer-Verlag, Berlin, 1978.
- [110] A.N. Shirayev. *Probability*. Springer-Verlag, New York, 1984.
- [111] H.M. Soner and S.E. Shreve. Regularity of the value function for a two-dimensional singular stochastic control problem. *SIAM J. Control Optimization*, 27:876–907, 1989.
- [112] P.E. Souganidis. Approximation schemes for viscosity solutions of Hamilton-Jacobi equations. *J. Diff. Eq.*, 56:345–390, 1985.
- [113] J.C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Wadsworth and Brooks/Cole, Pacific Grove, CA, 1989.

- [114] D.W. Stroock and S.R.S. Varadhan. On degenerate elliptic and parabolic operators of second order and their associated diffusions. *Comm. Pure Appl. Math.*, 25:651–713, 1972.
- [115] D.W. Stroock and S.R.S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, New York, 1979.
- [116] M. Taksar. Average optimal singular control and a related stopping problem. *Math. Oper. Res.*, 10:63–81, 1985.
- [117] H. Tanaka. Stochastic differential equations with reflecting boundary conditions in convex regions. *Hiroshima Math. J.*, 9:163–177, 1979.
- [118] L. M. Taylor and R. J. Williams. Existence and uniqueness of semi-martingale reflecting Brownian motions in an orthant. Technical report, Department of Mathematics, University of California, San Diego, 1992.
- [119] L.M. Taylor. *Existence and Uniqueness of Semimartingale Reflecting Brownian Motions in an Orthant*. PhD thesis, University of California, San Diego, 1990.
- [120] H.C. Tijms. *Stochastic Modelling and Analysis: A Computational Approach*. Wiley, New York, 1986.
- [121] H.M. Wagner. *Principles of Operations Research*. Prentice-Hall, Englewood Cliffs, NJ, 1975.
- [122] J. Warga. Relaxed variational problems. *J. Math. Anal. Appl.*, 4:111–128, 1962.
- [123] L.M. Wein. Optimal control of a two station Brownian network. *Math. Oper. Res.*, 15:215–242, 1990.
- [124] D.J. White. Dynamic programming, Markov chains and the method of successive approximations. *J. Math. Anal. Appl.*, 6:373–376, 1963.
- [125] P. Whittle. *Optimization Over Time, Dynamic Programming and Stochastic Control*. Wiley, New York, 1983.
- [126] W.M. Wonham. Some applications of stochastic differential equations to optimal nonlinear filtering. *SIAM J. Control*, 2:347–369, 1965.
- [127] L.C. Young. Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Compt. Rend. Soc. Sci. et Lettres Varsovie CI III*, 30:212–234, 1937.
- [128] M. Zakai. On optimal filtering of diffusion processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 11:230–243, 1969.

List of Symbols

$a(\cdot)$	13, 71	$f_x(x)$	13
$a_h(\cdot)$	71	$f_{xx}(x)$	13
α	20	$F(\cdot)$	413
$b_h(\cdot)$	71	$F^*(\cdot)$	414
$\mathcal{B}(S)$	10	$F_*(\cdot)$	414
$\mathcal{B}(\mathcal{U})$	86	\mathcal{F}_t	8
$\mathcal{B}(\mathcal{U} \times [0, \infty))$	86	$\mathcal{F}(A)$	10
β	37, 56	G	23, 69
$c(h)$	359	G^0	54
cov	21	G_h^0	71
$C(S)$	248	Γ	28
$C_b(S)$	248	γ	40
$C_0(S)$	248	$\gamma(u)$	195
$C^n(U)$	12	$\bar{\gamma}$	195, 320
$C^k[0, T]$	8	$\gamma^h(u)$	204
$C^k[0, \infty)$	8	$\gamma(x, m)$	320
$D^k[0, T]$	8	$\bar{\gamma}^h$	206
$D^k[0, \infty)$	8	$h(\cdot)$	29
$C(u)$	153	$J(t)$	27
$C^h(u)$	146	$J_f(t)$	31
$\Delta t^h(x, \alpha)$	70	$J^h(t)$	132, 291
Δt_n^h	70	$k^\eta(x, \alpha)$	367
$\Delta \hat{t}^{h,\delta}$	331	$k^{(0)}(x, \alpha)$	378
$\Delta \hat{t}_n^{h,\delta}$	334	$\bar{k}(x, \alpha)$	399
$\Delta \tau_n^h$	75	λ	28, 127
$\Delta \hat{\tau}_n^h$	308	$\lambda(x)$	27, 128
$\Delta \psi^h(t)$	75	\mathcal{L}	15
$\Delta \xi_n^h$	71	\mathcal{L}^α	20, 33, 106, 132
\mathbf{e}	40	$\mathcal{L}^{u(\cdot)}$	20, 33
E_t	341	\mathcal{L}_ω	254, 275
$E_{\mathcal{F}_t}$	8	\mathcal{L}^*	339
E_x	36	$m(\cdot)$	86
E_x^α	48	$m^h(\cdot)$	289
E_x^u	48	$m^{h,\delta}(\cdot)$	361
$E_{x,n}^u$	51	$m_t(\cdot)$	86
$E_{x,n}^{h,\alpha}$	71	$\mu^h(x, u)$	205
$E_{x,t}^{h,\alpha}$	75	μ_η	23
$ \eta (\cdot)$	23	N_h	72
$\eta(h)$	368	$N_\gamma(A)$	371
		$N(\cdot)$	29, 127

$N(t, H)$	127	σ_n^h	296
$N_h(\cdot)$	256, 291	$\Sigma_b(T)$	11
$\bar{N}(\cdot)$	272	Σ_b	11
ν_n	28, 127	Σ_b^*	11
ν_n^h	130	t_n^h	72
$p(x, y)$	36	$\hat{t}_n^{h,\delta}$	334
$p(x, y \alpha)$	48	$\text{tr } B$	13
$p^h(x, y \alpha)$	70	T_{Ju}	162
$p^{h,\delta}(x, y \alpha)$	328	$T_{GS,u}$	163
$\hat{p}^{h,\delta}(x, n; y, m \alpha)$	331	T_u^h	182
$\tilde{p}^{h,\delta}(x, y)$	334	$\hat{T}^h(t)$	308
$P_{\mathcal{F}_t}$	8	τ	54
$P_{x,n}^{h,\alpha}$	75	$\bar{\tau}$	261
$P_{x,t}^{h,\alpha}$	75	τ_h	72
$\mathcal{P}(S)$	249	τ_n^h	75, 138
$\pi(\cdot)$	40	$\hat{\tau}_n^h$	261, 280
$\pi(\cdot, \cdot)$	249	u_n^h	70
$\pi^h(x, u)$	204	$u^h(\cdot)$	72
$\Pi(\cdot)$	28, 127	$u^{h,\delta}(\cdot)$	334
$\bar{\Pi}(x, H)$	27, 128	\mathcal{U}	19
∂G_h^0	145	$\mathcal{U}^{h,\delta}$	353
∂G_h^+	137	$w^h(t)$	258, 290
∂S	36	$x_\delta(\cdot)$	259
∂S^+	39	ξ_n	36
$(\phi, m)_t$	275	ξ_n^h	70
$\psi^h(\cdot)$	75, 130	$\xi^h(\cdot)$	72
$\hat{\psi}^h(\cdot)$	308	$z_\delta(\cdot)$	274
$q(\cdot)$	28, 127	ζ_n^h	332
$q_h(\cdot)$	129	$\hat{\zeta}_n^{h,\delta}$	332
$Q^h(x, \alpha)$	107, 109	$\hat{\zeta}_{n,0}^{h,\delta}$	332
$r(x)$	23, 134, 412	$\hat{\zeta}_{h,\delta}(t)$	334
$r(x, y)$	37	$\zeta_n^{h,\delta}$	334
$r(x, y \alpha)$	50	\Rightarrow	248
$r^h(x, y \alpha)$	146		
$R(u)$	153		
$R^h(u)$	146		
$\mathcal{R}(\mathcal{U} \times [0, \infty))$	265		
\mathbb{R}^r_h	106		
ρ_n	28, 127		
S	36		
S_0	43		
$ S $	36		
S_h	70		
$S^h(\cdot)$	417		

Index

- Absorbing boundary condition 54
- Accelerated Gauss-Seidel (SOR)
 - 165
- Accelerated Jacobi 164
- Accelerated methods, comparisons
 - 166ff
- Adapted 11
- Admissible control law 19, 32, 47,
 - 71, 76
- Aggregation method 172ff
- Approximation in policy space 154
 - ergodic cost case 199
 - modified 159
- Approximation in value space 156
 - ergodic 197
- Bellman equation, continuous time
 - discounted cost 80ff, 130
 - finite time problem 335
 - formal derivations of PDE 53ff
 - heavy traffic model 241
 - Markov chain approximation
 - 152
 - optimal stopping 143, 300
 - reflecting boundary 147
 - singular control 244
 - target set 145
- Bellman equation, Markov chains
 - discounted cost 44, 48
 - finite time 51, 336
 - target set 50
 - unique solutions 154
- Calculus of variations 346
- Completely- \mathcal{S} condition 136
- Continuous time interpolation 72,
 - 74
 - convergence 292, 310
 - local properties 71, 76
 - reflecting boundary 306, 316
- representation 77, 207, 306,
 - 316
- singular control 316
- Contraction 46, 50
 - absence of contraction 367
- Control, pure Markov 45, 48
 - admissible 47, 71, 76
 - randomized 45
- Controlled jump diffusions 32, 271
 - convergence of costs 283
 - discontinuous dynamics 277
 - limits of 273
- Controlled transition probability
 - 47, 70
- Convergence in distribution 248
- Convergence of costs
 - deterministic problems 83, 346ff
 - lower bound 294, 312
 - upper bounds 295, 314
 - without local consistency 297
- Convergence
 - interpolated Markov chains 292
 - sequence of jump diffusions
 - 273
- Covariance matrix, skew 109ff
- Density of jump diffusion, approximation 340
- Deterministic problems
 - approximating cost functions
 - 82
 - discontinuous running costs
 - 395
- Discount factor
 - effective averaged 77
 - state dependent 38
- Domain decomposition 169, 337
- Dynamic programming equation
 - (see Bellman equation)
- Ergodic control problem

- approximation to 321
- Bellman equation 195, 206
- boundary cost and control 214
- centering 197, 200
- constraints 323
- contractions and 199
- convergence of costs 321
- cost function 194, 204, 205, 301
- Jacobi iteration 197
- numerical methods 201–203, 209–214
- sense of uniqueness 195
- singular control 236, 323
- stationarity of limit 321
- weak convergence 321
- Error bounds, discounted cost
 - Gauss-Seidel method 163
 - Jacobi method 162
- Exit time
 - continuity of 261, 279
 - convergence 74, 261, 280
 - randomized stopping 282
 - uniform integrability 262, 297
 - via Girsanov transformation 282
- Explicit Markov chain approximation
 - Bellman equation 336
 - example 326
 - general method 329
- Exterior cone condition 354
- Filtration 8
- Finite difference approximation
 - central difference (symmetric) 96
 - general method 106ff
 - Markov chain interpretation 92
 - one sided 93
 - uncontrolled Wiener process 91
- Finite time problem
 - Bellman equation 335
- decomposition 337
- Markov chain 41
- solution methods 337
- Fokker-Planck equation 339
- Gauss-Seidel method 157
 - ordering of states 159
 - red-black 170
 - smoothing properties of 179
- Girsanov transformation 17, 27, 282
- Grid refinement 177
- Heavy traffic 217
 - assumptions 226
 - boundary approximations 238
 - formal dynamic programming equation 234
 - limits 221, 222, 228, 233
 - Markov chain approximation 236ff
 - production system, impulse control 230
 - routing control 231
- Height function 379
- Implicit Markov chain approximation
 - Bellman equation 336
 - example 330
 - general method 332
 - representation of solution 334
- Impulsive control, heavy traffic limit 230
- Intensity function 379
- Interior cone condition 354
- Interpolated process 72, 75, 130, 138, 240
 - representation 132, 139, 289
- Interpolated time 72
- Interpolation interval 70
- Invariant measure, approximation of 206
- Itô's formula 12
 - diffusion 15

- jump diffusion 31
- vector case 13
- Jacobi method 156
 - ergodic control 197
- Jump term
 - approximating chain 128, 291
 - finite difference method 132
 - local consistency 129
 - properties 27, 127
- Jump term, approximation to 291
- Jump times, Markov chain interpolation 75
- Linear programming 186
 - basic feasible solution 187
 - basic variable 187
 - dual program 187, 191
 - Markov chain optimization 189
 - simplex procedure 187
- Local consistency 71, 128, 137, 238, 351
 - convergence without 297
- Markov chain approximations
 - sticky boundaries 147
 - tightness 292
 - weak convergence 292, 307, 359
- Markov chain approximation method 67ff
- Markov chains
 - controlled 47
 - discounted cost 37
 - ergodic cost problem 40
 - optimal stopping 42
 - reflecting boundary 39
 - stopping time 42
- Markov property 36, 48, 71
- Martingale 8
 - local 9
 - vector valued 8
- Martingale problem 253
- Modified approximation in policy space 159
- Multigrid method 179ff
 - interpolation operator 184
 - k -level procedure 184
 - projection operator 183
 - smoothing properties of Gauss-Seidel 179–182
- Nonanticipative 11
 - control 19, 32
- Nonlinear filtering 339ff
 - approximations to 344
 - convergence of approximations 344
 - Markov chain signal 342–344
 - representation of optimal filter 341
- Normalization of transition probabilities 100
- Optimal stopping
 - convergence of costs 300
 - existence of optimal 299
 - Markov chain 42, 300
- Optimal control
 - approximation of 284
 - existence of 278
 - nonexistence of 86
- Pathwise uniqueness 15
- Picard iteration 16
- Poisson random measure 29
 - characterization 254
- Principle of optimality 44, 48, 376
- Prohorov metric 249
- Queueing system
 - admission control 223
 - controlled heavy traffic limit 222, 228
 - controlled 222
 - heavy traffic approximation 219ff
 - heuristic approximation 219
- Randomized stopping 282

- Reflecting boundaries 21, 133, 303, 412
 approximations 136, 238, 306
 convergence of approximations 310
 convergence of costs 312ff
- Reflecting jump diffusions 32
- Reflecting states, eliminating 148
- Reflection directions 23, 134, 236
- Relaxed control, deterministic 85ff, 265, 355
 approximation of 355
 chattering theorem 88
 equivalence of values 86
 relation to ordinary controls 266
 representation as a measure 86
 topology 265
- Relaxed control, stochastic 267
 approximation of 277
 optimal 276
- Semiaccelerated Gauss-Seidel 164
- Shape-from-shading 378
- Simple function 11
- Singular control 223, 315
 admission control 223
 convergence of costs 318
 dynamic programming equation 244
 Markov chain approximation 242
 portfolio selection problem 235
 rescaled time 317
 routing control 231
 uniform integrability of reflection terms 318
- Singular points 379
- Skorokhod Problem 23, 304
 approximations 139ff
 bounds on reflection term 304
 convention in jump case 32
 local consistency 136
- tightness of approximations 310
 uniform integrability of reflection terms 311, 318
- Skorokhod representation 252
- Splitting the operator 103, 111
- Stationary measures, approximation of 323
- Stochastic differential equation (SDE) 9
 with reflection (SDER) 24
- Stochastic integral 11ff
- Stopping time 9
- Strong existence 14
 controlled diffusions 19
- Strong uniqueness 15
 controlled diffusions 20
- Submartingale problem 23
- Superlinear growth condition 348
- Target sets, isolated point example 385
- Tightness 250
 of Markov chain approximations 292, 318
- Time rescaling 308, 317
 inverse 312
- Topology for the set $[0, \infty]$ 260
- Uniqueness in the sense of probability law 15
- Variance control 100
- Variable grids 122ff
 lack of local consistency 123
- Viscosity solution 411
 subsolution 414
 supersolution 414
- Weak convergence 248ff
- Weak existence 14
 controlled diffusions 20
 controlled jump diffusions 32
 jump diffusions 30

Weak uniqueness 15
controlled diffusions 20
controlled jump diffusions 33
jump diffusions 30

Wiener process 10
approximation to 290
characterization 253, 276, 312
random walk approximation
78