Differential games in economics and management science

Engelbert Dockner · Steffen Jørgensen
Ngo Van Long · Gerhard Sorger

Differential Games in Economics and Management Science

Differential Games in Economics and Management Science is a comprehensive, self-contained survey of the theory and applications of differential games, one of the most commonly used tools for modelling and analysing economics and management problems which are characterized by both multiperiod and strategic decision making. Although no prior knowledge of game theory is required, a basic understanding of linear algebra, ordinary differential equations, mathematical programming, and probability theory is necessary to appreciate fully the authors' arguments.

Part I presents the theory of differential games, starting with the basic concepts of game theory and going on to cover control theoretic models, Markovian equilibria with simultaneous play, differential games with hierarchical play, trigger strategy equilibria, differential games with special structures, and stochastic differential games. Part II offers applications to capital accumulation games, industrial organization and oligopoly games, marketing, resources, and environmental economics.

ENGELBERT J. DOCKNER is based at the University of Vienna.
STEFFEN JØRGENSEN is based at University of Southern Denmark, Odense University.
NGO VAN LONG is based at McGill University and CIRANO, Montreal.
GERHARD SORGER is based at Queen Mary and Westfield College, London.

Differential Games in Economics and Management Science

Engelbert J. Dockner
University of Vienna

Steffen Jørgensen
University of Southern Denmark, Odense University

Ngo Van Long
McGill University and CIRANO, Montreal

Gerhard Sorger
Queen Mary and Westfield College, London

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521631259

First published 2000

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Differential games in economics and management science / Engelbert Dockner ... [et al.].
p. cm.
Includes bibliographical references and index.
ISBN 0 521 63125 4 – ISBN 0 521 63732 5 (pbk.)
1. Game theory. 2. Differential games. 3. Economic models. 4. Management. I. Dockner, Engelbert.

HB144.D53 2000 330'.01'5193–dc21 99-059881

ISBN-13 978-0-521-63732-9 paperback
ISBN-10 0-521-63732-5 paperback

Transferred to digital printing 2006

Contents

Figures

Tables

Preface

Game theory is a theory of conflict and cooperation between rational decision makers. It relies heavily on mathematical models and has proven useful to address problems of conflict and cooperation in economics and management science, as well as other areas of social science. Differential games are dynamic game models used to study systems that evolve in continuous time and where the system dynamics can be described by differential equations.

This book is a self-contained introduction to differential game theory and its applications in economics and management science. The book is introductory, but not elementary. It requires no prior knowledge of game theory, although to benefit fully from the theoretical developments in chapters 3 to 8, and to understand the various steps of the analysis of the applications in chapters 9 to 12, a basic knowledge of mathematical analysis, linear algebra, ordinary differential equations, mathematical programming, and probability theory is necessary. Whenever appropriate, we provide supporting references in these areas. The applications treated in chapters 9 to 12 presuppose that the reader is familiar with basic concepts in economics and management science. The selection of the topics dealt with in the applications chapters has to some extent been guided by the personal research interests of the authors. Nevertheless, the areas covered in these chapters comprise a fairly large subset of the interesting applications of differential games in economics and management science.

Our intended readership includes advanced undergraduate and first-year graduate students of economics, management science, operations research, and quantitative business administration who are specializing in strategic decision making, differential game theory, and applied differential games. Students in game theory in general, or in dynamic game theory and its applications, can also benefit from the book. The material herein may be used for one- or two-semester courses. Depending on the area of specialization, the mathematical background of the students, and the available time, different weights can be allocated to the book's two parts.

This volume should also prove valuable to economists, management scientists, operations researchers, and academics in other fields of the social sciences who work, or intend to work, in areas such as economic dynamics, differential games, industrial organization, or oligopoly theory. Researchers wishing to specialize in, for example, investment and capital accumulation, pricing and advertising competition, or the economics of natural resources should also find the book useful.

Differential Games in Economics and Management Science is a product of the joint efforts of four authors, located in Austria, Canada, Denmark, and the UK. Overcoming the geographic distances was not a major obstacle, thanks to modern information technology; but overcoming different conceptions concerning content and style of presentation proved to be more of a problem. In order to solve it, each chapter was written by one of the four authors and afterwards 'refereed' by the others. We hope that the outcome is acceptable to our readers.

We wish to thank Cambridge University Press, and Ashwin Rattan in particular, for their patience and support and for putting the pieces of our manuscript together to such good effect. The authors would also like to thank the Department of Business Administration and the Department of Economics at the University of Vienna, the Department of Management at Odense University, the Department of Economics at McGill University, and the Center for Interuniversity Research and Analysis on Organizations (CIRANO) in Montreal for providing the necessary research environments, and M. I. Kamien for useful comments. Engelbert J. Dockner acknowledges research support from the Austrian Science Foundation under contract P10850-OEK and wishes to thank M. Schuster for technical assistance. Steffen Jørgensen acknowledges support from SNF/SSF Denmark, and wishes to thank Kristian R. Miltersen and Bjarne G. Sørensen for technical assistance. Ngo Van Long wishes to acknowledge useful comments offered by Bach Long and Huilan Tian. He also thanks Gérard Gaudet, Koji Shimomura, and Hassan Benchekroun for useful discussions. Gerhard Sorger acknowledges support from the Austrian Science Foundation under grants J1179-SOZ and P10850-OEK.

1 Introduction

This book deals with the theory and applications of noncooperative differential games. A noncooperative game is a strategic situation in which decision makers (from now on: players) cannot make binding agreements to cooperate. In a noncooperative game, the players act independently in the pursuit of their own best interests. Confining our interest to noncooperative games should not be seen as an indication that cooperative games are less interesting. The reason simply is that, in the area of differential games, cooperative theory is far less developed than noncooperative theory and almost all applications in economics and management science are in the noncooperative setup. We do not deal with zero-sum games (these are games in which the players have completely opposite interests, that is, the gain for one player equals the loss for another player), because the zero-sum assumption is only plausible in rather special situations in economics and management science.

Differential games belong to a subclass of dynamic games called state space games. In a state space game, the modeller introduces a set of (state) variables to describe the state of a dynamic system at any particular instant during play. The hypothesis is that the payoff-relevant influence of past events is adequately summarized in the state variables. To illustrate, the state vector may consist of the current capital stocks of N oligopolistic firms and these stocks can be influenced by the firms through the choices of their individual investment rates. If a state space game is cast as a differential game, the assumption is that time evolves continuously and that the evolution over time of the state variables can be modelled by a set of differential equations. If the modeller prefers to use discrete time, a system of difference equations replaces the differential equations and the game is then a difference game. The choice between discrete time and continuous time often seems quite arbitrary. Discrete-time models involve the assumption that no decisions are made between the time instants that define the periods. This raises a problem of what should be the length of a period. The lengths of reaction lags or information lags often play a role in dynamic economic models and the case of negligible lags can serve as a benchmark. This can be accommodated in continuous-time models, whereas in specific discrete-time models the difference between one-period lags and no lags at all can be rather dramatic. Continuous-time models can give problems, too. An important one is the lack of a natural notion of a 'last time' before time instant t.

In most of the differential games considered in this book, both in theory (part I) and in applications (part II), ordinary differential equations are employed to describe the evolution over time of the state of a dynamical, many-player system. The theory of differential games with partial differential equations is technically involved and applications are few. In most areas of this book we employ deterministic dynamics, but chapter 8 is devoted to stochastic differential games.

Part II of the book, on applications, includes a number of studies in the following areas of economics and management science:

- capital accumulation and investments,
- R&D and technological innovations,
- Cournot oligopoly,
- pricing and advertising decisions in marketing,
- natural resource extraction,
- pollution control.

We have chosen these areas because they contain the major part of differential games applications. We do not pretend that other fields are less important. The economic applications treated in the book are mostly drawn from microeconomics and industrial organization. Macroeconomic models are given less emphasis, although examples can be found in chapters 7, 10, and 12. The reader who is particularly interested in differential games in macroeconomics is referred to Petit [190], in which a series of issues in macroeconomic policy design is addressed, using optimal control theory and difference games, as well as differential games.

The rapidly increasing use of game theory in the social sciences has produced a number of textbooks on game theory with economic applications, including Rasmusen [196], Kreps [154], Dixit and Nalebuff [41], Fudenberg and Tirole [104, 105], Friedman [100], McMillan [172, 173], Binmore [8], and Gardner [108]. The main part of this literature gives little, if any, attention to differential games. It may seem that differential games are viewed as a somewhat esoteric branch of game theory, being practised by control engineers and researchers of similar inclination. One aim of writing this book has been to correct this view and demonstrate that differential games are founded on the same notions and use the same concepts as 'mainstream' dynamic game theory. From their origins in optimal control theory and dynamic programming, differential games have enriched dynamic game theory with a number of useful concepts, and we maintain that differential games are a natural and interesting subclass of dynamic games. Its emphasis on differential games makes this book different from most of the textbook literature on game theory.

The historical development of the theory of differential games is documented by the works of, for example, Isaacs [130], Friedman [98], Leitmann [157], Krasovskii and Subbotin [153], and Başar and Olsder [4]. This literature shows that differential games originated as an extension of optimal control theory (which is concerned with one-person dynamic optimization problems in continuous time). Not surprisingly, the analytical tools of differential games have been influenced by those of optimal control (e.g., maximum principles, Hamilton–Jacobi–Bellman equations of dynamic programming, state space analysis). However, differential game theory has long since transcended its origins in one-person dynamic optimization and moved on to become a subclass in its own right of the broader field of dynamic game theory.

The theory of dynamic games evolved from static game theory which originated in the 1920s with seminal works of John von Neumann and made great progress in the 1940s and early 1950s with *The Theory of Games and Economic Behavior* by John von Neumann and Oskar Morgenstern [239] and the papers by John Nash (e.g., Nash [185]). Central concepts in game theory have been influenced by one-person decision theory (utility theory), the groundwork of which also was laid by von Neumann and Morgenstern. A considerable part of game theory's inventory of concepts and methods originates from the study of economic problems.[1] During the course of time, game theory has received important inputs from, and has been successfully applied in, other areas of social science (e.g., political science) as well as other sciences (mathematics, biology).

The textbook literature dealing both with differential games and with their economic and managerial applications is sparse and we believe that the present book fills a gap. No other existing textbook provides a rigorous treatment of the theory of differential games, having at the same time an extensive coverage of economic and management science applications. Case [16] is an early attempt to give an integrated treatment of differential game theory and economic applications. Başar and Olsder [4]

[1]Recently, John Nash, John Harsanyi, and Reinhard Selten were jointly awarded the Nobel Prize in economics for pathbreaking work in game theory.

give a thorough account of static and dynamic noncooperative game theory with some emphasis on differential games, but their examples are not drawn from economics or management science. Mehlmann [176] is a short treatise on differential games which includes a few economic examples. Petit [190] covers both optimal control theory and differential games with a focus on macroeconomic policy design. Clemhout and Wan [32] provide a brief survey of differential games in economics. Sethi and Thompson [216], Feichtinger and Hartl [85], and Kamien and Schwartz [148] are textbooks on optimal control theory that include a brief exposition of differential games and their applications.

We have divided the book into two parts. The first part (chapters 2 to 8) has the objective of giving a rigorous introduction to differential game theory. The second part (chapters 9 through to 12) presents various applications, mainly drawn from the journals literature. In the applications chapters we present the models and the analysis in a detail that should be sufficient to enable the reader to see the connections with the theoretical part of the book and to become familiar with important areas of applied differential games.

Part I begins with a brief introduction to static and dynamic game theory (chapter 2). We emphasize only game theoretic concepts that are relevant to the study of differential games. Readers who are familiar with game theory can skip this chapter. Chapter 3 gives an account of elements of optimal control theory that are useful in the study of differential games. Readers who are familiar with optimal control theory can skip chapter 3, although they are recommended to read at least section 3.5 which deals with the important concept of a strategy. Chapter 4 gives a precise definition of what a differential game is and states equilibrium conditions for Nash equilibria under simultaneous play. Important tools here are the maximum principle of optimal control theory and the Hamilton–Jacobi–Bellman equations of dynamic programming. Chapter 5 relaxes the assumption of simultaneity of decisions and deals with 'leader–follower' differential games. In such games one player announces his decision before the other player(s). Chapter 6 deals with the use of trigger strategies to sustain tacit collusion (an unwritten, cooperative agreement to provide monopoly profits). The idea of trigger strategies is to threaten by a harsh punishment any player who might be tempted to deviate from the tacit agreement. The problem of sustaining a collusive agreement has been extensively studied in oligopoly theory, using a special type of dynamic games known as repeated games. In differential games, the literature on collusion is somewhat scattered and is mainly applications oriented. Chapter 6 collects the most useful results of this literature. Chapter 7 provides a number of specific quantitative and qualitative techniques by which differential games with particular structures can be analysed. Much of the progress in analytical solvability of differential games has come on a 'type-by-type' basis, each time adding a new type of game to the inventory of tractable games. The reader should note that we do not deal with numerical methods for solving differential games. Algorithms and computer software are, however, available for such purposes. Chapters 4 to 7 all deal with deterministic differential games. The assumption of certainty is modified in chapter 8, which considers two important examples of stochastic modelling. The first one is a piecewise deterministic game. In such a game, the evolution of the state variables proceeds deterministically between discrete instants of time but, at unknown instants of time, random changes in the dynamic system occur. In the second case, uncertainty influences the dynamic system continuously and the random disturbances are modelled by a Wiener stochastic process.

Part II, chapters 9 to 12, is devoted to the application areas that we mentioned above. Chapter 9 deals with a well-known theme of economic theory, namely investment and the accumulation of capital stocks. Capital stocks can consist of physical capital (productive capacity), but stocks of human capital and consumer goodwill are also included in the general framework. Chapter 10 considers two selected issues coming from the broad fields of industrial organization and oligopoly theory. The first one deals with duopolistic output competition under sluggish market price adjustment, the second one with rivalry in research and development (R&D) when the dates of the successful completion of the firms' R&D projects are uncertain. Chapter 11 is devoted to marketing competition where we focus on two basic topics: pricing and advertising. Finally, chapter 12 presents a number of differential game models of natural resource extraction, both in exhaustible resources (e.g., oil fields) and renewable resources (e.g., fish stocks or forests).

Each chapter (except this one) includes a 'Further reading' section that provides some notes on the material covered and references to material for further studies, together with references to supplementary texts (for example, in the theory of differential equations, mathematical programming, optimal control theory, probability, and stochastic processes). Exercises are provided at the end of each chapter, and the reader is encouraged to work through this material. (Brief answers and hints to the exercises can be found at the end of the book.)

Part I Theory

2 Basic concepts of game theory

This chapter introduces those concepts in static and dynamic game theory that are particularly relevant for the study of differential games, the theory of which will be presented in chapters 4 to 8. In this chapter we proceed in a somewhat informal way and do not attempt to render a precise mathematical representation of each and every concept. The main idea is to provide an understanding of what game theory is about, and we have chosen not to complicate matters by insisting on mathematical rigour. Those wishing to study more precise accounts of game theory should consult the references mentioned in section 2.4.

We start by discussing the distinction between noncooperative and cooperative games and offer some comments on game theoretic modelling. The chapter proceeds by presenting the two types of game theoretic models: the strategic form (or normal form) and the extensive form. We introduce fundamental concepts such as a player's strategy, the Nash equilibrium, the role of the information available to the players, and the concept of subgame perfectness. Finally, a brief presentation of a standard differential game model is given, postponing the detailed description to chapter 4.

2.1 Axioms of game theory

Game theory is concerned with the study of situations involving two or more decision makers (individuals, organizations, or governments). Decision makers are designated as players. The players often have partly conflicting interests and make individual or collective decisions. In a game, the fortunes of the players are interdependent: the actions taken by one particular player influence not only his own fortune but also the fortunes of the other players. Such interdependence is well known from many areas of economics and management science: think about the interdependence between firms that compete in an oligopolistic market, between a firm and the consumers who are in the market for the firm's product, between a firm and its suppliers of inputs, between the management of a firm and the firm's employees, or between a firm and its lenders and potential investors. In contrast to one-person optimization (e.g., mathematical programming, optimal control theory, or decision theory), a key feature of game theory is the presence of two or more players in a game situation where each player takes into account the decisions of the other players when choosing his or her own course of action.[1]

Game theorists make a distinction between two kinds of games: cooperative and noncooperative. Supposing that a game is played in a noncooperative way means that the institutional environment is one in which the players cannot or will not make binding agreements to follow some joint course of action. Players are rivals and all players act in their own best interest, paying no attention whatsoever to the fortunes of the other players. A fundamental problem for any player in a noncooperative game is that of strategic uncertainty: when a player must act, he does not know for sure how the other players will act. (Otherwise, the decision problem of a player would be obvious.) Noncooperative game theory offers a formal methodology to try to resolve the strategic uncertainty and predict what could be the outcome when rational players have acted in accordance with their plans.

In a cooperative game the assumption is that the players realize that there are gains to be obtained if they can agree to act as a group and coordinate their actions for their mutual benefit. An important element in a cooperative game is the strategies available to groups of players, in contrast to a noncooperative game in which what matters are the strategies available to each individual player. Questions to be answered by the analysis of a cooperative game are, for instance: Under which circumstances can an agreement be established? What would be the terms of an agreement? To answer such questions, a starting point is to determine what the players could achieve by acting as a group, and then try to identify the different ways in which the gains of cooperation could be divided among the individual players. Then one can try to predict which particular agreement would be implemented.

The categorization of games as cooperative and noncooperative should be seen as a recognition of the fact that often there is more than just one way in which a particular game can be played. It is less advisable to think

[1]In the sequel, a single decision maker is often referred to as 'he'. In two-person games, the decision makers conveniently could be represented by a 'he' and a 'she'. In games with three or more players, the reader is free to develop his or her own system of indicating the gender of the players.

that the set of all games in some exogenous way has been divided into cooperative and noncooperative ones. Depending on the underlying institutional characteristics, a game may be analysed as cooperative or noncooperative. In some situations we are also interested in mixed cases; for instance, could cooperative behaviour emerge without any binding agreements in a game which is essentially noncooperative? Under which circumstances can cooperative behaviour be sustained? Problems of tacit collusion are addressed in chapters 6 and 12.

According to the aim of a game theoretic study, two different approaches are discerned. One branch of game theory addresses the question of how players actually act. The problem of how players should act to achieve their objectives is not an issue. The main part of game theory, however, takes its starting point in a set of hypotheses concerning the kind of behaviour that players are assumed to have. Fundamental axioms are that players are rational and think strategically. In this book we are only concerned with axiomatic game theory.

Being rational means that a player has clear preferences, represented by a payoff function. Payoff can be expressed in terms of utility, profit, sales revenue, negative cost, or any other such quantitative measure. In this book we confine our interest to quantitative payoffs. There is a theory of games in which payoffs are represented in a qualitative way: one player wins, the other player loses, or one player pursues another player who tries to evade being captured. The first case is well known from parlour games (e.g., chess, card games). The second is important in, for example, military applications (a missile pursuing an aircraft). Being rational also means that the player makes decisions in a way which is consistent with his objective, namely, to maximize his payoff in the game. Rationality includes that a player knows the number of opponents and the set of all possible strategies that are available to them, and that he can form probabilistic beliefs (expectations) about any uncertainty that may influence the play of the game.

The number of players, the sets of strategies available to them, and the payoffs are essential elements of what game theorists call the rules of the game. The rules are the theorist's formal description of a game and they should be derived from the institutional environment in which the game is supposed to be played, rather than being chosen on an ad hoc basis. The theory includes the assumption of common knowledge, which means that all players know the rules of the game and each player knows that his opponents know the rules, and that the opponents know that he knows the rules, and so forth, ad infinitum. All players are aware that they face rational opponents and all players think strategically. The latter means that, when designing his strategy for playing the game, a player takes into account any knowledge or expectation he may have regarding his opponents' behaviour.

The reader will have noticed that the behavioural assumptions of prescriptive game theory are quite strong. Various points of criticism have been raised against the assumptions, and attempts have been made to modify them.

First, in regard to the assumption of payoff-maximizing behaviour, some researchers have suggested that players are only boundedly rational: they satisfice rather than maximize. Satisficing behaviour means that a player is content with obtaining a certain level of payoff, not necessarily a maximal one. Bounded rationality also includes that, when uncertainties are involved, players are not able to take the full probabilistic view of the game. Rather, they resort to less sophisticated methods to deal with uncertainties.

Second, unlike multi-criteria decision problems, in game theory each player most often has a single-valued payoff function. In problems of economics and management science, the assumption of payoff maximization often translates into the classical objective of profit or, in general, utility maximization. Although this may be an adequate objective in tactical and operational problems, many strategic decisions involve long-term consequences, commitment to the future, and may even affect the viability of the firm. In such situations it is likely that a firm has many other objectives besides profit maximization.

Third, rationality includes the hypothesis that when maximizing his payoff, each player assumes that all other players act as payoff maximizers. The assumption that all players maximize their own payoff differs from the assumption that players expect payoff maximizing behaviour of their rivals. Experiments have produced evidence that apparently rational players are unwilling to rely on the maximizing behaviour of others, in particular when stakes are high.

Fourth, it is an assumption that players have the ability and time to collect and process any amount of information that we want them to, and that they have the skills to perceive all future contingencies.

Game theoretic models are sometimes criticized for employing too many unrealistic behavioural assumptions and for including only a small number of the features of a real-world institutional environment. People who insist on realistic models say that this produces the right solution to the wrong problem. Here we should be aware that models (including game theoretic models) are not supposed to be accurate representations of real-world phenomena, but even very simplified models do not necessarily produce useless predictions. The predictions that result from simple models are correct on their assumptions and one strength of formal modelling lies in the fact that everyone can verify the validity of the conclusions derived from the model. In contrast, a fair part of the strategic recommendations offered by management consultants and other advisors on strategic decision making cannot be verified and any faith placed in such advice is largely a matter of trust.

Working with simple models does not mean that we should not try to validate the model assumptions as far as possible. However, it is well known many of the standard model assumptions in economics and management, even critical ones, are not easily validated or justified in real-world terms. The questions raised in such models should be seen as being hypothetical, as will their answers be. But answering a hypothetical question can also be instructive (which perhaps is the reason why politicians are recommended never to answer hypothetical questions).

On the whole, it seems advisable to regard game theory's assumptions of optimizing and all-knowing players not as a claim that this *is* the way individuals and organizations behave, but rather as a way in which individuals and organizations *could* behave. In this hypothetical framework, game theory can make the strategic structure and opportunities of a particular situation transparent, and game theory has in fact produced many thought-provoking predictions of strategic behaviour. However, in view of the observed limited rationality in human behaviour, we should not try to apply game theory in a mechanistic way as the solution to a real-world decision problem.

Usually there will be three elements of a game theoretic study. First, it is necessary to scrutinize the institutional environment in which the game is supposed to be played in order to obtain a plausible set of rules of the game and to select the relevant variables and their relationships. Next, a mathematical structure must be designed, a game theoretic model that reflects in a simplified way the pertinent aspects of the strategic problem. Third, the interesting properties of the model must be rigorously deduced.

2.2 Game theoretic models

Noncooperative game theory uses two types of models: the strategic form and the extensive form. The strategic form includes the following three elements:

(1) A set of players $\mathbf{N} = \{1, 2, \dots, N\}$.
(2) For each player $i \in \mathbf{N}$ a set of feasible strategies U^i.

(3) For each player $i \in \mathbf{N}$ a real-valued function J^i such that the value $J^i(u^1, u^2, \ldots, u^N)$ represents the payoff of player $i \in \mathbf{N}$ if the N players use the strategies $(u^1, u^2, \ldots, u^N) \in U^1 \times U^2 \times \ldots \times U^N$.

A strategy profile is another name for a set of N feasible strategies. In what follows we always assume that $\mathbf{N}$ is a finite set.

The notion of a player's strategy is fundamental in game theory. We may think of a strategy as a player's contingent plan, to be determined before playing the game. A strategy is a function that tells the player how to select one of his feasible actions whenever he must make a move, for all possible events that may have occurred so far. Let us refer to all such possible events as the history of the game. Thus, a strategy is a mapping from the set of possible histories of the game to the set of feasible actions. It is important to note that a strategy prescribes a player's choice of action for all possible histories of the game, including those histories that will never be observed (because the players would not choose actions that generate such histories).

When the strategic form is used, the game theoretic model includes a list of all possible strategies of all players. Each player is supposed to select before the play of the game one of his feasible strategies. Each player makes this choice independently of any other player and there is no communication or cooperation among the players when they make their strategy choices. No player is informed about the choice of strategy of any other player and this is what causes the problem of strategic uncertainty.

There is no explicit element of time involved in a strategic form game. Nevertheless, a strategic form game can represent – although in a very general sense – a game that is played over several time periods (or stages). In a game played over time, a rational player can determine in advance a complete, contingent plan for all his actions that he must take during the whole game. Such a plan, a strategy, specifies what particular action the player should take in any situation that may possibly occur at any instant of time during the game. The actual play of the game then amounts to the implementation of the players' predetermined strategies.

The extensive form of a game is used for games played over time (or in successive stages) and is represented by a game tree. The extensive form includes a description of the sequence in which players have to take action as well as the instances at which possible chance events will occur during the game. Although these questions of timing are only implicit in the strategic form, the concept of a player's strategy can be seen as an object of both an extensive and a strategic form game. In the extensive form, a player waits to take his action until the game has

Table 2.1. *The strategic form of example 2.1*

	(v_1, v_1)	(v_1, v_2)	(v_2, v_1)	(v_2, v_2)
u_1	(2,2)	(2,2)	(0,0)	(0,0)
u_2	(1,1)	(4,4)	(1,1)	(4,4)

reached a certain instant. The rule by which he chooses his actions depending on the information he has gathered up to that instant is his strategy.

Depending on the specific game model one wishes to analyse, each of the two forms has its advantages and drawbacks. In most games that evolve over time, the extensive form is superior to the strategic form since the extensive form explicitly depicts the order of moves, which information is revealed during the course of the game, and how players take such information into account. On the other hand, in dynamic games of some complexity, the extensive form easily becomes unmanageable.

Example 2.1 The extensive form depicted in figure 2.1 shows a game in which player 1 moves first and has to choose between actions u_1 and u_2. Player 2 moves next and has to choose between v_1 and v_2. The numbered nodes of the tree represent the player who must make a move. The assumption in the figure is that player 2 knows the action taken by player 1 when player 2 has to take his own action. The black nodes are terminal nodes in which payoffs are received. The first number in the pairs associated with the terminal nodes is the payoff of player 1, the second number is the payoff of player 2.

The strategic form of the game is given in table 2.1 and shows that player 2 has a complete, contingent plan for taking an action when his turn comes. For example, the first column shows the payoffs when player 2's strategy is to play v_1 upon both u_1 and u_2. The second column represents the case where player 2 plays v_1 after u_1 and v_2 after u_2. The fourth column represents the case in which player 2 always plays v_2. Recall that the strategy of player 2 describes his plan of action also for histories that would not occur.

Sometimes the feasible strategy sets are finite, as in example 2.1, and we say that the game is finite. The strategic form of a finite, two-player game can be represented by a matrix. Feasible strategy sets could also be

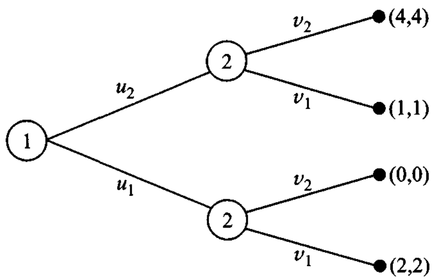

Figure 2.1 The game tree for example 2.1

infinite, for example, $U^i = \{u_i \in I\!R \mid u_i \geq 0\}$. Example 2.5 provides an illustration.

The simplest games are static games (also known as one-stage, one-period, or one-shot games). In such games that are played only once there is no history on which a strategy can be based and a strategy then coincides with the choice of an action, once and for all.

Example 2.2 To illustrate a static game, consider a duopolistic market in which two firms, 1 and 2, must decide simultaneously and independently whether or not to make their respective products compatible (say, with respect to some product dimension). Suppose that both products can only be designed in the two dimensions 'large' and 'small'. Feasible action (strategy) sets are identical, namely $U^1 = U^2 = \{\text{large, small}\}$. These sets as well as the payoffs are depicted in the 2×2 matrix in table 2.2. The first number in a cell is firm 1's payoff , the second number is the payoff of firm 2. The payoffs indicate that both firms gain if their products are compatible and that they stand to gain most if both firms choose 'large'. Each firm wishes to select a strategy (here: a single choice of action) so as to maximize its payoff, knowing that the rival firm will select its strategy so as to maximize its own payoff.

The example clearly illustrates the problem of strategic uncertainty: when a firm must make its choice, it is not informed about the choice of the other firm. But both firms know table 2.2 (the rules of the game), which means that each firm can reason not only about its own choice of action but also about that of its rival. The purpose of a prescriptive game theoretic analysis is to predict which one of the four feasible outcomes (listed in the four cells of the table) will be the one that is actually chosen by the players. One prediction could be that the firms do not choose opposite actions, another one is that both firms would choose 'large',

Table 2.2. *The strategic form of example 2.2*

	large	small
large	(4,4)	(1,1)
small	(1,1)	(2,2)

but as it stands we have no basis for a prediction of what specific outcome will be chosen by the players. In the next example, however, there is a fairly obvious prediction of the outcome of the game.

Example 2.3 The game is static and the players choose an action (a strategy) simultaneously and independently. Player 1 has the feasible strategy set $\{u_1, u_2, u_3\}$ and for player 2 the strategy set is $\{v_1, v_2, v_3\}$. Table 2.3 depicts the strategic form. Both players can reason as follows. No matter which strategy player 1 chooses, v_3 gives player 2 a strictly higher payoff than v_2 (since $2 > 1$, $6 > 4$, and $8 > 6$). Therefore, player 1 expects that player 2 will not play v_2. Observing this we note that the strategy u_1 gives player 1 a higher payoff than both u_2 and u_3, irrespectively of player 2's choice. Since player 2 knows that player 1 knows that player 2 will not play v_2, player 2 expects that player 1 will play u_1. Hence, player 2 should play v_1. Now we have obtained a unique prediction of the outcome of the game, namely, the payoff-pair (4,3).

The process by which we arrived at a prediction of the outcome of the game in example 2.3 is called iterated strict dominance. Since each player has a unique strategy choice we obtain a unique prediction of what would be the outcome of playing the game. This outcome is an equilibrium in the following sense. When player 1 (player 2) expects that player 2 (player 1) will use his strategy v_1 (u_1) (and he has every reason to believe this), then player 1 (player 2) can do no better than to choose the strategy u_1 (v_1).

How can we predict the outcome of a game which does not have an equilibrium as the one in example 2.3? Note that although player i has no knowledge of the rival players' decisions, player i can always determine what would be his best strategy, given any particular set of the rivals' strategies. This is a one-person optimization problem for player i. For notational convenience, define the vector of strategies of player i's rivals by $u_{-i} = (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$. We say that strategy u_i^b is a best reply of player i to the $(N-1)$-tuple u_{-i} if

Table 2.3. *The strategic form of example 2.3*

	v_1	v_2	v_3
u_1	(4,3)	(5,1)	(6,2)
u_2	(2,1)	(8,4)	(3,6)
u_3	(3,0)	(9,6)	(2,8)

$$J^i(u_i^b, u_{-i}) \geq J^i(u_i, u_{-i}) \qquad \text{for all } u_i \in U^i. \tag{2.1}$$

A best reply is a strategy that maximizes player i's payoff, given that his rivals play one particular strategy profile u_{-i}. Needless to say, this profile must be feasible.

A strictly dominant strategy u_i^d is player i's best reply to any feasible profile of the $N-1$ rivals:

$$J^i(u_i^d, u_{-i}) > J^i(u_i, u_{-i}) \qquad \text{for all } u_i \in U^i, u_{-i} \in U^1 \times \dots \times U^{i-1} \times U^{i+1} \times \dots \times U^N. \tag{2.2}$$

In example 2.3 we used the notion of a strictly dominant strategy to iteratively eliminate strictly dominated strategies and were led to a unique strategy choice of each player. In such a game the problem of strategic uncertainty is completely resolved. A notable class of games that admits an equilibrium in strictly dominating strategies are games of the prisoners' dilemma type. If such a game is played only once, the strategic form is as in table 2.4, where the payoff-values need to satisfy the inequalities $c > a > d > b$. Table 2.4 shows that the equilibrium in strictly dominating strategies is the pair (u_2, v_2). The interesting thing is that both players would be better off by playing (u_1, v_1), but in a one-shot game an agreement to do so would not be sustainable.

The notion of strictly dominant strategies is a strong one and we would expect that in most games dominant strategies fail to exist for all players. Being unable to rely on strict dominance, we need a weaker equilibrium concept to predict an outcome of a game. Here the fundamental concept of Nash equilibrium comes in. The Nash equilibrium is a cornerstone of noncooperative game theory and is defined as follows. Strategy profile $(u_1^*, u_2^*, \dots, u_N^*)$ is a Nash equilibrium if for each $i \in \mathrm{N}$ the condition

$$J^i(u_1^*, u_2^*, \dots, u_N^*) \geq J^i(u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_N^*) \quad \text{for all } u_i \in U^i \tag{2.3}$$

holds.

Table 2.4. *The prisoners' dilemma*

	v_1	v_2
u_1	(a,a)	(b,c)
u_2	(c,b)	(d,d)

Comparing (2.3) with (2.1) shows that a Nash equilibrium strategy of player i is his best response to the set of Nash equilibrium strategies of his $N-1$ rivals. Comparing (2.3) with (2.2) shows that a profile of strictly dominant strategies is also a Nash equilibrium (but the opposite is not true). Hence, Nash equilibrium is a weaker equilibrium concept than equilibrium in strictly dominant strategies.

In the case of Nash equilibrium, the use of the term 'equilibrium' is justified in the sense that an equilibrium outcome is a consistent prediction. When player i rationally expects that all his rivals will choose their Nash equilibrium strategies, player i can do no better than choose his own Nash equilibrium strategy. All players being rational, and all players knowing that all the rivals are rational, no player would wish to do anything else but choose his Nash equilibrium strategy. Thus, if all players predict that a particular Nash equilibrium outcome will occur, no player can do better than choose his own Nash equilibrium strategy associated with that outcome. The equilibrium in strictly dominating strategies is also a consistent prediction. Note that, although Nash and strictly dominant strategy equilibria are consistent predictions, this does not necessarily make them good predictions of the outcome of actual play. It may be that the outcome of actual play depends on more than what we decided to model in the strategic form!

An equilibrium in strictly dominant strategies is unique, but uniqueness is not guaranteed for Nash equilibrium. To get a prediction in the case of multiple Nash equilibria we need to assume that the players – in some way or another – can agree to focus on one specific Nash equilibrium outcome. The game in table 2.2 is an example where we have two Nash equilibria.

Example 2.4 This example shows a static and finite game which has a unique Nash equilibrium. First note that u_1, u_2, and u_3 are player 1's best replies if player 2 chooses v_1, v_2, and v_3, respectively. Next note that v_3, v_2, and v_1 are player 2's best replies to u_1, u_2, and u_3, respectively. By (2.3), the profile (u_2, v_2) is the unique Nash equilibrium and the corresponding payoffs are $(1, 1)$.

Table 2.5. *A game with a unique Nash equilibrium*

	v_1	v_2	v_3
u_1	(4,0)	(0,2)	(0,4)
u_2	(2,0)	(1,1)	(2,0)
u_3	(0,4)	(0,2)	(4,0)

The games that we have considered so far are finite games. To illustrate the use of Nash equilibrium in a static game with infinite strategy sets, the next example considers the classic Cournot duopoly game.

Example 2.5 In a market there are two firms, 1 and 2, and each firm produces the same homogeneous product. The outputs are sold at the market-clearing price $P(Q) = 1 - Q$ where $Q = Q_1 + Q_2$ is the total output supplied to the market and Q_i is the output of firm $i \in \{1, 2\}$. For simplicity, suppose that both firms' production costs are zero. Strategies are the output decisions of a single period, and feasibility requires that both firms' outputs are chosen from the set $[0, \infty)$. Thus, the assumption is that outputs are perfectly divisible. Strategies are chosen simultaneously. The profit of firm i is $J^i(Q_1, Q_2) = (1 - Q)Q_i$ and maximizing this profit with respect to Q_i yields firm i's best reply (also known as the firm's reaction function):

$$R_i(Q_j) = (1 - Q_j)/2, \qquad i, j = 1, 2;\ i \neq j. \tag{2.4}$$

This best reply represents firm i's optimal output choice, given any feasible Q_j. The unique Nash equilibrium outputs are $Q_1^* = Q_2^* = 1/3$, since in Nash equilibrium it must be true that $R_i(Q_j^*) = Q_i^*$. The outcome of the game (in terms of optimal profits) is $J^1 = J^2 = 1/9$.

Although in each of the two preceding examples we had one and only one Nash equilibrium, existence and uniqueness of Nash equilibrium is not guaranteed. However, in many games a Nash equilibrium does exist. There are theorems which deal with the existence of Nash equilibrium in various types of games (see section 2.4 for a remark on existence). It happens (in fact, more often than we would like) that we come up with more than just one Nash equilibrium and the question arises: How can the players conform to one particular equilibrium? In the case of non-uniqueness, the Nash equilibrium concept in itself is insufficient to provide a single prediction of what strategies will be used and the problem of strategic uncertainty remains unresolved. To claim that one particular

Nash equilibrium will be singled out as the one that is actually played we must invoke some mechanism which ensures that all players consistently predict that particular equilibrium. One such mechanism, which has an intuitive appeal, is the focal point effect. The idea of a focal point equilibrium is that one particular equilibrium has a property which conspicuously singles that equilibrium out from all the other equilibria. Then we would expect to see this equilibrium as the one which is actually played. There are other routes to uniqueness; for instance, strengthening the concept of Nash equilibrium or designing a formal procedure by which a single equilibrium can be selected (see section 2.4 for a brief discussion of nonuniqueness of Nash equilibrium).

2.3 Dynamic games

Examples 2.2 to 2.5 were static games. Obviously, many strategic problems in economics and management are not properly modelled as static games since firms can make decisions at more than just one point of time. A first question would be: How does one distinguish a dynamic game from a static one? There is no general agreement on the use of the terms 'static game' and 'dynamic game'. One might say that a game in which time is not explicitly involved is a static game, supposing that 'dynamic' refers to the fact that variables (e.g., actions) are explicitly dated. Consider, however, a Cournot duopoly game played in the following way. The firms have to choose their respective output levels independently of each other at each of T successive time instants $1, 2, \ldots, T$. After the firms have made their output choices at time instants $s \in \{1, 2, \ldots, t\}$, these choices will be known to both firms when they have to choose their outputs at the subsequent time instant $t+1$. Now suppose that, before the game starts, each firm must make an irrevocable choice of every output quantity that the firm will produce at time instants $1, 2, \ldots, T$. Thus, each firm must commit itself in advance to a fixed sequence of outputs. This game certainly includes time but one could maintain that such a game should not be called a dynamic game. The argument is that during the play of the game the firms get no opportunity to react strategically to the rival's actions, using incoming information on actions taken.

The following definition of a dynamic game takes into account the reasonable requirement that players should be able to select strategies that are based on information being revealed during the play of the game. A game is said to be dynamic if at least one player can use a strategy which conditions his single-period action at any instant of time on the actions taken previously in the game. Previous actions are those of the rivals but also a player's own actions.

To analyse a dynamic game, we need to start by describing in which order the players take their actions and what information is available to a player when he takes action. In what follows we confine our interest to dynamic games in which all players' actions are observable by all players. The game is said to be one of perfect information. Hence, any player, when taking an action at time t, has perfect knowledge of all previous actions. These are his own past actions and those of his rivals, but can also include acts of nature (chance events) if there are exogenous uncertainties in the game. In such a game we say that players move simultaneously at time t if no player – when taking his action at time t – knows about the actions that the other players take at time t. Notice that this terminology is not meant to exclude games in which no two players make decisions at the same time (typically, the players alternate in making moves), since one can include 'do nothing' as a feasible action at a particular instant.

Since all past actions till time $t-1$ are common knowledge among the players at decision-time t, it makes sense to speak of the history of actions by time t. The history of actions by time t, denoted by h_t, is a sequence of action-profiles $u_1, u_2, \dots, u_{t-1}$, where any such profile is a set of N individual actions of the players. The initial history h_0 is the history before the start of the game and is an empty set. The terminal history is the one after which no more actions occur. Payoffs of the players can be defined as functions of the terminal history but could also be taken as (discounted) sums of per-period payoffs.

As already said, in the analysis of dynamic games the strategic form may not be a satisfactory model. The extensive form is particularly designed for the analysis of dynamic games and can be thought of as the many-player extension of the decision tree, known from one-person decision analysis. The following simple example concerns a two-period dynamic game, cast in extensive form.

Example 2.6 Player E (a potential entrant) moves first and must decide whether or not to enter the market of player I (an incumbent firm). Player E selects an action from the set {'enter','stay out'}. Player I observes the choice of E and then it is player I's turn to make a decision. The incumbent chooses an action from the set {'collude with entrant', 'fight the entrant'}. When player I has made his choice, the game terminates. The payoffs depicted in table 2.6 reflect that player E must pay an entry fee of 2 (million dollars) to enter the market. Total monopoly profits are 20 and this amount is equally shared if E enters and I colludes.

Table 2.6. *The strategic form for example 2.6*

	collude	fight
enter	(8,10)	(–2,0)
stay out	(0,20)	(0,20)

If E enters and I fights the entrant, a price war is supposed to leave both firms with no profits at all. The extensive form is shown in figure 2.2. It shows that the players alternate in making decisions and there are four terminal histories, represented by the four terminal nodes. The assumption is that player I knows for sure in which of the two nodes (dots marked by I) he is when he must make his decision. In table 2.6 we show the strategic form of the entry game. The strategic form exhibits two Nash equilibria. One has payoffs (8,10), corresponding to the profile ('enter','collude'). The other equilibrium has payoffs (0,20) and corresponds to the pair ('stay out','fight').

From the incumbent's point of view it is best if E would choose not to enter since this leaves I with the monopoly profit of 20. Now suppose that the incumbent can communicate with the entrant before playing the game. It may be tempting for the incumbent to try to threaten the entrant by the 'fight' strategy, that is, to try to single out the equilibrium ('stay out','fight') as the one to be played. However, this story would only work if the incumbent could credibly commit himself to use the 'fight' strategy. Then the entrant would be discouraged from entering (-2 is worse than 0). Note that the threat to fight would actually not be called upon. But the threat of playing 'fight' is not credible since if E enters, it is not in the incumbent's best interest to fight. If E enters, the incumbent should rather collude because this is his best reply to 'enter'. Seeing through all this, the entrant knows that the incumbent is bluffing and the entrant can safely enter the market.

The reader has noticed that using the strategic form to analyse the entry game yields no indication of how to choose between the two equilibria. The extensive form revealed that one Nash equilibrium was flawed in the sense that it was supported by an incredible threat (an irrational plan). Using the extensive form we can analyse the game by starting from the terminal nodes. Doing so clearly reveals that, given E has entered, 'fight' is not a best reply of the incumbent.

The principle of backward induction can be applied in games with a finite number of time periods $T < \infty$ and finite strategy sets. In the entry

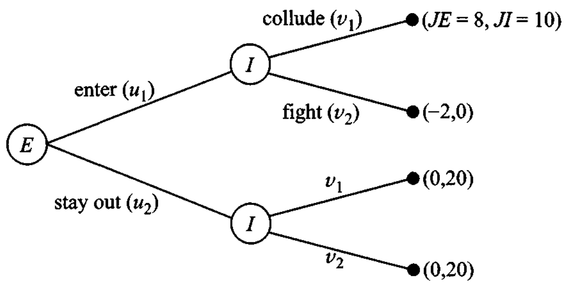

Figure 2.2. The game tree for example 2.6

game, backward induction starts by determining the optimal choice of the incumbent at $t = 2$, given any history leading to this decision point. In the example there are but two histories and given the choice 'enter', the incumbent's optimal action is 'collude'. Given 'stay out', the incumbent is free to choose between 'collude' and 'fight'. (The actual choice between these two is irrelevant: if the entrant stays out, there is no decision for the incumbent to make.) Working backward in time, we then consider the time instant $t = 1$ (at which the induction stops in the example) and determine the optimal choice of the entrant, given that the incumbent will act at $t = 2$ according to his strategy. The entrant's optimal choice is 'enter' (since $8 > 0$).

The idea of backward induction originated in dynamic programming. It has a natural extension in the property of subgame perfectness which is an important concept of dynamic game theory. The effect of invoking subgame perfectness is that incredible threats are ruled out. For an informal discussion of the notion of subgame perfectness, let us agree that in a dynamic game, a subgame is a 'truncated' version of the whole game. A subgame is a game in its own right and a subgame starts out at time instant t, after a particular history of actions h_t. Denote such a subgame by $\Gamma(h_t)$. A Nash equilibrium strategy profile for the whole game, say σ, induces a strategy profile in the subgame $\Gamma(h_t)$. This truncated profile is the restriction of σ to the subgame $\Gamma(h_t)$. A Nash equilibrium strategy profile σ for the whole game is subgame perfect if for any history h_t it holds that the restriction of σ to the subgame $\Gamma(h_t)$ is a Nash equilibrium of $\Gamma(h_t)$. Thus, subgame perfectness requires not only that the profile σ must be a Nash equilibrium in the whole game, but also that the relevant restrictions of the profile σ must be Nash equilibria in every subgame. Recall that one of the axioms of normative game theory is that rationality of players is common knowledge. This means that no player should expect anything else than rational behaviour on the part of his opponents. With a view to irrational threats this leads to the requirement that an equilibrium should at least be subgame perfect.

It is important to note that the definition of subgame perfectness includes the condition 'for any history h_t'. This means that the restriction of the overall equilibrium profile must be an equilibrium also in subgames that will never be played. Thus, subgame perfectness induces Nash equilibrium behaviour in all subgames, not only along the overall equilibrium path but also in all games off the overall equilibrium path.

Subgame perfectness was developed to sharpen the predictions of Nash equilibria. Thus, subgame perfectness can be viewed as a stronger equilibrium concept than Nash equilibrium, a refinement of Nash equilibrium. There exist many refinements other than subgame perfectness but the general idea is that Nash equilibrium is not restrictive enough in the following two senses. First, it does not necessarily produce a unique prediction of the outcome of a game and, second, some Nash equilibria are flawed as rationally acceptable equilibria (e.g., supported by incredible threats). Using a refinement of Nash equilibrium does not guarantee a unique outcome under all circumstances but serves at least to exclude certain less realistic equilibria, by sharpening the rationality requirements.

Let us illustrate the analysis of a dynamic game by another example. The situation depicted in figure 2.2 can be used as an example of a two-player leader–follower game in which the aggressive player (the leader) decides upon his action first. The other player (the follower) observes the action of the leader and has to react to it. Typically, leader–follower games are played over two periods and the setup is traditionally known as a Stackelberg game.[2] Note that a Stackelberg game is a dynamic game since the follower's behaviour cannot be described by a single action but must be represented by a strategy (namely, the follower's rational reaction to any announced action of the leader).

Example 2.7 Consider the Cournot duopoly game of example 2.5 but suppose that firm 1 is now the leader who chooses his output Q_1 first. Firm 2 observes Q_1 before firm 2 makes its own output decision. Feasible strategy sets and payoffs are as in example 2.5. The leader chooses Q_1 as a choice of a single action whereas the follower chooses a strategy, being a map of the leader's set of feasible actions into the follower's set of feasible

[2]The use of the Stackelberg setup in differential games is the topic of chapter 5.

actions. Denote the follower's strategy by S_2. The outcome of the leader–follower game is a pair of outputs $(Q_1, S_2(Q_1))$ and payoffs $J^i(Q_1, S_2(Q_1))$.[3]

We look for Nash equilibria of the leader–follower game. One particular Nash equilibrium is determined by letting the follower maximize his profit for any feasible Q_1. This yields the best reply given by (2.4) in which we put $i = 2$:

$$R_2(Q_1) = (1 - Q_1)/2. \tag{2.5}$$

In Nash equilibrium, the leader maximizes his profit, given the strategy $S_2 = R_2$ of the follower. The solution of this maximization problem is $Q_1 = 1/2$, which by (2.5) implies $Q_2 = 1/4$. The corresponding profits are $J^1 = 1/8$ and $J^2 = 1/16$. Denote these output levels and profits as the Stackelberg outcome. Compared to the simultaneous-move Cournot game, the Stackelberg leader has greater output and profit. The follower's output and profit are lower than those in the Cournot game.

Now suppose that the players use the strategies $Q_1 = 1/3$ and $S_2(\bar{Q}_1) = 1/3$ for any feasible $\bar{Q}_1$. Would this be a Nash equilibrium? Given that the follower's output is equal to 1/3 for any choice of the leader's output, the leader can only maximize his profit by choosing his Cournot output level, which is equal to 1/3. Given that the leader puts his output equal to 1/3, the follower can select any strategy S_2 which satisfies $S_2(1/3) = 1/3$ and maximizes the follower's profit. Among these strategies is the constant strategy $S_2(\bar{Q}_1) = 1/3$. This strategy is not necessarily a best reply to outputs $\bar{Q}_1$ that differ from 1/3 but this does not affect that the pair (1/3, 1/3) is a Nash equilibrium. Thus, the Cournot outcome is another Nash equilibrium in the leader–follower game.

We saw that the follower's constant strategy $S_2(\bar{Q}_1) = 1/3$ may not be a best reply to some feasible output levels that the leader might, but does not choose in Nash equilibrium. However, the Cournot–Nash equilibrium of the leader–follower game has a more serious shortcoming. The follower would prefer the Cournot outcome over the Stackelberg one and might consider threatening the leader by saying that he will use the constant strategy. But this threat is not credible since the leader should not believe that the follower will use his constant strategy irrespective of the leader's output choice. To see this, suppose that the leader announces his Stackelberg output (which is equal to 1/2). Faced with this *fait accompli*, the follower can do no better than to react in accordance with (2.5), that is, to choose his Stackelberg output (equal to 1/4). The reason for the incredibility of the threat is that if the leader selects any

[3]The leader–follower duopoly game is treated in more detail in chapter 5.

output which differs from $Q_1 = 1/3$, it is not in the follower's best interest to stick to the constant output 1/3.

In this informal exposition we do not give a precise description of the extensive form. It is possible to do so but the presentation is long and quite technical. Although the extensive form has a major role in the modelling and analysis of dynamic games in general, it is only implicit in differential games. In these games, the strategic form is the preferred means of modelling a game. What then is a differential game?

A differential game is a dynamic game, played in continuous time. Two distinguishing features of a differential game are:

(i) the modeller introduces a set of variables to characterize the state of the dynamical system at any instant of time during the play of the game, and
(ii) the evolution over time of the state variables is described by a set of differential equations.

Feature (i) makes the dynamic game a state space game and feature (ii) makes the game a differential game (as opposed to, for instance, a difference game).

Postponing a precise description of a differential game till chapter 4, we only state here the basic ingredients of such a game. Denote the time variable by t and suppose that the game is played over the time interval $[0, T]$, which is referred to as the planning period or horizon. The horizon can be finite or infinite and can be either exogenously fixed or determined optimally as a result of playing the game.

To characterize the current state of a dynamical system, observed at any time t, we introduce an n-vector $x(t)$, called the state vector. If a simple characterization of the dynamical system suffices, the state vector is of dimension one or two. If the system is more complex, the state vector has several dimensions and the resulting game is much harder to analyse. Denote by $u_i(t)$ the action taken at time t by player i. The variable $u_i(t)$ is referred to as the control of player i and is in general an m_i-dimensional vector, $m_i \geq 1$. In many differential games one assumes simultaneous actions, i.e., all players make their individual decisions at the same time t. Any action chosen by a player at any instant of time must be selected from the player's set of feasible actions. This set depends in general on the current time t, the current state $x(t)$ and the set of current actions of the player's rivals. In many games, however, this set is constant or depends in a simple way on $x(t)$.

A characteristic feature of a differential game – and what has coined this type of game – is that the evolution over time of the state vector is determined by differential equations.[4] Various choices exist, among them partial differential equations and differential equations with time lags. In this book – with the exception of chapter 8 in which uncertainties are introduced – we consider deterministic, ordinary differential equations of the type

$$\dot{x}(t) = f(t, x(t), u_1(t), u_2(t), \ldots, u_N(t)), \; x(0) = x_0. \tag{2.6}$$

These equations are known as the state equations, the system dynamics, the kinematic equations, or the equations of motion. They show that in general all players have the possibility of influencing the rate of change of the state vector through the choice of their current actions (controls). In addition, the rate of change of the state at any instant of time depends on the current position of the system, as represented by time t and state $x(t)$. Each player i seeks to maximize his total payoff over the planning horizon, discounted at the rate $r_i \geq 0$:

$$J^i = \int_0^T e^{-r_i t} F^i(t, x(t), u_1(t), u_2(t), \ldots, u_N(t))\, dt. \tag{2.7}$$

In (2.7), function F^i represents the instantaneous payoff of player i, for example, the player's profit or utility rate at time t. Equation (2.7) shows that, in general, all players can influence the payoff of player i through the choice of their current actions. The payoff J^i is to be maximized by player i through his choice of the control $u_i(t)$ for $t \in [0, T]$, subject to (2.6) and feasibility of his control vector. A differential game so defined is a strategic form game. To get a first idea of what is a differential game, consider the following simple example of advertising competition in a duopolistic market.

Example 2.8 In a two-firm differential game with one state variable $x(t)$, the state evolves over time according to the differential equation

$$\dot{x}(t) = u(t)[M - x(t)] - v(t)x(t)$$

in which $u(t)$ and $v(t)$ are scalar control variables of firm 1 and 2, respectively. The state variable $x(t)$ represents the number of customers that firm 1 has at time t and $M > 0$ is the constant size of the total market. Hence $M - x(t)$ is the number of customers of firm 2. The control variables $u(t)$ and $v(t)$ are the firms' respective advertising effort rates at time t. The interpretation of the differential equation is that the number of customers of firm 1 tends to increase by the advertising efforts of firm 1

[4]There is a related theory of difference games in which the evolution of the state is described by difference equations.

since these efforts attract customers from firm 2. On the other hand, the advertising efforts of firm 2 tend to draw away customers from firm 1. Payoffs are given by

$$J^1 = \int_0^T e^{-r_1 t}[q_1 x(t) - c_1 u(t)^2]\, dt,$$

$$J^2 = \int_0^T e^{-r_2 t}\{q_2[M - x(t)] - c_2 v(t)^2\}\, dt$$

in which $q_i > 0$ represent firm i's unit revenues. The second term in the integrand of J^i is a convex advertising cost function of firm i. Feasibility requires that $u(t)$ and $v(t)$ are nonnegative. Each firm wishes to choose its advertising strategy over $[0, T]$ so as to maximize its payoff. The payoff is simply the present value of a firm's profit stream over the horizon. Note that in this game, the rival firm's actions do not influence a firm's payoff directly but only indirectly through the state dynamics.

In differential games we need to specify upon which information a player conditions his strategy. This issue is referred to as the choice of a strategy space or an information structure. (The problem arises since we employ the strategic form, and not the extensive form, to model a differential game.) Choosing an appropriate strategy space is neither a question of rationality nor something which should be done by the modeller in a more or less ad hoc fashion. The choice should be guided by the description of the institutional environment in which the game is played. Differences in institutional setups lead to different games that may call for different strategy spaces.

We may assume, as one extreme, that the players use a minimum of information and base their strategies on time alone. As another extreme, the players can base their strategies on the whole history (here: trajectory) of actions. An intermediate case is one in which the players base their strategies on the current value of the state vector. In state space games – of which differential games are a subset – the assumption is that the previous action history is adequately summarized in the current state vector.

Suppose that player i chooses action $u_i(t)$ at time t. As mentioned, this choice can be based upon different sets of information, but we shall always assume that all players know the value of the initial state vector.[5] An open-loop strategy is a strategy which is conditioned on current time

[5]In some books on differential games, the initial state is treated as a separate piece of (state) information. Here we include it in the rules of the game that are common knowledge (along with the number of players, the length of the planning period, the discount rates, and so forth).

only, that is, a minimal amount of information. Like any other type of strategy, an open-loop strategy is fixed at the start of the game. What particular action to take at a specific instant of time then depends only on the instant of time at which the action is to be taken. The open-loop assumption means that the players leave all information except time out of consideration, or that they must choose open-loop strategies since they cannot observe anything other than their own actions and time.

The use of open-loop strategies has been criticized for being static in nature, not allowing for genuine strategic interaction between the players during the play of the game. There are, however, some arguments in favour of open-loop strategies. We have just mentioned an obvious reason, namely that the players lack any information other than their own actions and time. Thus, they may be unable to observe the state vector, let alone the actions of their rivals. In problems in the economics of renewable resources and the environment, the commitment that lies in open-loop strategies can be seen as a reflection of far-sightedness and concern for the conservation of resources and the environment (see also chapter 12). Further, if the planning horizon is short, an open-loop strategy could be employed as a representation of a rigid strategy for short-term operational or tactical planning. Finally, an open-loop equilibrium could be used as an approximation in a game with a large number of small players (a situation resembling perfect competition in microeconomics).

Markovian strategies are decision rules in which the choice of a player's current action is conditioned on current time t and state vector $x(t)$. Thus, one imposes the Markovian assumption that the game history in itself is not relevant for a choice of action at time t, only the consequences of the history are important and they are reflected in the current value of the state vector. Thus, the use of Markovian strategies is a natural choice in the setup of state space games where the history of the game till time t is summarized in the value of the state vector at time t. The choice of Markovian strategies is also motivated by their simplicity: players react only to factors which are payoff-relevant and constitute an intertemporal link in the game (namely, the state variables). The Nash equilibrium in a game played with Markovian strategies is called a Markovian Nash equilibrium. The reader should be aware that a Markovian Nash equilibrium often is referred to as a feedback (or closed-loop) Nash equilibrium, an inheritance from optimal control theory where feedback (closed-loop) controllers play an important role.

Suppose that all players know the history of actions till time t; this means that they can calculate the value of the state vector at time t. But let us suppose that the players wish to condition their strategies on the history of actions itself. In this context, the strategy of a player is a mapping which associates a control u_i with every instant of time t and every history of the game up to time t. History consists of the restriction of control functions $u_j(\cdot)$, $j \in N$, to the interval $[0, t)$. Strategies based on action history are useful in problems in which the players should be given a chance to react if they observe deviations by their rivals from some tacitly agreed course of action. To observe deviations, the players normally need to know the action history.

Strategies based on action history have been extensively used in dynamic games of tacit collusion where the players use threats of punishment as a means to sustain collusive behaviour. Collusion to secure monopolistic profits is in practice almost universally illegal, but history has seen a remarkably high number of such arrangements (concerning prices, product specifications, advertising, output limitations, market shares, and exclusive selling areas). The general structure of tacit collusion problems is as follows. It would be better if all players stick to their collusive strategies than if all players played noncooperatively, but for each player it is better unilaterally to maximize his payoff (i.e., to cheat on the collusive outcome) if all the other players stick to their collusive strategies. Since actions are observable, the players will know whether cheating has taken place.

What is interesting – from a theoretical and a real-world point of view – is that tacit collusion can occur in an essentially noncooperative game environment without any need for contractual, binding agreements. The reason why a collusive outcome occurs is that the outcome emerges as a Nash equilibrium, that is, the players collude because it is in everyone's best interest to do so. This is game theory's explanation of tacit collusion.

The tacit collusion problem has been particularly popular in repeated oligopoly games (discrete-time games in which the same game is played in each period), but also in differential games a theory of threats and punishments has been developed. In chapter 6 we shall see that important instruments to enforce a tacit agreement are trigger strategies. These are strategies based on action history and they embody a threat to punish a defector if he should dare to cheat on the tacitly agreed collusive outcome.

2.4 Further reading

Cooperative behaviour in games

The idea of taking cooperative behaviour as the starting point of a study of strategic interactions has been questioned by some game theorists, arguing that most interactions do not take place in such friendly environments. Basically a game should be viewed as a noncooperative one. If cooperation emerges, it should be as a Nash equilibrium outcome of a noncooperative game. Thus, if collusion occurs it does so, not as a bargained agreement among players, but because collusive behaviour is in the best self-interest of any individual player.

Many studies of tacit collusion are cast in the framework of a repeated game, i.e., a game in which a one-period game (the constituent game) is identically repeated a finite or an infinite number of times. It can be shown that if the constituent game has a unique Nash equilibrium and the number of repetitions is finite, then there is a unique and subgame perfect Nash equilibrium which simply is the identical repetition of the Nash equilibrium of the constituent game. This result is driven by the assumption of a known and finite horizon of the game and in such a game the players cannot cooperate at all. On the other hand, if the number of repetitions is infinite and players have sufficiently small discount rates, there arise a vast number of collusive equilibria. The so-called folk theorems show that any collusive outcome which is feasible and individually rational can be enforced as a noncooperative Nash equilibrium.

In the differential games literature, cooperative outcomes have been analysed in, for instance, Hämäläinen et al. [113], Ehtamo et al. [70], Haurie and Tolwinski [125, 126], Haurie and Pohjola [124], Kaitala and Pohjola [141], Mehlmann [177]. Quite a few of the differential game studies of cooperative outcomes are concerned with bargaining to reach an agreement on the exploitation of natural resources (e.g., fish stocks) or energy supplies. Other applications, e.g., in natural resources and policy coordination, take the noncooperative approach to collusion.

Rationality in games

We already mentioned bounded rationality as one suggestion to relax game theory's strong behavioural assumptions. A more radical proposal comes from Gilboa and Schmeidler [111] who assume that each player considers the behaviour of the other players as a phenomenon of nature, rather than the result of rational strategic decisions. The players ignore complicated strategic considerations and simply update some statistics on the opponents' behaviour. Then they choose a best response to this behaviour. Quite another alternative is to draw upon the theory of evolutionary selection and claim that also in situations of economic competition, decision makers behave in order to increase the probability that they survive in the game of economic selection. Economic selection results in the growth and survival of the most profitable firms and those firms will earn maximal profits – as if they had behaved as payoff maximizers. A reference on evolutionary games in general is Maynard Smith [170].

Nonuniqueness of Nash equilibrium

As we have seen, multiplicity of equilibria occurs in static games and in repeated games. The phenomenon is also well known in so-called games of incomplete information (cf. section 2.4). We discussed one refinement of Nash equilibrium, subgame perfectness, but many other refinements exist. For details see Myerson [184], Van Damme [236], Friedman [101]. Using a refinement raises questions such as (see Myerson [184]): does the refinement satisfy an existence theorem and will the refinement exclude all intuitively unreasonable equilibria? Working in a different direction than the refinements, attempts have been made to develop a procedure (a theory) that selects a unique equilibrium (Harsanyi and Selten [118]). A general task for an equilibrium selection theory is to provide an objective standard that singles out an equilibrium that everyone would expect to be played. In contrast to the refinements, equilibrium selection theory is not very often applied to specific problems in, e.g., economics or industrial organization. Preplay communication has been suggested as a means to determine which of many equilibria actually will be played. Assuming that preplay communication is a possibility, one needs to address questions such as the number of messages that can be sent, their timing, and possible restrictions on the contents of messages. The idea of using a focal point to select an equilibrium goes back to Schelling [211].

Leader–follower games

In most applications of the leader–follower game setup, the specification of roles is exogenously given. Informational or psychological dominance can be invoked to determine the roles. One might also regard the determination of roles as a result of previous plays of the game. It is not necessarily the case, however, that a dominant firm would wish to choose its designated role of leader, nor does a subordinate firm automatically accept the role as follower. Dowrick [66] and Hamilton and Slutsky [114] address the interesting problem of providing an endogenous determination of the roles in leader–follower games.

Incomplete information games

The classification into complete and incomplete information games was introduced by von Neumann and Morgenstern [239]. In a game of complete information, all players know all relevant information, expressed in the rules of the game. In an incomplete information game, this is not the case. Games of incomplete information were not well understood until Harsanyi [117] showed that a game of incomplete information can be redefined as a game of complete information. Harsanyi's trick was to add an initial chance move in which nature makes a random choice among different sets of rules. This opening move, together with the original game of incomplete information, yields a new game which has complete but imperfect information.

The standard interpretation of an incomplete information game is the following. The players have private information.[6] In a duopoly, for example, one firm is uncertain about the cost structure (in general: the 'type') of the second firm, and vice versa. Each firm knows its own cost (its own type) and it is common knowledge that each player must belong to a set of possible types. Which specific type, however, is not known for sure. Each player has an a priori probability distribution over the sets of possible types of all other players. Thus, if firm 1 is uncertain about the unit cost of firm 2, then firm 1 makes a list of all possible values of this cost and assigns a probability to each value. Now firm 1 can determine which strategy 2 will use, given any value of 2's cost. These considerations are then used by firm 1 to design its strategy. In the course of the game, the players can use observations of the actions of their rivals to make inferences about those things they initially were uncertain about. Games of incomplete information capture in a simple way the important idea that history matters, i.e., players use the history of actions of the game to predict future behaviour, intentions, or capabilities. In recent years there has been a quite dramatic increase in the applications of incomplete information games, particularly in microeconomics and industrial organization.

Miscellaneous

The reader who wishes to proceed to more advanced material in game theory in general should consult Myerson [184], Fudenberg and Tirole [105], Friedman [100], or Osborne and Rubinstein [188].

Existence of Nash equilibrium has been established for specific classes of games. We give three examples:

[6]Player i has private information if he knows something which the other players do not know, but the other players know that player i knows something they do not know, etc.

(i) In every finite, strategic form game there is a Nash equilibrium in mixed strategies (Osborne and Rubinstein [188, proposition 33.1]). A mixed strategy (a randomized strategy) is a probability distribution over a player's set of feasible strategies (also called the pure strategies).

(ii) In every strategic form game in which feasible action sets are compact and convex and payoffs are continuous and quasi-concave, there is a Nash equilibrium (Osborne and Rubinstein [188, proposition 20.3]).

(iii) Every finite, extensive form game with perfect information has a subgame perfect equilibrium (Osborne and Rubinstein [188, proposition 99.2]).

The duopoly output game goes back to Cournot [35] who found a unique Nash equilibrium in the game. The prisoners' dilemma game is standard inventory in almost all textbooks on game theory, for example, Myerson [184]. The story goes back to Luce and Raiffa [166]. Finally, we want to point out that, to our knowledge, the first economic application of differential game theory dates back to 1925 when C. F. Roos [208] introduced his 'mathematical theory of competition' in which he studied a dynamic version of the classical Cournot model with quantity competition (see also Roos [209]).

2.5 Exercises

1. Consider the extensive form game depicted in figure 2.3. Determine the strategic form (a matrix) and find its Nash equilibria. What can backward induction or subgame perfectness tell us here?
2. Consider the static Cournot duopoly game from example 2.5 (p. 20), but assume that the inverse demand function is given by $P(Q) = 100 - 4Q + 3Q^2 - Q^3$. Furthermore, there are positive costs, given by functions $c_1(Q_1) = 4Q_1$, $c_2(Q_2) = 2Q_2 + 0.1Q_2^2$. Determine a Nash equilibrium by identifying the best reply functions.
3. Consider a two-person game with perfect information, played over two periods. Denote the players by 1 and 2. In the first period, both players simultaneously choose actions from sets A_1 and A_2, respectively. In the second period they choose actions from sets B_1 and B_2, respectively. Action sets are open subsets of the set of real numbers. Payoff functions do not change from period 1 to period 2, they are differentiable and strictly concave in a player's own action. Determine an open-loop Nash equilibrium for the two-period game. Discuss what changes would need to be made in the first order necessary conditions

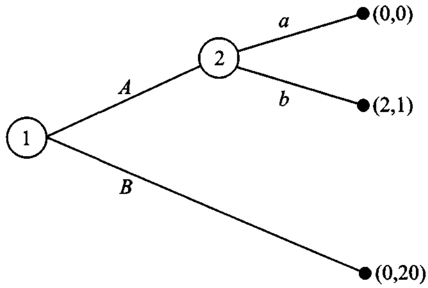

Figure 2.3. The game tree for exercise 1

if history-dependent strategies were played. Hint: Now the players must recognize – when determining their actions in period 1 – that their actions in period 2 will depend on the actions chosen in period 1.

4. Verify that ('enter','collude') and ('stay out','fight') are Nash equilibria in the game in example 2.6.
5. Invoke the principle of backward induction in example 2.7 to show that this principle leads to the Stackelberg equilibrium outcome, and not the Cournot outcome.
6. A chain store has a branch in each of N different cities. Players $1, 2, \ldots, N$ are potential entrants in each of the N cities. In each city, an entry game as described in example 2.6 is played. Recall that this game has two Nash equilibria but only one is subgame perfect. The chain store game is played in N successive stages, each stage corresponding to the game in example 2.6. Thus, in stage $i \in \{1, 2, \ldots, N\}$ entrant i decides whether or not to enter the chain store's market in city i. The chain store observes the decision of an entrant and then chooses between 'collude' and 'fight'. Any entrant i knows the outcomes of the games in cities $1, 2, \ldots, i-1$. What are the Nash equilibria of the chain store game (if there are any)? Show that there is a unique subgame perfect equilibrium. Is this equilibrium intuitive?

3 Control theoretic methods

In a differential game each player maximizes his objective functional subject to a number of constraints which include, in particular, a differential equation describing the evolution of the state of the game. Optimization problems of this type are known as optimal control problems and are widely used in economic theory and management science. The present chapter introduces two basic solution techniques for optimal control problems which are used extensively throughout the book: the Hamilton–Jacobi–Bellman equation and Pontryagin's maximum principle. We start by introducing these tools in a standard model with smooth functions and a finite time horizon and illustrate their application by an example. It is then shown that optimal solutions can be represented in many different ways and that the choice of the representation, also called the strategy, depends on the informational assumptions of the model. Sections 3.6 and 3.7 deal with generalized versions of the Hamilton–Jacobi–Bellman equation and Pontryagin's maximum principle, which are valid for optimal control problems defined on unbounded time domains and for non-smooth problems.

3.1 A simple optimal control problem

Let us assume that the differential game is defined over the time interval $[0, T]$, where $T > 0$ denotes the terminal instant of the game. All players can take actions at each time $t \in [0, T]$, thereby influencing the evolution of the state of the game as well as their own and their opponents' objective functionals. In this chapter we focus on one particular player and assume that the other players' actions are once and for all fixed.

Suppose the state of the game at each instant t can be described by an n-dimensional vector $x(t) = (x_1(t), x_2(t), \dots, x_n(t)) \in X$ where $X \subseteq \mathbb{R}^n$ is a set containing all possible states. The set X will be referred to as the state space of the game. Furthermore, assume that the evolution of the state can be described by the ordinary differential equation

$$\dot{x}(t) = f(x(t), u(t), t) \tag{3.1}$$

and the initial condition

$$x(0) = x_0 \in X, \tag{3.2}$$

where $u(t) = (u_1(t), u_2(t), \ldots, u_m(t)) \in \mathbb{R}^m$ is the vector of actions chosen by the decision maker at time t.[1] The set of all actions that are feasible at time t, if the state of the system is equal to x, is given by $U(x, t) \subseteq \mathbb{R}^m$. This implies that the decision maker has to obey the constraint

$$u(t) \in U(x(t), t). \tag{3.3}$$

The constant x_0 in (3.2) is a given initial state and the function f in (3.1) is defined on the set $\Omega = \{(x, u, t) \mid x \in X, u \in U(x, t), t \in [0, T]\}$ and takes values in $\mathbb{R}^n$ so that $f(x, u, t)$ is an n-dimensional vector $(f_1(x, u, t), f_2(x, u, t), \ldots, f_n(x, u, t))$. Equation (3.1) is the system dynamics and describes how the current state $x(t)$ and the player's actions at time t influence the rate of change of the state at time t.

Equations (3.1)–(3.3) are the constraints of the optimal control problem. Note that we do not consider $x(t) \in X$ as a separate constraint. Instead we assume that, whenever (3.1)–(3.3) hold, the validity of $x(t) \in X$ is implied. This is the case if the set $\{f(x, u, t) \mid u \in U(x, t)\}$ does not contain any directions pointing out of the state space X.[2] For example, if $X = [-1, 1]$ and $f(x, u, t) = u \in \mathbb{R}$ then it must hold that $U(-1, t) \cap (-\infty, 0) = U(1, t) \cap (0, \infty) = \emptyset$.

On a first pass, readers may think of the state $x(t)$ and the control $u(t)$ as real numbers instead of an n-dimensional vector and an m-dimensional vector, respectively. We shall state all results for the general case of vector-valued variables, but it may be easier to get the intuition for the results if one interprets x and u as single-valued variables. No notational discrimination will be made between vectors and real variables.

Example 3.1 In common property renewable resource games the state variable of the game is typically the resource stock $x \geq 0$.[3] The set of possible states at each time t is given by $X = [0, \infty)$. Let us assume that the natural growth rate of the resource is given by the function $g(x)$ so

[1]By 'decision maker' we mean the single player on which we focus this chapter.
[2]More precisely, the set $\{f(x, u, t) | u \in U(x, t)\}$ must be contained in the tangent cone of X at x.
[3]See chapter 12 for a treatment of common property resource games.

that the dynamics of the resource stock under the assumption of no harvesting can be described by the differential equation $\dot{x}(t) = g(x(t))$. Now assume that there are N players who harvest the resource stock, and denote by $u^i(t)$ the harvesting rate of player i at time t. If we neglect negative harvesting (like breeding fish and throwing them into the ocean) then we have to specify the feasible control set for each player by $U(x, t) = [0, \infty)$ if $x > 0$ and $U(x, t) = \{0\}$ if $x = 0$. The system dynamics under harvesting are given by

$$\dot{x}(t) = g(x(t)) - \sum_{i=1}^{N} u^i(t).$$

It has to be noted, however, that this equation is not necessarily the system dynamics of the control problem of any particular player. To derive the function f that has to be used by a particular player, say player 1, on the right-hand side of (3.1), we have to substitute the strategies of all other players into the above equation. For example, if the strategy of player $i \in \{2, 3, \dots, N\}$ is given by the rule $u^i(t) = \phi^i(x(t), t)$, then we obtain the system dynamics $\dot{x}(t) = f(x(t), u^1(t), t)$, where the function f is defined by $f(x, u^1, t) = g(x) - \sum_{i=2}^{N} \phi^i(x, t) - u^1$. If player 1 is a monopolist we have an optimal control problem with $f(x, u^1, t) = g(x) - u^1$.

The goal of the decision maker is to choose the control path $u : [0, T] \mapsto \mathbb{R}^m$ in an optimal way. More precisely, we assume that the decision maker has the objective functional

$$J(u(\cdot)) = \int_0^T e^{-rt} F(x(t), u(t), t)\, dt + e^{-rT} S(x(T)), \tag{3.4}$$

where $r \geq 0$ denotes the constant time preference rate (or discount rate). The term $F(x(t), u(t), t)$ measures the instantaneous utility derived by choosing the control value $u(t)$ at time t when the current state of the game is $x(t)$. The expression $S(x(T))$ denotes the terminal value associated with the state $x(T)$. Depending on the context, the function F is called utility function, felicity function, or profit function. Alternatively, one can call $-F$ the cost function or loss function. In any case, F maps the set Ω defined above into the real numbers. The function S is called scrap value function, salvage value function, or terminal value function and maps the state space X into the real numbers. Note that $F(x(t), u(t), t)$ is the rate at which profits or utility flow so that it has the dimension 'profit per unit of time' or 'utility per unit of time', respectively. The value $S(x(T))$, however, is not a rate but has the dimension 'profit' or 'utility'.

A basic optimal control problem consists of maximizing the functional J defined in (3.4) over all control paths $u(\cdot)$ which satisfy (3.3) while taking into account that the evolution of the state is determined by the system dynamics (3.1) and the initial condition (3.2).

In control theory one usually makes the assumptions that the exogenously given functions f and F are sufficiently smooth and satisfy certain boundedness conditions to ensure that solutions to (3.1)–(3.2) are uniquely defined and that the integral in (3.4) makes sense. In the framework of a differential game the functions f and F are not exogenously given but depend on the strategies of the opponents. As in example 3.1, the system dynamics for player 1's optimal control problem, $f(x, u, t)$ are defined in terms of the strategies $\phi^i(x, t)$ for $i \in \{2, 3, \ldots, N\}$. Any smoothness or boundedness assumption on f would therefore implicitly restrict the possible strategies of the players 2, 3, …, N. Since we do not want to impose strong restrictions on the set of available strategies, we cannot easily assume any restrictive properties for the functions f and F. On the other hand, it is well known that solutions to (3.1)–(3.2) need not exist or may be nonunique and that the integral in (3.4) may not be defined unless the right-hand side of (3.1) and the integrand in (3.4) behave sufficiently well. The minimal requirement, in order to deal with our control problem, is to restrict the set of control paths $u(\cdot)$ in such a way that the objective functional $J(u(\cdot))$ is well defined. This leads to the following definition.

Definition 3.1 A control path $u : [0, T] \mapsto \mathbb{R}^m$ is feasible for the optimal control problem stated above if the initial value problem (3.1)–(3.2) has a unique absolutely continuous solution $x(\cdot)$ such that the constraints $x(t) \in X$ and $u(t) \in U(x(t), t)$ hold for all t and the integral in (3.4) is well defined.[4] The control path $u(\cdot)$ is optimal if it is feasible and if the inequality $J(u(\cdot)) \geq J(\tilde{u}(\cdot))$ holds for all feasible control paths $\tilde{u}(\cdot)$.

It is clear that an optimal control problem can have none, one, or many feasible control paths and that it can have none, one, or many optimal control paths. In the following sections we present results which allow us to verify that a particular feasible control path is optimal, i.e., we are dealing with sufficient optimality conditions.[5]

[4]The function $x : [0, T] \mapsto \mathbb{R}^n$ is said to be absolutely continuous if for every $\epsilon > 0$ there exists $\delta > 0$ such that the following holds: if $(a_l, b_l), l = 1, 2, \ldots, k$, are disjoint intervals contained in $[0, T]$ such that $\Sigma_{l=1}^k (b_l - a_l) < \delta$, then it holds that $\Sigma_{l=1}^k \|x(b_l) - x(a_l)\| < \epsilon$. Absolute continuity is stronger than continuity but weaker than differentiability. The integral is called well defined if its value is a unique real number.

[5]For a general discussion of optimality conditions and, in particular, the distinction between necessary and sufficient conditions, see section 3.4 below.

3.2 The Hamilton–Jacobi–Bellman equation

The Hamilton–Jacobi–Bellman (HJB) equation lies at the heart of the dynamic programming approach to optimal control problems. This approach is based on the important principles of embedding and recursion. To explain these two principles, recall that the problem formulated in the previous section starts at time 0 in the initial state x_0. Formally, we denote that problem by $P(x_0, 0)$. The principle of embedding says that we should solve not only the given problem $P(x_0, 0)$ but rather the entire family of problems $\{P(x, t) \mid x \in X, t \in [0, T]\}$. Here, $P(x, t)$ is the problem that starts at time t in initial state x and can be stated as follows:

$$\begin{aligned} &\text{Maximize } \int_t^T e^{-r(s-t)} F(x(s), u(s), s)\, ds + e^{-r(T-t)} S(x(T)) \\ &\text{subject to } \dot{x}(s) = f(x(s), u(s), s) \\ &\qquad x(t) = x \\ &\qquad u(s) \in U(x(s), s). \end{aligned} \tag{3.5}$$

Thus problem $P(x_0, 0)$ is embedded in the family $\{P(x, t) \mid x \in X, t \in [0, T]\}$.

The principle of embedding alone does not help us in any way: it tells us to solve infinitely many problems instead of a single one. However, if combined with the principle of recursion it leads to the powerful HJB equation. Recursion means that we start at the 'smallest' problems of the entire family (these are the problems $P(x, T)$, $x \in X$) and work our way backwards to the 'largest' problems, which are the problems $P(x, 0)$, $x \in X$. The knowledge of the solution of all small problems will help to find the solution of any larger problem. Let us explain this in more detail.

The problems $P(x, T)$ are trivial because they are not decision problems. As a matter of fact, the objective functional in (3.5) collapses to $S(x)$ if $t = T$. Since x is the exogenously given initial state of $P(x, T)$, the decision maker cannot influence this value at all. For further reference let us denote the only feasible (and hence the optimal) value of the objective functional of $P(x, T)$ by $V(x, T)$, i. e., $V(x, T) = S(x)$. Analogously, we denote by $V(x, t)$ the optimal value of the objective functional of problem $P(x, t)$ in (3.5). We now present an intuitive argument for the fact that the optimal value function V satisfies the partial differential equation

$$rV(x, t) - V_t(x, t) = \max\{F(x, u, t) + V_x(x, t) f(x, u, t) | u \in U(x, t)\},$$

which is called the Hamilton–Jacobi–Bellman equation.[6]

Suppose that you want to solve $P(x, t)$, where $x \in X$ and $t < T$. In other words, the system is observed to be in state x and your clock shows t. Instead of attempting the formidable task of determining the entire control path $u : [t, T] \mapsto \mathbb{R}^m$ it may be wiser just to consider the immediate future $[t, t + \Delta)$ where $\Delta > 0$ is a very small but positive real number. Every feasible control path $u : [t, t + \Delta) \mapsto \mathbb{R}^m$ takes the system from the present state x to some state $x(t + \Delta)$ as described by the differential equation (3.1). If you behave optimally from $t + \Delta$ onwards, the total utility derived during the interval $[t + \Delta, T]$, discounted back to time $t + \Delta$, is given by $V(x(t + \Delta), t + \Delta)$. This follows simply from the way in which we have defined the optimal value function V. Therefore, the total utility derived by choosing $u : [t, t + \Delta) \mapsto \mathbb{R}^m$ until time $t + \Delta$, and behaving optimally after time $t + \Delta$, discounted back to the present time t, is equal to

$$\int_t^{t+\Delta} e^{-r(s-t)} F(x(s), u(s), s)\, ds + e^{-r\Delta} V(x(t + \Delta), t + \Delta).$$

If we choose the control path $u : [t, t + \Delta) \mapsto \mathbb{R}^m$ optimally, the discounted utility should equal the maximal utility we can obtain, $V(x, t)$. This leads to the equation

$$\begin{aligned} V(x, t) = \max\Bigg\{ & \int_t^{t+\Delta} e^{-r(s-t)} F(x(s), u(s), s)\, ds \\ & + e^{-r\Delta}\ V(x(t + \Delta), t + \Delta)\Bigg\}, \end{aligned}$$

where the maximum is taken with respect to all feasible control paths $u : [t, t + \Delta) \mapsto \mathbb{R}^m$ and subject to the constraints $\dot{x}(s) = f(x(s), u(s), s)$, $x(t) = x$, and $u(s) \in U(x(s), s)$ for all $s \in [t, t + \Delta)$. Now subtract $V(x, t)$ from both sides of the above equation and divide the resulting equation by Δ. This yields

$$\begin{aligned} 0 = \max\Bigg\{ & \frac{1}{\Delta} \int_t^{t+\Delta} e^{-r(s-t)} F(x(s), u(s), s)\, ds \\ & + \frac{e^{-r\Delta} V(x(t + \Delta), t + \Delta) - V(x, t)}{\Delta}\Bigg\}. \end{aligned} \tag{3.6}$$

[6]Because V is a real valued function and x is an n-dimensional vector, the gradient $V_x(x, t)$ is also a vector with n components. The term $V_x(x, t) f(x, u, t)$ is therefore the scalar product of two n-dimensional vectors.

Assuming that all functions appearing in (3.6) are sufficiently smooth we can ask the question of what happens if Δ approaches 0. From the mean value theorem we know that $\lim_{\Delta\to 0}(1/\Delta)\int_t^{t+\Delta} e^{-r(s-t)}F(x(s), u(s), s)ds = F(x(t), u(t), t)$. To derive the limit of the second term on the right-hand side of (3.6) we use familiar rules of differential calculus to obtain

$$\begin{aligned}
&\lim_{\Delta\to 0}\frac{e^{-r\Delta}V(x(t+\Delta), t+\Delta) - V(x, t)}{\Delta}\\
&= \frac{d}{d\Delta}\left[e^{-r\Delta}V(x(t+\Delta), t+\Delta)\right]\Big|_{\Delta=0}\\
&= -rV(x(t), t) + V_x(x(t), t)\dot{x}(t) + V_t(x(t), t).
\end{aligned}$$

Putting everything together, we see that for $\Delta \to 0$ equation (3.6) becomes

$$0 = \max\left\{F(x(t), u(t), t) - rV(x(t), t) + V_x(x(t), t)\dot{x}(t) + V_t(x(t), t)\right\}.$$

Recall that the maximization has to be carried out with respect to all feasible control paths $u : [t, t+\Delta) \to \mathbb{R}^m$ and subject to $\dot{x}(s) = f(x(s), u(s), s)$, $x(t) = x$, and $u(s) \in U(x(s), s)$ for $s \in [t, t+\Delta)$. Since we let Δ go to 0, our only decision variable is the control value at time t, $u(t) = u$. Hence, by substituting for $\dot{x}(t)$ from (3.1) we obtain

$$0 = \max\{F(x, u, t) - rV(x, t) + V_x(x, t)f(x, u, t) + V_t(x, t)|u \in U(x, t)\}.$$

This equation is identical to the HJB equation as stated above and our heuristic argument for the validity of the equation is therefore complete.

Why is this a heuristic argument? Well, the most important reason is that we have made smoothness assumptions which are not necessarily satisfied. In particular, we have not made sure that the partial derivatives of the optimal value function occurring in the HJB equation exist. It is quite plausible that the optimal value function V is not smoother than the problem fundamentals f and F and, in a differential game, these functions depend on the strategies of the rival players on which we cannot impose restrictive assumptions. But even if f and F are infinitely many times differentiable it can happen that the optimal value function is not differentiable. This is illustrated by the following example.

Example 3.2 Consider the problem $P(x, t)$ defined as follows:

$$\begin{aligned}
&\text{Maximize } \int_t^T x(s)u(s)\,ds \\
&\text{subject to } \dot{x}(s) = u(s), \\
&\qquad x(t) = x, \\
&\qquad u(s) \in [-1, 1].
\end{aligned}$$

The state space is $X = \mathbb{R}$ and the discount rate is $r = 0$. Note that the functions $F(x, u, t) = xu$ and $f(x, u, t) = u$ are infinitely many times differentiable. To find the optimal value function we observe that for every feasible control path

$$\int_t^T x(s)u(s)\,ds = \int_t^T x(s)\dot{x}(s)\,ds = x(s)^2/2\Big|_{s=t}^{T} = \left[x(T)^2 - x^2\right]/2.$$

To maximize the objective functional we should therefore maximize the distance between the numbers $x(T)^2$ and x^2. This can be achieved by choosing $u(s) = 1$ for all $s \in [t, T]$ if $x \geq 0$, and $u(s) = -1$ for all $s \in [t, T]$ if $x \leq 0$. The optimal final state is therefore given by $x(T) = x + T - t$ if $x > 0$, $x(T) = x - T + t$ if $x < 0$, and either of the two values if $x = 0$. It follows that the optimal value function V is given by

$$V(x, t) = \begin{cases} [(x + T - t)^2 - x^2]/2 & \text{if } x \geq 0, \\ [(x - T + t)^2 - x^2]/2 & \text{if } x \leq 0. \end{cases}$$

This function can also be written as $V(x, t) = (T - t)^2/2 + (T - t)|x|$ and it is easily seen that it is not differentiable at $x = 0$ whenever $t < T$.

Because of examples like this one we cannot obtain a theorem which states that the optimal value function V is continuously differentiable and solves the HJB equation. There are various ways out of this problem. First, the HJB equation can be derived as a necessary optimality condition when the partial derivatives V_x and V_t are replaced by weaker forms of derivatives. This would require a level of mathematical sophistication above that of this book. Another possibility would be to state the HJB equation only as a sufficient optimality condition. This way we could simply assume that V is continuously differentiable, although we would not be able to handle many interesting applications in a rigorous manner. Our approach is somewhere in between these two possibilities. In this section we present the HJB equation as a sufficient optimality condition under the assumption that the optimal value function is continuously differentiable. In section 3.7 we introduce more general forms of derivatives and state a sufficiency theorem for optimal control models which covers most of the examples occurring in the economics and management literature.[7]

Theorem 3.1 Let $V : X \times [0, T] \mapsto \mathbb{R}$ *be a continuously differentiable function which satisfies the HJB equation*

$$rV(x, t) - V_t(x, t) = \max\{F(x, u, t) + V_x(x, t)f(x, u, t) \mid u \in U(x, t)\} \tag{3.7}$$

and the terminal condition

$$V(x, T) = S(x) \tag{3.8}$$

for all $(x, t) \in X \times [0, T]$. *Let* $\Phi(x, t)$ *denote the set of controls* $u \in U(x, t)$ *maximizing the right-hand side of (3.7). If* $u(\cdot)$ *is a feasible control path with corresponding state trajectory* $x(\cdot)$ *and if* $u(t) \in \Phi(x(t), t)$ *holds for almost all* $t \in [0, T]$ *then* $u(\cdot)$ *is an optimal control path.*[8] *Moreover,* $V(x, t)$ *is the optimal value of problem* $P(x, t)$.

Proof Let $\tilde{u}(\cdot)$ be any feasible control path with corresponding state trajectory $\tilde{x}(\cdot)$. We have to show that $J(u(\cdot)) \geq J(\tilde{u}(\cdot))$. Because of the HJB equation and the feasibility of $\tilde{u}(\cdot)$ it holds for all $t \in [0, T]$ that $F(\tilde{x}(t), \tilde{u}(t), t) \leq rV(\tilde{x}(t), t) - V_t(\tilde{x}(t), t) - V_x(\tilde{x}(t), t)f(\tilde{x}(t), \tilde{u}(t), t)$. Multiplying by e^{-rt} and using the system dynamics $\dot{\tilde{x}}(t) = f(\tilde{x}(t), \tilde{u}(t), t)$ we see that this inequality can be written as

$$e^{-rt}F(\tilde{x}(t), \tilde{u}(t), t) \leq -\frac{d}{dt}\left[e^{-rt}V(\tilde{x}(t), t)\right]. \tag{3.9}$$

Since $u(\cdot)$ is also a feasible control path the same inequality holds if we replace $\tilde{x}(t)$ and $\tilde{u}(t)$ by $x(t)$ and $u(t)$, respectively. However, because of the assumption $u(t) \in \Phi(x(t), t)$ the equality sign instead of the inequality sign must hold for almost all $t \in [0, T]$, i.e.,

$$e^{-rt}F(x(t), u(t), t) = -\frac{d}{dt}\left[e^{-rt}V(x(t), t)\right]. \tag{3.10}$$

Substituting these two relations into the objective functional (3.4) and using (3.8) we obtain

[7] We do not formally prove the HJB equation as a necessary optimality condition at any place in this book. See section 3.4 for a further justification for restricting ourselves to sufficient optimality conditions.

[8] A condition holds 'for almost all $t \in [0, T]$' if the set of those t-values in $[0, T]$ where it does not hold has Lebesgue measure 0. This is true, for example, if the condition holds for all but finitely many values of $t \in [0, T]$.

$$
\begin{aligned}
&J(u(\cdot)) - J(\tilde{u}(\cdot)) \\
&= \int_0^T e^{-rt}[F(x(t), u(t), t) - F(\tilde{x}(t), \tilde{u}(t), t)]dt + e^{-rT}[S(x(T)) - S(\tilde{x}(T))] \\
&\geq \int_0^T \frac{d}{dt}\{e^{-rt}[V(\tilde{x}(t), t) - V(x(t), t)]\}dt + e^{-rT}[V(x(T), T) - V(\tilde{x}(T), T)] \\
&= V(x(0), 0) - V(\tilde{x}(0), 0).
\end{aligned}
$$

The expression on the last line of this formula is equal to 0 because the two state trajectories are feasible and, therefore, must start at the same initial point $x(0) = \tilde{x}(0) = x_0$. This shows $J(u(\cdot)) \geq J(\tilde{u}(\cdot))$.

The last assertion of the theorem follows by integrating (3.10) over the time interval $[t, T]$ and using (3.8). ■

Note that the condition $u(t) \in \Phi(x(t), t)$ involves only the values $u(t)$ and $x(t)$ along the candidate path $u(\cdot)$ and its state trajectory $x(\cdot)$, respectively, whereas the HJB equation (3.7) and the boundary condition (3.8) have to be satisfied for all $(x, t) \in X \times [0, T]$. Thus, it is not sufficient to verify the HJB equation only along the candidate trajectory. To avoid this common pitfall we recommend that any HJB equation is always stated in the form (3.7), i.e., '$rV(x, t) - V_t(x, t) = \ldots$' and not in the form '$rV(x(t), t) - V_t(x(t), t) = \ldots$', since the latter gives the wrong impression that the HJB equation must hold only along a certain trajectory $x(\cdot)$.

3.3 Pontryagin's maximum principle

Pontryagin's maximum principle is a necessary optimality condition for optimal control problems which has found many applications in economics and management science. It is a first order condition for smooth problems, comparable to the condition that the gradient vector of a function $g : \mathbb{R}^n \mapsto \mathbb{R}$ must vanish at a local maximum of g. It is well known that the latter condition is also satisfied at local minima and other critical points. Only if we have some additional information on the global curvature properties of g (like concavity of g) can we infer from the condition $g'(x) = 0$ that x is indeed a maximum. The situation is similar in the case of an optimal control problem. Even if a certain control path $u(\cdot)$ satisfies the maximum principle it need not be an optimal path. If the problem has certain curvature properties, however, then any control path which satisfies the maximum principle is an optimal path. We now discuss this augmented maximum principle as a sufficient optimality condition for the control problem of section 3.1. As in the previous section we first present a heuristic argument in order to motivate the conditions and then give a rigorous proof of their sufficiency.

Let us start by defining a real-valued function H by[9]

$$H(x, u, \lambda, t) = F(x, u, t) + \lambda f(x, u, t).$$

The domain of H is the set $\{(x, u, \lambda, t) \mid x \in X, u \in U(x, t), \lambda \in \mathbb{R}^n, t \in [0, T]\}$. The function H is called the (current-value) Hamiltonian function and plays a prominent role in Pontryagin's maximum principle. The variable λ is called the (current-value) costate variable associated with the state variable x, or the (current-value) adjoint variable. Finally, we define the maximized Hamiltonian function $H^* : X \times \mathbb{R}^n \times [0, T] \mapsto \mathbb{R}$ by

$$H^*(x, \lambda, t) = \max\{H(x, u, \lambda, t) \mid u \in U(x, t)\}.$$

Now consider the HJB equation (3.7). We have seen in theorem 3.1 that controls which maximize the right-hand side of (3.7) are optimal controls. For the sake of our heuristic argument let us assume that the maximum on the right-hand side is attained at the unique point $\phi(x, t) \in U(x, t)$. Using the definition of the maximized Hamiltonian function we can rewrite the HJB equation as

$$rV(x, t) - V_t(x, t) = H^*(x, V_x(x, t), t) = H(x, \phi(x, t), V_x(x, t), t).$$

Since this equation must hold for all $(x, t) \in X \times [0, T]$ we can differentiate it with respect to x to obtain[10]

$$rV_x(x, t) - V_{tx}(x, t) = H_x^*(x, V_x(x, t), t) + V_{xx}(x, t)H_\lambda^*(x, V_x(x, t), t). \tag{3.11}$$

There is a mathematical result called the envelope theorem which says that under our assumptions it must hold that $H_\lambda^*(x, V_x(x, t), t) = H_\lambda(x, \phi(x, t), V_x(x, t), t)$; that is, the derivative of the maximized Hamiltonian with respect to λ is equal to the derivative of the Hamiltonian with respect to λ if the latter is evaluated at the maximizing point $u = \phi(x, t)$. Moreover, from the definition of the Hamiltonian function we see that $H_\lambda(x, \phi(x, t), V_x(x, t), t) = f(x, \phi(x, t), t)$. Using these results we can rewrite (3.11) as

$$rV_x(x, t) - V_{tx}(x, t) = H_x^*(x, V_x(x, t), t) + V_{xx}(x, t)f(x, \phi(x, t), t). \tag{3.12}$$

[9]Note that the second term on the right-hand side of this equation is the scalar product of the two vectors $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $f(x, u, t) = (f_1(x, u, t), f_2(x, u, t), \ldots, f_n(x, u, t))$.

[10]Because the vector x has n components, the Hessian matrix of second order derivatives, $V_{xx}(x, t)$, is an $n \times n$ matrix. The last term in equation (3.11) is the product of this $n \times n$ matrix with the n-dimensional gradient vector $H_\lambda^*(x, V_x(x, t), t)$.

Assume that $u(\cdot)$ is an optimal control path with corresponding state trajectory $x(\cdot)$ and define the costate trajectory $\lambda : [0, T] \mapsto \mathbb{R}^n$ by $\lambda(t) = V_x(x(t), t)$. Since optimal controls maximize the right-hand side of the HJB equation (3.7) for almost all $t \in [0, T]$, we have $u(t) = \phi(x(t), t)$ and the following condition must hold for almost all $t \in [0, T]$

$$H(x(t), u(t), \lambda(t), t) = H^*(x(t), \lambda(t), t).$$

That is, for any given triple $(x(t), \lambda(t), t)$ with $\lambda(t) = V_x(x(t), t)$, an optimal control must maximize the Hamiltonian function in the set of feasible controls, $U(x(t), t)$. We call this equation the maximum condition. In the next step we derive a differential equation for the costate trajectory $\lambda(\cdot)$. To this end, differentiate in the definition $\lambda(t) = V_x(x(t), t)$ totally with respect to t to obtain

$$\begin{aligned}\dot{\lambda}(t) &= \frac{d}{dt} V_x(x(t), t)\\ &= V_{xx}(x(t), t)\dot{x}(t) + V_{xt}(x(t), t)\\ &= V_{xx}(x(t), t) f(x(t), \phi(x(t), t), t) + V_{xt}(x(t), t).\end{aligned}$$

Here we have used the system dynamics (3.1) and the fact that $u(t) = \phi(x(t), t)$. Since $V_{xt}(x, t) = V_{tx}(x, t)$ if V is sufficiently smooth, the above equation and (3.12) yield

$$\dot{\lambda}(t) = r\lambda(t) - H_x^*(x(t), \lambda(t), t).$$

This is called the adjoint equation or the costate equation of the optimal control problem. If the feasible set $U(x, t)$ is independent of x, that is, if for all $t \in [0, T]$ there exists a set $\tilde{U}(t) \subseteq \mathbb{R}^m$ such that $U(x, t) = \tilde{U}(t)$ holds for all $x \in X$, then the adjoint equation becomes

$$\dot{\lambda}(t) = r\lambda(t) - H_x(x(t), u(t), \lambda(t), t).$$

This is so because from the envelope theorem it follows that in the case where $U(x, t) = \tilde{U}(t)$ is satisfied for all $x \in X$, the identity $H_x^*(x, \lambda, t) = H_x(x, \phi(x, t), \lambda, t)$ must hold. The latter form of the adjoint equation is sometimes more convenient to use than the adjoint equation involving the maximized Hamiltonian, as will be seen in the applications in part II.

Finally, by differentiating (3.8) with respect to x, and using the definition of $\lambda(t)$, we obtain

$$\lambda(T) = S'(x(T)),$$

which is called the transversality condition.

The maximum principle states that under certain assumptions there exists for every optimal control path $u(\cdot)$ a costate trajectory $\lambda(\cdot)$ such that the maximum condition, the adjoint equation, and the transversality condition are satisfied. To obtain a sufficiency theorem we augment these conditions by convexity assumptions. This yields the following theorem.

Theorem 3.2 *Consider the optimal control problem of section 3.1 and define the Hamiltonian function H and the maximized Hamiltonian function H^* as above. Assume that the state space X is a convex set and that the scrap value function S is continuously differentiable and concave. Let $u(\cdot)$ be a feasible control path with corresponding state trajectory $x(\cdot)$. If there exists an absolutely continuous function $\lambda : [0, T] \mapsto \mathbb{R}^n$ such that the maximum condition*

$$H(x(t), u(t), \lambda(t), t) = H^*(x(t), \lambda(t), t), \tag{3.13}$$

the adjoint equation

$$\dot{\lambda}(t) = r\lambda(t) - H_x^*(x(t), \lambda(t), t), \tag{3.14}$$

and the transversality condition

$$\lambda(T) = S'(x(T)) \tag{3.15}$$

are satisfied, and such that the function $x \mapsto H^*(x, \lambda(t), t)$ is concave and continuously differentiable with respect to x for all $t \in [0, T]$, then $u(\cdot)$ is an optimal path. If the set of feasible controls, $U(x, t)$, does not depend on x, this result remains true if equation (3.14) is replaced by

$$\dot{\lambda}(t) = r\lambda(t) - H_x(x(t), u(t), \lambda(t), t). \tag{3.16}$$

Proof Let $\tilde{u}(\cdot)$ be an arbitrary feasible control path with corresponding state trajectory $\tilde{x}(\cdot)$. We have to show that $J(u(\cdot)) \geq J(\tilde{u}(\cdot))$. Using the definition of the Hamiltonian function and the system dynamics (3.1) we obtain

$$\begin{aligned}
&J(u(\cdot)) - J(\tilde{u}(\cdot)) \\
&= \int_0^T e^{-rt}F(x(t), u(t), t)\,dt - \int_0^T e^{-rt}F(\tilde{x}(t), \tilde{u}(t), t)\,dt + e^{-rT}[S(x(T)) - S(\tilde{x}(T))] \\
&= \int_0^T e^{-rt}[H(x(t), u(t), \lambda(t), t) - \lambda(t)\dot{x}(t)]dt \\
&- \int_0^T e^{-rt}\Big[H(\tilde{x}(t), \tilde{u}(t), \lambda(t), t) - \lambda(t)\dot{\tilde{x}}(t)\Big]dt + e^{-rT}[S(x(T)) - S(\tilde{x}(T))].
\end{aligned}$$

Because of the feasibility of $\tilde{u}(\cdot)$ and the definition of H^* we have $H(\tilde{x}(t), \tilde{u}(t), \lambda(t), t) \leq H^*(\tilde{x}(t), \lambda(t), t)$. Using this inequality and the maximum condition (3.13) in the above equation we obtain

$$\begin{aligned} &J(u(\cdot)) - J(\tilde{u}(\cdot)) \\ &\geq \int_0^T e^{-rt}[H^*(x(t), \lambda(t), t) - \lambda(t)\dot{x}(t)]dt \\ &- \int_0^T e^{-rt}\Big[H^*(\tilde{x}(t), \lambda(t), t) - \lambda(t)\dot{\tilde{x}}(t)\Big]dt + e^{-rT}[S(x(T)) - S(\tilde{x}(T))]. \end{aligned}$$

The differentiability and concavity of the function $H^*(x, \lambda(t), t)$ with respect to x imply $H^*(x(t), \lambda(t), t) - H^*(\tilde{x}(t), \lambda(t), t) \geq H^*_x(x(t), \lambda(t), t)$ $[x(t) - \tilde{x}(t)]$. Because of the adjoint equation (3.14) this can be written as $H^*(x(t), \lambda(t), t) - H^*(\tilde{x}(t), \lambda(t), t) \geq [r\lambda(t) - \dot{\lambda}(t)][x(t) - \tilde{x}(t)]$. Using this in the above inequality we obtain

$$\begin{aligned} &J(u(\cdot)) - J(\tilde{u}(\cdot)) \\ &\geq \int_0^T e^{-rt}\Big\{[r\lambda(t) - \dot{\lambda}(t)][x(t) - \tilde{x}(t)] - \lambda(t)[\dot{x}(t) - \dot{\tilde{x}}(t)]\Big\}dt \\ &+ e^{-rT}[S(x(T)) - S(\tilde{x}(T))] \\ &= \int_0^T \frac{d}{dt}\{e^{-rt}\lambda(t)[\tilde{x}(t) - x(t)]\}dt + e^{-rT}[S(x(T)) - S(\tilde{x}(T))] \\ &= e^{-rT}\{\lambda(T)[\tilde{x}(T) - x(T)] + S(x(T)) - S(\tilde{x}(T))\} - \lambda(0)[\tilde{x}(0) - x(0)]. \end{aligned}$$

The last term on the last line of this formula is equal to 0 because $x(0) = \tilde{x}(0) = x_0$ (both state trajectories are feasible). To complete the proof of optimality of $u(\cdot)$ note that differentiability and concavity of S, together with condition (3.15), imply that the first term on the last line of the above formula is nonnegative.

The proof of the theorem with (3.14) replaced by (3.16) is not carried out here. It follows quite easily from the envelope theorem. ■

The heuristic argument used to motivate the conditions of theorem 3.2 has a useful by product. It provides us with an intuitive economic interpretation of the adjoint function λ. Indeed, we have seen that $\lambda(t) = V_x(x(t), t)$, where V is the value function of the optimal control problem. Therefore, $\lambda(t)$ measures the marginal utility of the state at time t along the optimal trajectory. In other words, $\lambda(t)$ is the highest hypothetical price which a rational decision maker would be willing to pay for an additional, infinitesimally small unit of the state variable at time t. Because of this, the adjoint variable is often called the shadow price of x. Depending on the structure of the model, the shadow price can be positive, negative, or equal to 0.

The proof of theorem 3.2 also gives us a hint on how transversality conditions for slightly modified problems can be derived. The basic inequality that we have derived in the proof is

$$J(u(\cdot)) - J(\tilde{u}(\cdot)) \geq e^{-rT}\{\lambda(T)[\tilde{x}(T) - x(T)] + S(x(T)) - S(\tilde{x}(T))\} - \lambda(0)[\tilde{x}(0) - x(0)].$$

Assume now that the initial state is not fixed but that it is free. In this case we must have $\lambda(0) = 0$ for the last term on the right-hand side to become zero independently of the choice of $\tilde{x}(0)$. Thus, for a problem with a free initial state we have an additional transversality condition

$$\lambda(0) = 0. \tag{3.17}$$

Other cases can be dealt with in an analogous way. If, for example, only some initial states $x_i(0)$, $i \in \{1, 2, \ldots, k-1\}$, are fixed but the remaining initial states $x_i(0)$, $i \in \{k, k+1, \ldots, n\}$, are free then we have to require $\lambda_i(0) = 0$ only for $i \in \{k, k+1, \ldots, n\}$ and there is no transversality condition at $t = 0$ for $i \in \{1, 2, \ldots, k-1\}$. One could also consider problems in which the initial state $x(0)$ is free and the objective functional contains an additional term depending on $x(0)$, that is,

$$J(u(\cdot)) = G(x(0)) + \int_0^T e^{-rt} F(x(t), u(t), t)\, dt + e^{-rT} S(x(T))$$

for some function $G : X \mapsto \mathbb{R}$. Using exactly the same logic as above one can show that theorem 3.2 remains valid if G is concave and differentiable and the additional transversality condition

$$\lambda(0) = -G'(x(0)) \tag{3.18}$$

holds. The reader should not find it difficult to derive appropriate transversality conditions for other variations of the standard model.

Instead of the current-value Hamiltonian some authors use the so-called present-value Hamiltonian function which is defined by

$$\tilde{H}(x, u, \tilde{\lambda}, t) = e^{-rt} F(x, u, t) + \tilde{\lambda} f(x, u, t).$$

The maximized present-value Hamiltonian function is defined by

$$\tilde{H}^*(x, \tilde{\lambda}, t) = \max\{\tilde{H}(x, u, \tilde{\lambda}, t) \mid u \in U(x, t)\}.$$

Using the present-value approach the conditions of theorem 3.2 have to be modified in the following way. The maximum condition is $\tilde{H}(x(t), u(t), \tilde{\lambda}(t), t) = \tilde{H}^*(x(t), \tilde{\lambda}(t), t)$ and the present-value costate variable satisfies the adjoint equation $\dot{\tilde{\lambda}}(t) = -\tilde{H}_x^*(x(t), \tilde{\lambda}(t), t)$ and the transversality condition $\tilde{\lambda}(T) = e^{-rT}S'(x(T))$. As before, the adjoint equation can also be written as $\dot{\tilde{\lambda}}(t) = -\tilde{H}_x(x(t), u(t), \tilde{\lambda}(t), t)$ when $U(x, t)$ does not depend on x. It is easy to verify that the above conditions are equivalent to those stated in theorem 3.2 and that the present-value costate variable $\tilde{\lambda}(t)$ is related to the current-value costate variable $\lambda(t)$ by the equation $\tilde{\lambda}(t) = e^{-rt}\lambda(t)$.

3.4 How to solve an optimal control problem

In this section we give some ideas on how to solve optimal control problems and illustrate them by means of a worked-out example.

Basically, there are three approaches to solving any optimization problem. The first one is to make it obvious what the optimal solution must be. This can be done, for example, by rewriting the problem in a clever way – an approach that was used in example 3.2 above. Of course, it works only for very simple problems and requires a good intuition about what the solution might look like. The other two approaches are more systematic and can, in principle, be applied to problems of arbitrarily complex structure. One of them uses necessary optimality conditions whereas the other is based on sufficient optimality conditions.

Consider first the approach which uses necessary optimality conditions, i.e., conditions that every optimal solution must satisfy. Let us call a feasible solution which satisfies a certain set of necessary optimality conditions an n-candidate. It follows from the very definition of necessity that any optimal solution must be an n-candidate. Consequently, we can restrict our search for optimal solutions to the set of n-candidates. If there are only a few of them, we can simply compare the objective values of the n-candidates and pick the one with the largest value. It is clear that stronger necessary optimality conditions yield fewer n-candidates so that this approach works the better the stronger the conditions are. The alert reader, however, will notice that this approach may fail even if the set of n-candidates is very small. The problem is that there might not exist an optimal solution at all. Nonexistence of an optimal solution can manifest itself in an empty set of n-candidates or, even worse, in a nonempty set of n-candidates, none of which is an optimal solution. In the former case we do realize the nonexistence of optimal solutions but in the latter we might think that the n-candidate with the highest objective value is an optimal solution while it may in fact be a very poor solution. The lesson to be learned is that the approach based on necessary optimality conditions should only be used in conjunction with an existence theorem for optimal solutions.

The other systematic way to solve an optimization problem is to apply sufficient optimality conditions. Again one tries to find feasible solutions which satisfy the sufficient conditions (s-candidates). In contrast to the approach based on necessary conditions, we know that every s-candidate is guaranteed to be an optimal solution. The problem with this approach is that we could miss some, or all, of the optimal solutions if the sufficient optimality conditions are too strong.

Actually, there is a fourth approach to solving an optimization problem. This approach will frequently be used in the present book. It consists of an 'informed guessing stage' and a 'verification stage'. In the former stage one uses intuition, heuristic arguments, experience with similar problems, and necessary or 'almost necessary' optimality conditions to find a set of candidates. In the second stage one uses sufficient optimality conditions or clever transformations to verify that a certain candidate is indeed an optimal solution. Let us illustrate this approach by means of an example.

Example 3.3 Consider the following optimal control problem with state space $X = [0, \infty)$:

$$\text{Maximize } \int_0^T e^{-rt}\left[-x(t) - \frac{\alpha}{2}u(t)^2\right]dt$$
$$\text{subject to } \dot{x}(t) = \beta(t) - u(t)\sqrt{x(t)},$$
$$u(t) \geq 0, \qquad x(0) = x_0,$$

where α, r, T, and x_0 are positive constants and $\beta : [0, T] \mapsto \mathbb{R}$ is a positive-valued function.

One may interpret this control problem as one of finding an optimal maintenance policy for a machine, building, or piece of equipment which is subject to continuous deterioration. The variable $x(t)$ measures the state of deterioration at time t whereas $u(t)$ denotes the intensity, or rate, at which maintenance activities are carried out at time t. If the equipment is not maintained at all it deteriorates at the (time dependent) rate $\beta(t)$. By using a positive rate of maintenance one can slow down, or even reverse, the process of deterioration. This explains the system dynamics of the model. The fact that $-u(t)$ on the right-hand side of the system dynamics is multiplied by an increasing function of $x(t)$ captures the assumption of economies of scale for the maintenance activities: the higher the state of deterioration is, the more efficient is one unit of maintenance effort. The reason why we have chosen the increasing function of $x(t)$ to be the square root function is analytical tractability. Hence, we do not give an economic interpretation of this particular choice. The objective functional of the optimal control problem is defined in such a way that the state of deterioration leads to linear costs whereas the maintenance activities incur quadratic costs. The goal is to minimize the present value of the sum of these costs over the finite time interval $[0, T]$.

Faced with such a control problem, it should first be ascertained that it is indeed a nontrivial problem. Scrutinizing the objective functional we see that it would be best if the state $x(t)$ and the control $u(t)$ both were equal to 0. Thus, one conjecture for a trivial solution is $u(t) = 0$ for all t. However, we see from the system dynamics that if $u(t) = 0$ for all t then the state trajectory $x(\cdot)$ will be increasing because the function β has positive values. This will make the integrand smaller so that $u(t) = 0$ for all t is most likely not an optimal control path. On the other hand, we could try to keep $x(t)$ as close to 0 as possible. Since $x_0 > 0$ this requires choosing $u(t)$ very high for small t, which also decreases the value of the objective functional. Thus, it seems that the optimal solution must consist of some (nontrivial) compromise between keeping $x(t)$ close to 0 and keeping $u(t)$ close to 0.

Let us continue the informed guessing stage by using Pontryagin's maximum principle. We argued that under certain assumptions conditions (3.13), (3.14), and (3.15) are necessary for the optimality of a feasible control path.[11] So let us try to find candidates which satisfy these conditions. The Hamiltonian function is given by

$$H(x, u, \lambda, t) = -x - \frac{\alpha}{2}u^2 + \lambda\big[\beta(t) - u\sqrt{x}\big].$$

According to condition (3.13) this function should be maximized with respect to $u \geq 0$, which yields $u = \max\{0, -\lambda\sqrt{x}/\alpha\}$. Along an optimal control path we cannot have $u(t) = 0$ except possibly at $t = T$. This follows because the marginal cost of control at $u = 0$ is equal to 0 but the marginal effect on the rate of change of the state is $\sqrt{x} > 0$. One could also arrive at this conjecture by using the intuitive argument from above concerning the nontriviality of the problem. Yet another intuitive reason for $u(t) > 0$ is that λ, the shadow price of the state, must be negative for all $t < T$ since the state of deterioration of the equipment has a negative influence on overall utility. However, if $\lambda < 0$ then $-\lambda\sqrt{x}/\alpha > 0$ so that the maximum of 0 and $-\lambda\sqrt{x}/\alpha$ is strictly positive.[12] Therefore, we con-

[11] In the particular example under consideration these assumptions are indeed satisfied, but let us proceed as if we did not know this fact.

[12] If you do not have any experience with or a good intuition for problems of this kind you should consider the possibility $u(t) = 0$ as well. You will discover that it does not satisfy sufficient optimality conditions.

jecture that the optimal control $u(t)$ is in the interior of the feasible set and given by

$$u(t) = -\lambda(t)\sqrt{x(t)}/\alpha. \tag{3.19}$$

Substituting this into the Hamiltonian function we obtain the maximized Hamiltonian

$$H^*(x, \lambda, t) = -x + \frac{1}{2\alpha}\lambda^2 x + \beta(t)\lambda.$$

The adjoint equation (3.14) and the transversality condition (3.15) yield the terminal value problem

$$\dot{\lambda}(t) = 1 + r\lambda(t) - \frac{1}{2\alpha}\lambda(t)^2, \qquad \lambda(T) = 0. \tag{3.20}$$

Note that instead of (3.14) we could also have used the second form of the adjoint equation, (3.16), together with (3.19), to obtain the same result because the set of feasible controls, $U(x, t) = [0, \infty)$, does not depend on the state variable x. Equation (3.20) is a differential equation of the Riccati type which can be solved explicitly. The unique solution is

$$\lambda(t) = \tilde{\lambda}(t) := \frac{2\left[1 - e^{C(t-T)}\right]}{(r - C)e^{C(t-T)} - (r + C)}, \tag{3.21}$$

where $C = \sqrt{r^2 + 2/\alpha}$. Since $C > r > 0$ it is easily seen that $\lambda(t) \leq 0$ holds for all $t \in [0, T]$, which confirms our intuition about the negativity of the shadow price. It follows that $u(t)$ defined in (3.19) is nonnegative; hence it maximizes the Hamiltonian $H(x(t), u, \lambda(t), t)$ over all feasible $u \geq 0$. Substituting (3.19) and (3.21) into the system dynamics we obtain the linear, nonautonomous initial value problem for the state trajectory

$$\dot{x}(t) = \frac{1}{\alpha}\tilde{\lambda}(t)x(t) + \beta(t), \qquad x(0) = x_0.$$

This differential equation, too, can be solved explicitly (for example by the variation of constants method). This yields

$$x(t) = \tilde{x}(t) := e^{\int_0^t \tilde{\lambda}(\tau)/\alpha\, d\tau}\left[\int_0^t \beta(s)e^{-\int_0^s \tilde{\lambda}(\tau)/\alpha\, d\tau}\, ds + x_0\right]. \tag{3.22}$$

To summarize, we have obtained a unique candidate $u(\cdot)$ for an optimal control path. It is given by (3.19) with $\lambda(t) = \tilde{\lambda}(t)$ from (3.21) and $x(t) = \tilde{x}(t)$ from (3.22).

Is this candidate actually an optimal solution? The answer is yes because the maximized Hamiltonian function H^* is linear (and hence concave) with respect to the state variable and the scrap value function is identically equal to 0 (and hence concave) so that theorem 3.2 applies.

Now let us see how we can solve the same optimal control problem by using the HJB equation instead of Pontryagin's maximum principle. To apply theorem 3.1 we have to find a continuously differentiable function $V(x, t)$ satisfying conditions (3.7) and (3.8). The theorem also tells us that, if such a function exists, it has to be the optimal value function. Since we already know the optimal solution of our problem we could simply compute the optimal value function by evaluating the objective functional along the optimal control path from (3.19). We could then verify that this function indeed satisfies the conditions stated in theorem 3.1. But what could we do if we had not already solved the problem by theorem 3.2?

The HJB equation (3.7) for the problem at hand is

$$rV(x,t) - V_t(x,t) = \max\left\{-x - \frac{\alpha}{2}u^2 + V_x(x,t)[\beta(t) - u\sqrt{x}] \,\middle|\, u \geq 0\right\}.$$

Carrying out the maximization on the right-hand side and assuming that it leads to an interior maximum[13] yields

$$u = \phi(x,t) := -V_x(x,t)\sqrt{x}/\alpha. \tag{3.23}$$

The similarity of this formula with (3.19) should not be surprising, given the interpretation of the adjoint variable as a shadow price of the state variable. By substituting (3.23) into the HJB equation stated above we obtain the partial differential equation

$$rV(x,t) - V_t(x,t) = -x + \frac{1}{2\alpha}xV_x(x,t)^2 + \beta(t)V_x(x,t). \tag{3.24}$$

Finding a closed form solution of such an equation is a difficult task and requires a good deal of experience and mathematical creativity. There are a few systematic approaches, but it is more common to try to find solutions which are separable in some way (like $V(x,t) = v(x) + w(t)$ or $V(x,t) = v(x)w(t)$) or which have simple functional forms. Let us take the latter route and suppose that $V(x,t)$ is a polynomial of degree k with respect to the state variable x. If this assumption is correct then the left-hand side of (3.24) is a polynomial of degree k in x. The right-hand side, however, is a polynomial of degree $2k - 1$ in x. The polynomial solution can therefore be valid only if $k = 2k - 1$, i.e., if $k = 1$. In other words, the hypothesized optimal value function must be linear with respect to x:

[13] In assuming an interior solution we are using again the intuition mentioned in the previous footnote.

$V(x, t) = A(t)x + B(t)$. Substituting this into the boundary condition (3.8) we obtain the condition $A(T)x + B(T) = 0$ for all $x \geq 0$ (recall that in the present problem $S(x) = 0$ for all x). This condition can be satisfied if and only if

$$A(T) = B(T) = 0. \tag{3.25}$$

Now let us substitute the conjectured optimal value function $V(x, t) = A(t)x + B(t)$ into (3.24). After some rearranging this yields

$$x\left[1 + rA(t) - \frac{1}{2\alpha}A(t)^2 - \dot{A}(t)\right] + rB(t) - \beta(t)A(t) - \dot{B}(t) = 0.$$

A necessary and sufficient condition for this equation to hold for all $x \geq 0$ and all $t \in [0, T]$ is that the two differential equations

$$\dot{A}(t) = 1 + rA(t) - \frac{1}{2\alpha}A(t)^2$$

and

$$\dot{B}(t) = rB(t) - \beta(t)A(t)$$

are satisfied. The equation for $A(t)$ together with the boundary condition $A(T) = 0$ from (3.25) coincides with the boundary value problem (3.20).[14] We already know that the unique solution is $A(t) = \tilde{\lambda}(t)$ with $\tilde{\lambda}(t)$ from (3.21). The differential equation for $B(t)$, with $A(t) = \tilde{\lambda}(t)$ and the boundary condition $B(T) = 0$ from (3.25), is a nonautonomous linear equation which has the unique solution $B(t) = \int_t^T e^{-r(s-t)}\beta(s)\tilde{\lambda}(s)\, ds$. Our polynomial assumption has therefore led to the optimal value function

$$V(x, t) = \tilde{\lambda}(t)x + \int_t^T e^{-r(s-t)}\beta(s)\tilde{\lambda}(s)\, ds \tag{3.26}$$

with $\tilde{\lambda}(t)$ defined in (3.21). Since $A(t) = \tilde{\lambda}(t) \leq 0$ for all $t \in [0, T]$, the maximum of the right-hand side of the HJB equation is indeed given by (3.23), so all conditions of theorem 3.1 are satisfied.

We have proved that if the decision maker determines $u(t)$ according to the law

$$u(t) = -\tilde{\lambda}(t)\sqrt{x}/\alpha, \tag{3.27}$$

where x denotes the observed state of the system at time t, then this yields an optimal control path. Note that the decision maker does not need to know x in advance but he can observe it as the system evolves. This is in contrast to the formula $u(t) = -\tilde{\lambda}(t)\sqrt{\tilde{x}(t)}/\alpha$ that we derived before (using

[14]This should not come as a surprise since $A(t) = V_x(x(t), t) = \lambda(t)$.

the maximum principle). In that formula $\tilde{x}(t)$ is meant to be a value that is computed in advance from (3.22). We shall elaborate on this subtle but important difference in the following section.

3.5 Information, commitment, and strategies

Imagine that the manager of a company hires you to develop a maintenance strategy for the equipment used by the company. After a detailed analysis of the properties of the equipment and the production processes in which it is involved, you come up with a mathematical model which describes the process of deterioration and the costs incurred by maintenance activities and deteriorated equipment. For simplicity, assume that this model is given by the optimal control problem of example 3.3, in which $u(t)$ denotes the intensity of maintenance activities at time t and $x(t)$ measures the state of deterioration of the machinery at time t.[15] Furthermore, assume that the parameters α, r, x_0, and T as well as the function $\beta(\cdot)$ are known to you. The manager seems to be convinced by your arguments for this model and asks you to determine the optimal time profile of maintenance activities, i.e., the form of the optimal control path $u(\cdot)$. Since you have followed the book, at least up to this point, you know the solution of example 3.3 and you are eager to explain to the manager what he should do. But how to describe the optimal solution? Is there a unique way of doing it and, if not, what is the 'best' way?

One possibility would be to write a computer program which computes the functions $\tilde{\lambda}(\cdot)$ and $\tilde{x}(\cdot)$ from equations (3.21) and (3.22), respectively, then asks the user to input a number t from the interval $[0, T]$, and finally outputs the number $u(t) = -\tilde{\lambda}(t)\sqrt{\tilde{x}(t)}/\alpha$ (see equation (3.19)). Then you tell the manager that he should input the current time t and choose the intensity of maintenance activities according to the computer output $u(t)$. In other words, you give the manager a rule ϕ (your program) which allows him to compute the optimal maintenance intensity at a certain time t, knowing nothing but the time t. Formally, we describe this by the equation $u(t) = \phi(t)$, which is called an open-loop representation of the optimal control path, or simply an open-loop strategy. Let us emphasize once more that in order to determine $u(t)$ the manager does not need to observe the current state $x(t)$, nor does he need to know past states $x(\tau)$, $\tau < t$. Nor do you need to observe the state variable when you write the program, except for the initial state x_0 which occurs in equation (3.22) and is therefore a parameter of the program.

[15]Other models of production and maintenance processes can be found in the optimal control literature mentioned in section 3.8.

A second possibility would be to write another computer program which only computes $\tilde{\lambda}(\cdot)$ from equation (3.21), then asks the user to input two numbers $t \in [0, T]$ and $x \in [0, \infty)$, and finally outputs the number $u(t) = -\tilde{\lambda}(t)\sqrt{x}/\alpha$ (see equation (3.27)). You tell the manager that he should input the current time t and the current state of deterioration of the machinery x and choose the maintenance rate corresponding to the computer output $u(t)$. If you follow this approach you do not have to know the initial state x_0 in order to write the computer program because x_0 does not occur in the formula for $\tilde{\lambda}(t)$. On the other hand, the manager has to know at each time t the current state $x(t)$. This representation of the optimal control path is of the form $u(t) = \phi(x(t), t)$ and is called a Markovian strategy, a closed-loop strategy, or a feedback strategy. Unfortunately, the terminology is not consistent and some authors make distinctions between the three terms. We shall use the terminology 'Markovian strategy' throughout the book. Of course, an open-loop strategy is a degenerate form of a Markovian strategy in which the function ϕ happens to be independent of the state x. A Markovian strategy which is not an open-loop strategy will be called a nondegenerate Markovian strategy.

Before we proceed to discuss other possible representations of the solution, you should pause for a moment and make sure that you have understood that the two strategies mentioned above generate exactly the same state trajectory and the same maintenance time path provided the model is correct. Whichever of the two programs you write, if the model is correct and the manager follows your instructions he will end up choosing the same time path for the maintenance rate, $u(\cdot)$. On the other hand, if the model does not describe the actual process of deterioration correctly, a manager who uses an open-loop strategy will, in general, implement a different maintenance rate than a manager employing a nondegenerate Markovian strategy. When you have convinced yourself of this fact you should also try to understand the differences between the two representations.

One important point to note is that different information is required for the implementation of the two strategies. If you advise the manager to use the open-loop strategy $\phi(t)$, the manager needs only his clock and your computer program in order to calculate the optimal maintenance rate. On the other hand, to use the Markovian strategy $\phi(x, t)$, someone has to tell the manager the current state $x(t)$ at any time t. The Markovian strategy is therefore informationally more demanding for the manager.

A second important issue is that of commitment. The open-loop strategy does not give the manager any flexibility to react to signals from the production process: the program outputs the number $u(t)$ and that is the rate of maintenance the manager has to implement. In the case of the nondegenerate Markovian strategy, the commitment is not so strong because the manager can react to different states of deterioration with different rates of maintenance activities.[16] If, for some reason, the state of deterioration of the equipment were different from the value $\tilde{x}(t)$ then the program would tell the manager to choose a maintenance rate which differs from the one computed in example 3.3. If the model is correct this will not happen. But most likely the model is only an approximation of the real world and the true value of $x(t)$ might be different from the predicted value $\tilde{x}(t)$. A non-degenerate Markovian strategy gives the manager the possibility to react to such a difference.

Let us discuss some other types of strategies and the corresponding informational requirements. One possibility would be to write a computer program which computes $\tilde{\lambda}(\cdot)$ and $\tilde{x}(\cdot)$ (as in the open-loop case), then asks the manager to input the current time t and state x, and finally outputs the number $u(t) = -\tilde{\lambda}(t)\sqrt{\gamma x + (1-\gamma)\tilde{x}(t)}/\alpha$, where γ is an arbitrary real constant. It is easy to see that this yields exactly the same path of maintenance activities. The information necessary to implement this policy consists of time t and the current state $x(t)$ for every time $t \in [0, T]$. Thus, it is again a Markovian strategy of the form $u(t) = \phi(x(t), t)$. In contrast to the Markovian strategy discussed before, this one depends on the initial state x_0, which you need to know when you write the computer program. Since γ can be any real constant we obtain infinitely many different representations of the same (unique) solution of the optimal control problem, i.e., infinitely many optimal Markovian strategies. In fact, γ need not be a constant, it may be a time-function built into the program, or it may be a time-function which has to be provided by the manager as an input to the program. In the latter case we would write $u(t) = \phi(x(t), t, \gamma(t))$. For example, the program may ask the manager to input the current time t, the current state $x(t)$, and his current blood pressure $\gamma(t)$ at each time t. This shows that among the many different representations of the same solution there are some which are based on seemingly irrelevant data like blood pressure (or the outside temperature in Copenhagen, or the current rate of exchange between the Canadian dollar and the Euro).

Not only can information be irrelevant, it can also be inaccessible. This is the case, for example, if γ is not the manager's own blood pressure but someone else's who refuses to provide this information. A very important

[16]Now he commits to the feedback rule, that is, to the Markovian strategy.

case of inaccessible information can be illustrated by the following example. Consider equation (3.22) and note that it can also be written as

$$x(t) = \tilde{x}(t) = e^{-\int_t^T \tilde{\lambda}(\tau)/\alpha \, d\tau} x(T) - \int_t^T \beta(s) e^{-\int_t^s \tilde{\lambda}(\tau)/\alpha \, d\tau} \, ds.$$

In contrast to (3.22), the initial state x_0 does not appear in this equation but the final state $x(T)$ does. Using this formula for $\tilde{x}(t)$, we obtain a strategy of the form $u(t) = \phi(t, x(T))$ which would require that someone tells you or the manager the state variable at the end of the planning horizon, $x(T)$. Not only is $x(T)$ something which is unknown at any time $t < T$ but it is something which is influenced by the manager's decisions during the interval $[t, T]$. In other words, requiring the knowledge of $x(T)$ at time $t < T$ violates a basic principle of causality: you cannot condition your decisions on the effects of these decisions.

From the above discussion it should be clear that, although many representations of the same solution exist, not all of them make sense and yield implementable strategies. A sensible strategy can only involve endogenous variables from the past up to the present, or exogenous variables which are known in advance (like the parameter r, the time variable t, or the function β). If you denote the set of all endogenous variables at time t by $\eta(t)$, then a valid strategy must be of the form $u(t) = \phi(\{\eta(\tau) \mid \tau \leq t\}, t)$. Obviously, it depends on the particular model and on the informational assumptions which variables actually occur as arguments of the strategy ϕ. In some situations it might be reasonable to assume that the current state variable can be observed so that a nondegenerate Markovian strategy is feasible. In other settings the state variable could be unobservable, or only observable after a certain time lag τ, so that open-loop strategies or strategies of the form $\phi(x(t - \tau), t)$ are more realistic than a nondegenerate Markovian strategy. We shall see in later chapters that different informational assumptions (and hence different strategies) usually lead to completely different solutions as soon as we leave (single decision maker) control theory and enter into the field of differential games.

3.6 Infinite time horizon

In many situations, the end of the planning period, T, is either very far in the future or unknown. For example, a firm which wants to maximize the present value of its lifetime profits probably does not know the time at which it will go out of business. In this case, setting $T = \infty$ may very well be the best approximation for the true problem horizon. But even if the firm's management restricts itself to maximizing the discounted profits over the next year it should value its asset position at the end of the year by the optimal value that can be earned starting from that asset position and continuing to be in existence in the years to come. In this situation, too, it is clear that the appropriate problem horizon of the firm is $T = \infty$.

The present section modifies the results that we have proved so far to make them applicable to optimal control problems over the unbounded planning period $[0, \infty)$. Of course, it does not make sense to have a terminal value function in this case because there is no finite terminal time, and we therefore set $S(x) = 0$ throughout this section.

If $T = \infty$ then the integral in (3.4) does not necessarily converge. The integral does converge if, for example, the utility function F is bounded and continuous and $r > 0$. In many examples, however, boundedness of the utility function is too strong an assumption. If the integral does not converge, our definition of an optimal path is not very useful. For example, if $F(x, u, t) = 1 + u$, $r = 0$, and $U(x, t) = [0, \infty)$ then every feasible control path $u(\cdot)$ has infinite value, $J(u(\cdot)) = \infty$, and could be called an optimal path. On the other hand, we would intuitively prefer a constant control path defined by $u(t) = 2$ to another constant control path defined by $u(t) = 1$. To handle this and similar situations in a rigorous way, several definitions of optimality for problems with unbounded time domain have been proposed in the literature. We list a few of them in the following definition.

Definition 3.2 Consider an optimal control problem in which the objective functional is given by

$$J(u(\cdot)) = \int_0^\infty e^{-rt} F(x(t), u(t), t)\, dt. \tag{3.28}$$

The T-truncation[17] of the objective functional, $J_T(u(\cdot))$, is defined by

$$J_T(u(\cdot)) = \int_0^T e^{-rt} F(x(t), u(t), t)\, dt.$$

A feasible control path $u(\cdot)$ is called

- overtaking optimal if for every feasible control path $\tilde{u}(\cdot)$ there exists a finite number τ such that $J_T(u(\cdot)) - J_T(\tilde{u}(\cdot)) \geq 0$ holds for all $T \in [\tau, \infty)$,

[17]The capital T here indicates terminal time; cf. t-truncation in section 6.1 and elsewhere, where t refers to an intermediate point in time. The meaning of both notations is, however, the same.

- catching up optimal if $\liminf_{T\to\infty}[J_T(u(\cdot)) - J_T(\tilde{u}(\cdot))] \geq 0$,
- sporadically catching up optimal if $\limsup_{T\to\infty}[J_T(u(\cdot)) - J_T(\tilde{u}(\cdot))] \geq 0$.

It is obvious that in the case where the objective functional converges for all feasible paths, all of these definitions of optimal control paths collapse to the one from definition 3.1. It is also easy to see that the definitions are ranked in order of decreasing strength, that is, any overtaking optimal path is catching up optimal and any catching up optimal path is also sporadically catching up optimal. The reverse implications do not hold in general. It can happen that an optimal control problem has a solution with respect to a weaker criterion (say, the catching up criterion) but has no solution with respect to a stronger criterion (say, the overtaking criterion).

Now let us see what has to be modified in our theorems in order to cope with the new situation of an unbounded time interval. First of all, the boundary condition (3.8) does not make any sense any more because $T = \infty$. Neither does the transversality condition (3.15) make sense since it was derived from (3.8). One might be tempted to replace these conditions by $\lim_{T\to\infty} V(x, T) = 0$ and $\lim_{T\to\infty} \lambda(T) = 0$, respectively, because the scrap value function has formally been set equal to 0. However, it turns out that this requirement is quite often too strong as a sufficient optimality condition.

A much better alternative is easily discovered if one looks carefully at the proofs of theorems 3.1 and 3.2, respectively. As a matter of fact, the very same arguments that were used in the proofs of these theorems show that

$$J_T(u(\cdot)) - J_T(\tilde{u}(\cdot)) \geq e^{-rT}[V(\tilde{x}(T), T) - V(x(T), T)]$$

and

$$J_T(u(\cdot)) - J_T(\tilde{u}(\cdot)) \geq e^{-rT}\lambda(T)[\tilde{x}(T) - x(T)].$$

Here, $\tilde{u}(\cdot)$ is an arbitrary feasible control path and $u(\cdot)$ is a feasible control path which satisfies $u(t) \in \Phi(x(t), t)$ (in case of theorem 3.1) or (3.13) and (3.14) (in case of theorem 3.2). Together with the above definitions this provides immediately the following sufficiency theorem for problems with unbounded time intervals.

Theorem 3.3 *Consider the optimal control problem with the objective functional (3.28) and the constraints (3.1)–(3.3).*

(i) If optimality is understood in terms of the overtaking criterion, then theorem 3.1 remains valid provided the boundary condition (3.8) is replaced by the assumption that for every feasible control path $\tilde{u}(\cdot)$ there exists a finite number τ such that

$$V(\tilde{x}(T), T) - V(x(T), T) \geq 0 \tag{3.29}$$

holds for all $T \in [\tau, \infty)$. Similarly, theorem 3.2 remains valid for the overtaking criterion if the transversality condition (3.15) is replaced by the assumption that for every feasible control path $\tilde{u}(\cdot)$ there exists a finite number τ such that

$$\lambda(T)[\tilde{x}(T) - x(T)] \geq 0 \tag{3.30}$$

holds for all $T \in [\tau, \infty)$.
(ii) Theorems 3.1 and 3.2 remain valid if optimality is understood in the sense of the catching up criterion if condition (3.8) is replaced by

$$\liminf_{T\to\infty} e^{-rT}[V(\tilde{x}(T), T) - V(x(T), T)] \geq 0 \tag{3.31}$$

and condition (3.15) is replaced by

$$\liminf_{T\to\infty} e^{-rT}\lambda(T)[\tilde{x}(T) - x(T)] \geq 0. \tag{3.32}$$

(iii) Theorems 3.1 and 3.2 remain valid if optimality is understood in the sense of the sporadically catching up criterion if condition (3.8) is replaced by

$$\limsup_{T\to\infty} e^{-rT}[V(\tilde{x}(T), T) - V(x(T), T)] \geq 0 \tag{3.33}$$

and condition (3.15) is replaced by

$$\limsup_{T\to\infty} e^{-rT}\lambda(T)[\tilde{x}(T) - x(T)] \geq 0. \tag{3.34}$$

Conditions (3.29)–(3.34) are usually not easy to verify because they have to be true for all feasible control paths $\tilde{u}(\cdot)$. In some situations, however, they are trivially satisfied. This is the case for conditions (3.31) and (3.33), for example, if the function V is a bounded function and $r > 0$. Analogously, conditions (3.32) and (3.34) are automatically true if all feasible state trajectories remain bounded (for example, because X is bounded) and $\lim_{T\to\infty} e^{-rT}\lambda(T) = 0$. A somewhat less obvious result is stated in the following lemma.

Lemma 3.1 *(i) Assume that V is bounded from below and that $r > 0$. Then condition (3.31) is implied by*

$$\limsup_{T\to\infty} e^{-rT} V(x(T), T) \le 0,$$

and condition (3.33) is implied by

$$\liminf_{T\to\infty} e^{-rT} V(x(T), T) \le 0.$$

(ii) Suppose there exists $a \in \mathbb{R}^n$ *such that* $x \ge a$ *holds for all* $x \in X$ *and that* $\lambda(T) \ge 0$ *holds for all sufficiently large* T. *Then (3.32) is implied by*

$$\limsup_{T\to\infty} e^{-rT}\lambda(T)[x(T) - a] \le 0,$$

and (3.34) is implied by

$$\liminf_{T\to\infty} e^{-rT}\lambda(T)[x(T) - a] \le 0.$$

(iii) If $\lim_{T\to\infty} e^{-rT}\lambda(T)\tilde{x}(T) = 0$ *holds for all feasible state trajectories* $\tilde{x}(\cdot)$, *then conditions (3.32) and (3.34) are satisfied.*

Proof (i) Assuming $V(x, t) \ge \alpha > -\infty$ for all $(x, t) \in X \times [0, \infty)$ we have

$$e^{-rT}[V(\tilde{x}(T), T) - V(x(T), T)] \ge \alpha e^{-rT} - e^{-rT} V(x(T), T).$$

Thus, $r > 0$ and $\limsup_{T\to\infty} e^{-rT} V(x(T), T) \le 0$ imply condition (3.31), and $r > 0$ and $\liminf_{T\to\infty} e^{-rT} V(x(T), T) \le 0$ imply condition (3.33).

(ii) For all sufficiently large T it holds that

$$e^{-rT}\lambda(T)[\tilde{x}(T) - x(T)] = e^{-rT}\lambda(T)\{\tilde{x}(T) - a - [x(T) - a]\} \ge -e^{-rT}$$
$$\lambda(T)[x(T) - a].$$

Thus, $\limsup_{T\to\infty} e^{-rT}\lambda(T)[x(T) - a] \le 0$ implies that condition (3.32) holds and $\liminf_{T\to\infty} e^{-rT}\lambda(T)[x(T) - a] \le 0$ implies condition (3.34).

(iii) The proof of this statement is obvious. ■

Conditions (i) and (ii) in this lemma are much easier to check than the conditions in theorem 3.3 because they involve only the given candidate solution, $u(\cdot)$ and $x(\cdot)$, instead of all feasible solutions $\tilde{u}(\cdot)$ and $\tilde{x}(\cdot)$. Condition (iii) does not require the state space X to be bounded from below (like condition (ii)) but has to hold for all feasible state trajectories.

Now consider an optimal control problem for which $X = \mathbb{R}^n$, V is not bounded below, and there exist feasible state trajectories $\tilde{x}(\cdot)$ growing at an arbitrarily large rate. Obviously, conditions (i) and (ii) of lemma 3.1 are not satisfied. But neither can part (iii) of the lemma or any of the conditions of theorem 3.3 be applied because it cannot be ruled out that $\lim_{T\to\infty} e^{-rT}\lambda(T)\tilde{x}(T) = -\infty$ or $\lim_{T\to\infty} e^{-rT} V(\tilde{x}(T), T) = -\infty$. This situation arises, for example, in linear quadratic optimal control problems with an unconstrained state space which are frequently used in applications.[18] Fortunately, these problems can often be solved by an approach which is based on finite horizon approximations. The following theorem provides the details.

Theorem 3.4 *Consider the optimal control problem with constraints (3.1)–(3.3) and objective functional (3.28). We make the following assumptions.*

(i) For all sufficiently large numbers $T > 0$ there exists a continuously differentiable function $V(\cdot, \cdot; T) : X \times [0, T] \mapsto \mathbb{R}$ which solves the HJB equation

$$rV(x, t; T) - V_t(x, t; T) = \max\{F(x, u, t) + V_x(x, t; T)f(x, u, t) \mid u \in U(x, t)\}$$

and the terminal condition

$$V(x, T; T) = 0.$$

Denote by $\Phi(x, t; T)$ the set of controls $u \in U(x, t)$ maximizing the right-hand side of this HJB equation.

(ii) For all $(x, t) \in X \times [0, \infty)$ the limit $V(x, t) := \lim_{T\to\infty} V(x, t; T)$ exists, is finite, and is a continuously differentiable function $V : X \times [0, \infty) \mapsto \mathbb{R}$ which solves the HJB equation, that is,

$$rV(x, t) - V_t(x, t) = \max\{F(x, u, t) + V_x(x, t)f(x, u, t) \mid u \in U(x, t)\}.$$

Denote the set of controls $u \in U(x, t)$ maximizing the right-hand side of this equation by $\Phi(x, t)$.

(iii) For all sufficiently large T there exists a control path $u_T(\cdot)$ and a corresponding state trajectory $x_T(\cdot)$ satisfying the constraints (3.1)–(3.3) and the condition $u_T(t) \in \Phi(x_T(t), t; T)$ for all $t \in [0, T]$. Analogously, there exists a control path $u(\cdot)$ and a corresponding state trajectory $x(\cdot)$ satisfying the constraints (3.1)–(3.3) and the condition $u(t) \in \Phi(x(t), t)$ for all $t \in [0, \infty)$.

(iv) It holds that $\limsup_{T\to\infty} e^{-rT} V(x(T), T) \leq 0$.

Then the control path $u(\cdot)$ is catching up optimal.

[18] We shall encounter exactly this problem in our discussion of linear quadratic differential games in section 7.1 below.

Proof Using the definition of J_T and the fact that V solves the HJB equation we obtain

$$\begin{aligned} J_T(u(\cdot)) &= \int_0^T e^{-rt}F(x(t), u(t), t)\,dt \\ &= \int_0^T e^{-rt}[rV(x(t), t) - V_t(x(t), t) - V_x(x(t), t)\dot{x}(t)]\,dt \\ &= \int_0^T \frac{d}{dt}\left[-e^{-rt}V(x(t), t)\right] dt \\ &= V(x_0, 0) - e^{-rT}V(x(T), T). \end{aligned}$$

Applying an analogous argument and noting that $V(x, T; T) = 0$ holds for all $x \in X$ we obtain $J_T(u_T(\cdot)) = V(x_0, 0; T)$. Therefore it holds that

$$\begin{aligned} &\liminf_{T\to\infty}[J_T(u(\cdot)) - J_T(u_T(\cdot))] = \liminf_{T\to\infty}[V(x_0, 0) - \\ &V(x_0, 0; T) - e^{-rT}V(x(T), T)] \geq 0. \end{aligned} \tag{3.35}$$

The last inequality in (3.35) follows from the convergence of $V(x_0, 0; T)$ to $V(x_0, 0)$ and from assumption (iv) in the theorem. Finally, note that the conditions of the theorem together with theorem 3.1 imply that $u_T(\cdot)$ is an optimal path for the finite horizon problem, that is, $J_T(u_T(\cdot)) \geq J_T(\tilde{u}(\cdot))$ holds for all feasible control paths $\tilde{u}(\cdot)$ and for all sufficiently large T. Together with (3.35) this implies $\liminf_{T\to\infty}[J_T(u(\cdot)) - J_T(\tilde{u}(\cdot))] \geq 0$ for all feasible paths $\tilde{u}(\cdot)$, which proves the theorem. ■

We conclude this section with two examples. The first is a classical example from optimal growth theory originally set up by Frank P. Ramsey in 1928. This example can also be interpreted as a renewable resource model describing the same situation as example 3.1 in the special case where there is only a single firm in the market. The second example is a very simple linear quadratic optimal control problem with a single state variable. These examples are used to illustrate the verification of the transversality conditions and the boundary conditions for the optimal value function.

Example 3.4 Consider an infinite-horizon economy producing output from a single capital stock. Output can be either consumed or invested. Denoting by $x(t)$ the capital stock at time t, by $u(t)$ the consumption rate at time t, and by g the net production function (i.e., output minus depreciated capital), we obtain the system dynamics

$$\dot{x}(t) = g(x(t)) - u(t), \ x(0) = x_0. \tag{3.36}$$

This is an accounting identity stating that net investment, $\dot{x}(t)$, must be equal to net production minus consumption. It is assumed that the production function $g : [0, \infty) \mapsto \mathbb{R}$ is continuous, twice continuously differentiable on $(0, \infty)$, and strictly concave. In addition, we assume that $g(0) = 0$, $\lim_{x \to 0} g'(0) = \infty$, and that there exists a unique capital stock $\bar{x} > 0$ such that $g(\bar{x}) = 0$. This implies that $g(x) > 0$ for all $x \in (0, \bar{x})$, and $g(x) < 0$ for all $x > \bar{x}$.

The goal of the decision maker is to maximize the discounted utility derived over the infinite planning interval $[0, \infty)$. That is, the objective functional is

$$\int_0^\infty e^{-rt} F(u(t))\, dt, \tag{3.37}$$

where $F : [0, \infty) \mapsto \mathbb{R}$ is the utility function. Although the analysis can be carried out for a very general class of utility functions, we restrict ourselves to functions of the form

$$F_\beta(u) = \begin{cases} \dfrac{u^\beta - 1}{\beta} & \text{if} \beta \in (0, 1), \\ \ln u & \text{if} \beta = 0. \end{cases}$$

These utility functions have a constant elasticity of intertemporal substitution $1/(1 - \beta)$ and are very often used in studies of economic growth. The optimal control problem consists in maximizing (3.37) subject to the system dynamics (3.36) and the nonnegativity constraints $x(t) \geq 0$ and $u(t) \geq 0$.

The Hamiltonian function for this problem is $H(x, u, \lambda, t) = F_\beta(u) + \lambda[g(x) - u]$. A necessary and sufficient condition that $u(t)$ is an interior maximum of $H(x(t), u, \lambda(t), t)$ is

$$F'_\beta(u(t)) = u(t)^{\beta - 1} = \lambda(t), \tag{3.38}$$

which shows (i) that the costate $\lambda(t)$ has to be positive if there is an interior maximum of the Hamiltonian and (ii) that the maximizing value of u does not depend on the state $x(t)$ and can be written as $u(t) = \lambda(t)^{1/(\beta-1)}$. Substituting this into the Hamiltonian function we get $H^*(x, \lambda, t) = F_\beta(\lambda^{1/(\beta-1)}) - \lambda^{\beta/(\beta-1)} + \lambda g(x)$. It follows from our assumptions that $H^*(x, \lambda(t), t)$ is strictly concave with respect to x whenever $\lambda(t) > 0$. The adjoint equation becomes

$$\dot{\lambda}(t) = \lambda(t)[r - g'(x(t))]. \tag{3.39}$$

Differentiating (3.38) with respect to t and using (3.39) we obtain

$$\dot{u}(t) = u(t)[g'(x(t)) - r]/(1 - \beta). \tag{3.40}$$

The system consisting of the two differential equations (3.36) and (3.40) is the main tool for the analysis of optimal solutions to the model. Note that we have only one initial condition for this two-dimensional system, namely, $x(0) = x_0$. This implies that there exist infinitely many solutions of the system (3.36), (3.40). This situation is typical for infinite-horizon optimal control problems and the following arguments may help to solve many similar problems arising in various applications.

The key observation is that not all solutions of (3.36), (3.40) satisfy a transversality condition. The transversality conditions listed in theorem 3.3 usually help to reduce the number of candidates for optimality to a small number, quite often to a unique candidate. To see how this works, we first draw a phase diagram of the system (3.36), (3.40) (see figure 3.1). The relevant phase space is the set $\{(x, u) \mid x \geq 0, u \geq 0\}$. The solid concave curve is $u = g(x)$, which is the locus of all points at which the right-hand side of (3.36) becomes 0. It is called the $(\dot{x} = 0)$-isocline. The vertical line is the locus $x = \hat{x}$, where $\hat{x}$ is the unique capital stock satisfying $g'(\hat{x}) = r$. This line together with the horizontal coordinate axis is the set of all points where the right-hand side of (3.40) becomes 0. We call this set the $(\dot{u} = 0)$-isocline.

The two isoclines divide the phase space into four regions each of which is characterized by a unique direction of the flow determined by (3.36) and (3.40). These directions are indicated in the figure by arrows. The three points of intersection of the two isoclines are the origin $(0, 0)$, the point $(\hat{x}, g(\hat{x}))$, and the point $(\bar{x}, 0)$. These are the only three fixed points of system (3.36), (3.40). The non-trivial fixed point $(\hat{x}, g(\hat{x}))$ is a saddle point, as can be seen by computing the Jacobian matrix, J, and its eigenvalues.

More specifically, we have

$$J = \begin{pmatrix} r & -1 \\ A & 0 \end{pmatrix},$$

where $A = g(\hat{x})g''(\hat{x})/(1 - \beta) < 0$. Consequently, the eigenvalues are $\sigma_1 = r/2 - \sqrt{(r/2)^2 - A}$ and $\sigma_2 = r/2 + \sqrt{(r/2)^2 - A}$. Because $A < 0$ we know that σ_1 is negative and σ_2 is positive, which proves that $(\hat{x}, g(\hat{x}))$ is indeed a saddle point. It follows that for every initial state $x_0 \in (0, \infty)$ there exists a unique solution of the system (3.36), (3.40) which converges to the saddle point.[19] Along this solution we have $u(t) > 0$ and

[19]The stable saddle point path is depicted in figure 3.1 as a dotted line.

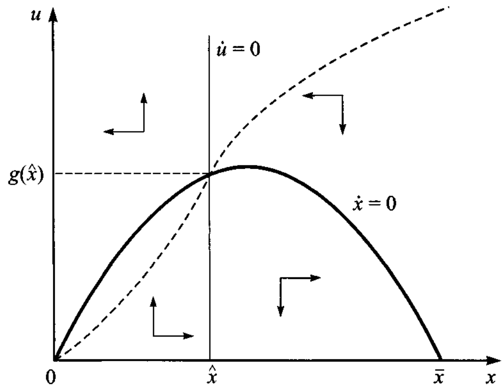

Figure 3.1. Phase diagram of system (3.36), (3.40)

$x(t) > 0$, and the corresponding costate trajectory, which can be derived from (3.38), is also strictly positive. By construction, the maximum condition and the adjoint equation are satisfied for this triple $(x(\cdot), u(\cdot), \lambda(\cdot))$. It remains to verify a transversality condition. To this end we invoke lemma 3.4(ii). The state space is the nonnegative real line, the costate trajectory is nonnegative for all $t \geq 0$, and both $x(t)$ and $\lambda(t)$ converge to finite limits $\hat{x}$ and $g(\hat{x})^{\beta-1}$, respectively. Therefore, we have $\lim_{t\to\infty} e^{-rt}\lambda(t)x(t) = 0$ and lemma 3.1 shows that the transversality condition (3.32) is satisfied. This proves that the stable branch of the saddle point describes a catching up optimal solution to the problem.

Example 3.5 Consider the optimal control problem defined by $X = \mathbb{R}$, $U(x, t) = \mathbb{R}$, $f(x, u, t) = \alpha x + u$, and $F(x, u, t) = -u^2/2 - \beta x^2/2$, where α and β are real constants. The nonnegative discount rate is denoted by r. We shall see shortly that, in the case where β is positive, the optimal value function for this optimal control problem is a strictly concave quadratic function such that $\lim_{|x|\to\infty} V(x, t) = -\infty$. Furthermore, there are neither state nor control constraints so that feasible state trajectories can diverge to ∞ or $-\infty$ arbitrarily fast. Thus, we are in the situation described in the paragraph preceding theorem 3.4 and we have to apply the finite horizon approximation approach described in that theorem. For the sake of completeness, we shall also discuss the case where $\beta \leq 0$.

The HJB equation is given by

$$\begin{aligned} rV(x,t) - V_t(x,t) &= \max\{-u^2/2 - \beta x^2/2 + V_x(x,t)(\alpha x + u) \mid u \in \mathbb{R}\} \\ &= V_x(x,t)^2/2 - \beta x^2/2 + \alpha x V_x(x,t) \end{aligned}$$

and the right-hand side is maximized at $u = V_x(x,t)$. We conjecture that the optimal value function for the problem with finite time horizon T is of the form $V(x,t;T) = A(t;T)x^2$. Substituting this into the HJB equation we see that $A(\cdot;T)$ must satisfy the Riccati differential equation[20]

$$\begin{aligned} \dot{A}(t;T) &= \beta/2 + (r-2\alpha)A(t;T) - 2[A(t;T)]^2 \\ &= 2[A(t;T) - A_1][A_2 - A(t;T)], \end{aligned} \tag{3.41}$$

where

$$A_1 = \frac{r - 2\alpha - \sqrt{(r-2\alpha)^2 + 4\beta}}{4}, \quad A_2 = \frac{r - 2\alpha + \sqrt{(r-2\alpha)^2 + 4\beta}}{4}.$$

Note that A_1 and A_2 are fixed points of (3.41). The boundary condition for the finite horizon problem requires $A(T;T) = 0$. Although equation (3.41) can be solved explicitly, we use graphical analysis to derive the solution.

Suppose first that $\beta > 0$. In this case the utility function is strictly concave and the fixed points of (3.41) satisfy $A_1 < 0 < A_2$. The corresponding $(A, \dot{A})$-diagram is depicted in figure 3.2(a). The arrows on the horizontal axis indicate the direction of the flow determined by (3.41). If $A_1 < A(t;T) < A_2$ then $\dot{A}(t;T)$ is positive so that $A(t;T)$ must be increasing. If $A(t;T) \in (-\infty, A_1,) \cup (A_2, \infty)$ then $\dot{A}(t;T)$ is negative and $A(t;T)$ must be decreasing. The unique solution of (3.41) that satisfies the boundary condition $A(T;T) = 0$ must therefore also satisfy $A(t;T) \in (A_1, 0)$ for all $t \in [0, T]$. This solution determines an optimal Markovian strategy for the finite horizon problem, namely $u_T(t) = \phi(x(t), t; T)$ with $\phi(x,t;T) = V_x(x,t;T) = 2A(t;T)x$. Moreover, it is clear that $\lim_{T\to\infty} A(t;T) = A_1$ holds for all $t \in [0,\infty)$ so that $V(x,t) = A_1 x^2$ is the limit of the optimal value functions $V(x,t;T)$. Since A_1 is a fixed point (and hence a solution) of (3.41), the limit $V(x,t)$ solves the HJB equation. The corresponding control path is defined by the Markovian strategy $u(t) = \phi(x(t))$ with $\phi(x) = 2A_1 x$. It remains to verify condition (iv) of theorem 3.4. To this end note that application of the strategy ϕ

[20]We denote by $\dot{A}(t;T)$ the derivative of $A(t;T)$ with respect to the first argument.

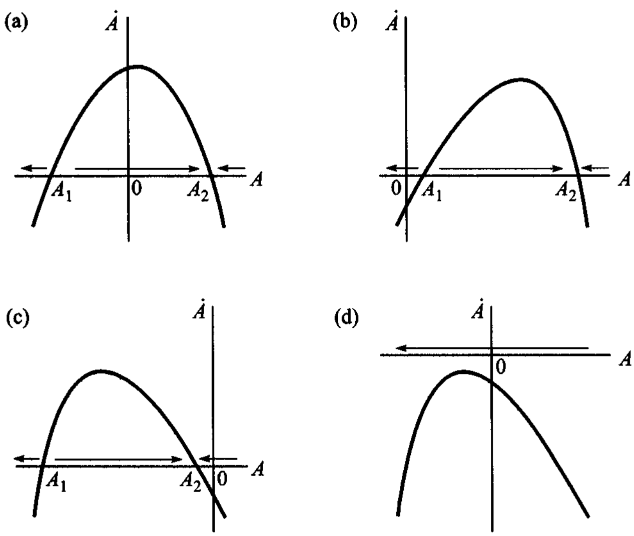

Figure 3.2. Analysis of the $(A, \dot{A})$-diagram

yields $\dot{x}(t) = \alpha x(t) + \phi(x(t)) = (\alpha + 2A_1)x(t)$, which implies $x(t) = x_0 \exp[(\alpha + 2A_1)t]$. Since

$$\alpha + 2A_1 = \frac{r - \sqrt{(r - 2\alpha)^2 + 4\beta}}{2} < \frac{r}{2}$$

we see that $e^{-rt}V(x(t), t) = A_1 e^{-rt}x(t)^2 = A_1 x_0^2 \exp[(2\alpha + 4A_1 - r)t]$ converges to 0 as t approaches infinity. Therefore all conditions of theorem 3.4 are satisfied and $\phi(x) = 2A_1 x$ is an optimal Markovian strategy. It should be noted that the optimal state trajectory need not be bounded. In fact, $x(\cdot)$ remains bounded if and only if $\alpha + 2A_1 \leq 0$, which is equivalent to $\beta \geq \alpha(r - \alpha)$.

Now suppose that $-(r - 2\alpha)^2/4 \leq \beta \leq 0$ and $r \geq 2\alpha$. In this case it holds that $0 \leq A_1 < A_2$. The corresponding $(A, \dot{A})$-diagram is depicted in figure 3.2(b). By the same argument as above we see that the unique solution $A(t; T)$ satisfying the terminal condition $A(T; T) = 0$ must be decreasing with $0 \leq A(t; T) \leq A_1$ and $\lim_{T\to\infty} A(t; T) = A_1$ for all $t \in [0, \infty)$. Consequently, $\phi(x) = 2A_1 x$ is an optimal Markovian strategy in this case, too. As before, $x(\cdot)$ remains bounded if and only if $\beta \geq \alpha(r - \alpha)$.

The next case is $-(r - 2\alpha)^2/4 \leq \beta \leq 0$ and $r < 2\alpha$, which is depicted in figure 3.2(c). Both fixed points are nonpositive so that the unique solution of equation (3.41) satisfying $A(T; T) = 0$ must be decreasing and strictly positive. In this case, however, since there are no fixed points of (3.41) on the positive real axis, it holds that $\lim_{T\to\infty} A(t; T) = +\infty$ for all $t \in [0, \infty)$. Because there is no finite limit, theorem 3.4 cannot be applied.[21]

The final case is characterized by $\beta < -(r - 2\alpha)^2/4$. Clearly, A_1 and A_2 are not real and the $(A, \dot{A})$ diagram is as in figure 3.2(d). As in the previous case, the unique solution of (3.41) satisfying $A(T; T) = 0$ is strictly positive and decreasing and it holds that $\lim_{T\to\infty} A(t; T) = +\infty$ for all $t \in [0, \infty)$. The conditions of theorem 3.4 cannot be satisfied.

Let us conclude the discussion of this example by introducing additional state constraints of the form $-\underline{x} \leq x(t) \leq \bar{x}$ into the model where $\underline{x}$ and $\bar{x}$ are sufficiently large positive numbers. Since we write all constraints in the form of (3.3), define $X = [-\underline{x}, \bar{x}]$ and

$$U(x, t) = \begin{cases} [0, \infty) & \text{if } x = -\underline{x}, \\ \mathbb{R} & \text{if } x \in (-\underline{x}, \bar{x}), \\ (-\infty, 0] & \text{if } x = \bar{x}. \end{cases}$$

In those cases where we obtained a bounded optimal state trajectory for the original problem, it remains the optimal solution also in the problem with the additional state constraints. In these cases one can also verify the conditions of theorem 3.3 since the optimal value function V remains bounded on the bounded interval X. The optimal value function for the model with state constraints is the same as the one for the unconstrained problem.

In those cases where we obtained an optimal strategy of the unconstrained problem leading to an unbounded state trajectory, this strategy is no longer optimal in the problem with constraints. This follows simply from the fact that the strategy is not feasible because it does not generate a state trajectory with $x(t) \in X$ for all t. Another way to see that our proof for the unconstrained model does not carry over to the constrained model in this case is to note that the quadratic function V used above does not solve the HJB equation at $x = -\underline{x}$ and $x = \bar{x}$. The true optimal value function for this case is not a quadratic function and will not be derived here.

[21] In fact, there does not exist an overtaking optimal solution either in this case or in the next one.

3.7 Conditions for nonsmooth problems

In theorems 3.1–3.4 we have assumed differentiability of certain functions. In the results using the HJB equation the function V was assumed to be continuously differentiable; in the results using the maximum principle it was the maximized Hamiltonian $H^*(\cdot, \lambda(t), t)$ and the scrap value function S. Example 3.2, on the other hand, clearly showed that the optimal value function is not necessarily smooth even if the utility function F and the system dynamics f are smooth. For the same example one can also show that the maximized Hamiltonian is given by $H^*(x, \lambda, t) = |x + \lambda|$, which is neither continuously differentiable nor concave. Consequently, we cannot solve this simple example by means of the techniques introduced so far. In the present section we state more general (i.e., weaker) sufficient optimality conditions that are able to handle problems similar to example 3.2. To this end we introduce a concept of smoothness which is weaker than the notion of continuous differentiability. It is called local Lipschitz continuity and is defined as follows.

Definition 3.3 Consider any real-valued function $g : \mathbb{R}^k \mapsto \mathbb{R}$. We say that g is Lipschitz continuous on the set $S \subseteq \mathbb{R}^k$ if there exists a constant $K_S > 0$ such that $|g(x) - g(y)| \leq K_S \|x - y\|$ holds for all $x, y \in S$. The function g is said to be locally Lipschitz continuous on S, if for every $x \in S$ there exists an open neighbourhood of x such that g is Lipschitz continuous on this neighbourhood.

It can be shown that a function g, which is defined and locally Lipschitz continuous on an open subset of $\mathbb{R}^k$, is differentiable almost everywhere, that is, on a set whose complement in the domain of g has Lebesgue measure 0. Let us denote by Z_g the set of all points at which the function g fails to be differentiable and by $\nabla g(x)$ the gradient vector of g at any point $x \notin Z_g$.

Definition 3.4 Assume that the function $g : \mathbb{R}^k \mapsto \mathbb{R}$ is Lipschitz continuous in an open neighbourhood of x. The set

$$\partial g(x) = \text{co}\left\{ \lim_{i \to \infty} \nabla g(x_i) \,\middle|\, x_i \to x, x_i \notin Z_g \right\}$$

is called the generalized gradient of the function g at the point x. Here $\text{co}\, A$ denotes the convex hull of a set A, i.e., the set of all convex combinations of elements in A.

If the function is continuously differentiable in a neighbourhood of x then the generalized gradient $\partial g(x)$ coincides with the usual gradient $\nabla g(x)$. To see what happens if g is not differentiable let us consider the function V from example 3.2.

Example 3.2 (continued) The optimal value function $V(x, t) = (T-t)^2/2 + (T-t)|x|$ fails to be differentiable on the set $Z_V = \{(x, t) \mid x = 0, t \in [0, T)\}$. If $(x, t) \notin Z_V$ then

$$\partial V(x, t) = \nabla V(x, t) = (V_x(x, t), V_t(x, t)) = ((T-t)\text{sgn}(x), t - T - |x|).$$

Now consider any sequence $(x_i, t_i)_{i=1}^{\infty}$, such that $(x_i, t_i) \notin Z_V$ and $\lim_{i\to\infty}(x_i, t_i) = (0, t)$, for which the sequence $\nabla V(x_i, t_i)$ converges. It is easy to check that for such a sequence we have either $\lim_{i\to\infty} \nabla V(x_i, t_i) = (T-t, t-T)$ or $\lim_{i\to\infty} \nabla V(x_i, t_i) = (t-T, t-T)$. The convex hull of these two limits is the set $[t-T, T-t] \times \{t-T\}$. We therefore obtain the generalized gradient

$$\partial V(x, t) = \begin{cases} \{((T-t)\text{sgn}(x), t-T-|x|)\} & \text{if } (x, t) \notin Z_V, \\ [t-T, T-t] \times \{t-T\} & \text{if } (x, t) \in Z_V, \end{cases}$$

for the optimal value function of the problem at hand.

Equipped with the notion of generalized gradients we can state the following generalization of theorem 3.1.

Theorem 3.5 *Let $V : X \times [0, T] \mapsto \mathbb{R}$ be a locally Lipschitz continuous function which satisfies equation (3.8) and the generalized HJB equation*

$$rV(x, t) = \max\{F(x, u, t) + \alpha f(x, u, t) + \beta \mid u \in U(x, t), (\alpha, \beta) \in \partial V(x, t)\} \tag{3.42}$$

for all $(x, t) \in X \times [0, T]$. Denote by $\Phi(x, t)$ the set of all $(u, \alpha, \beta) \in U(x, t) \times \partial V(x, t)$ which maximize the right-hand side of (3.42). Let $u(\cdot)$ be a feasible control path with corresponding state trajectory $x(\cdot)$ and assume that for almost all $t \in [0, T]$ there exists $(\alpha(t), \beta(t)) \in \mathbb{R}^{n+1}$ such that $(u(t), \alpha(t), \beta(t)) \in \Phi(x(t), t)$, and

$$\frac{d}{dt} V(x(t), t) = \alpha(t)\dot{x}(t) + \beta(t). \tag{3.43}$$

Then $u(\cdot)$ is an optimal control path.

Proof It can be shown[22] that for any absolutely continuous path $\tilde{x}(\cdot)$ it holds that

$$\frac{d}{dt}\left[e^{-rt}V(\tilde{x}(t), t)\right] \in \left\{e^{-rt}\left[-rV(\tilde{x}(t), t) + \alpha\dot{\tilde{x}}(t) + \beta\right] \,\middle|\, (\alpha, \beta) \in \partial V(\tilde{x}(t), t)\right\}.$$

If $\tilde{x}(\cdot)$ is a feasible state trajectory we have $\dot{\tilde{x}}(t) = f(\tilde{x}(t), \tilde{u}(t), t)$ so that the above inclusion can also be written as

$$\begin{aligned}&\frac{d}{dt}\left[e^{-rt}V(\tilde{x}(t), t)\right] + e^{-rt}F(\tilde{x}(t), \tilde{u}(t), t) + re^{-rt}V(\tilde{x}(t), t)\\ &\in \left\{e^{-rt}[F(\tilde{x}(t), \tilde{u}(t), t) + \alpha f(\tilde{x}(t), \tilde{u}(t), t) + \beta] \,\middle|\, (\alpha, \beta) \in \partial V(\tilde{x}(t), t)\right\}.\end{aligned}$$

Because of the HJB equation (3.42), the right-hand side contains only elements which are smaller than or equal to $re^{-rt}V(\tilde{x}(t), t)$. This implies that the inequality

$$e^{-rt}F(\tilde{x}(t), \tilde{u}(t), t) \leq -(d/dt)[e^{-rt}V(\tilde{x}(t), t)]$$

must hold. Now consider the given candidate path $u(\cdot)$ with its state trajectory $x(\cdot)$. From (3.43) we obtain

$$\begin{aligned}&\frac{d}{dt}\left[e^{-rt}V(x(t), t)\right] + e^{-rt}F(x(t), u(t), t) + re^{-rt}V(x(t), t)\\ &= e^{-rt}[F(x(t), u(t), t) + \alpha(t)f(x(t), u(t), t) + \beta(t)].\end{aligned}$$

Together with the assumption $(u(t), \alpha(t), \beta(t)) \in \Phi(x(t), t)$ this implies that the equation $e^{-rt}F(x(t), u(t), t) = -(d/dt)[e^{-rt}V(x(t), t)]$ must hold. Thus we have established conditions (3.9) and (3.10). The rest of the proof is identical to the corresponding part in the proof of theorem 3.1.

Since condition (3.43) in the above theorem is new, as compared to the corresponding result in theorem 3.1, one or two remarks concerning this condition are in order. First of all, it can be shown that $(d/dt)V(x(t), t) = \alpha\dot{x}(t) + \beta$ always holds for some $(\alpha, \beta) \in \partial V(x(t), t)$. The non-trivial part of (3.43) therefore is that these are the same values $\alpha = \alpha(t)$ and $\beta = \beta(t)$ which lead to the maximization of the HJB equation (3.42). If $V(x, t)$ is differentiable at $(x, t) = (x(t), t)$ and, hence, the generalized gradient $\partial V(x(t), t)$ is a singleton, condition (3.43) is automatically satisfied. Also, we emphasize that $(u(t), \alpha(t), \beta(t)) \in \Phi(x(t), t)$ and (3.43) are required to hold only for almost all $t \in [0, T]$. This implies, in particular, that if $V(x, t)$ is nondifferentiable at $(x, t) = (x(t), t)$ only for

[22] See the references given in section 3.8.

finitely many $t \in [0, T]$ then there is no need to check condition (3.43). This situation arises quite often in applications.

If $X \subseteq \mathbb{R}^n$ is a convex set with non-empty interior and $g : X \mapsto \mathbb{R}$ is a concave function then g is known to be locally Lipschitz continuous on the interior of X. Consequently, the generalized gradient of g, $\partial g(x)$, is well defined for all x in the interior of X.[23] In the following generalization of theorem 3.2 we assume that the maximized Hamiltonian is concave and locally Lipschitz continuous as a function of the state x. Because of the properties just mentioned, local Lipschitz continuity is only an additional assumption as far as boundary points of X are concerned.

Theorem 3.6 *Assume that the state space X is a convex set and define the maximized Hamiltonian function H^* as in section 3.3. Furthermore, assume that the scrap value function S is locally Lipschitz continuous and concave. Let $u(\cdot)$ be a feasible control path with corresponding state trajectory $x(\cdot)$. If there exists an absolutely continuous function $\lambda : [0, T] \mapsto \mathbb{R}^n$ such that the maximum condition (3.13), the generalized adjoint equation*

$$r\lambda(t) - \dot{\lambda}(t) \in \partial_x H^*(x(t), \lambda(t), t), \tag{3.44}$$

and the generalized transversality condition

$$\lambda(T) \in \partial S(x(T)) \tag{3.45}$$

are satisfied, and such that the function $x \mapsto H^*(x, \lambda(t), t)$ is concave and locally Lipschitz continuous with respect to x for all $t \in [0, T]$, then $u(\cdot)$ is an optimal path. Here $\partial_x H^*(x(t), \lambda(t), t)$ denotes the generalized gradient of the function $H^*(\cdot, \lambda(t), t)$ evaluated at the point $x = x(t)$.

Proof The proof of this result is exactly the same as the proof of theorem 3.2. Just note that the generalized gradient of a concave function coincides with the subdifferential in the sense of convex analysis (see footnote 23). ■

Having extended the conditions of dynamic programming and the maximum principle to nonsmooth problems we can now reconsider example 3.2.

Example 3.2 (continued) The optimal value function of this example is given by $V(x, t) = (T - t)^2/2 + (T - t)|x|$. For $t = T$ we have $V(x, t) = 0 = S(x)$, which shows that condition (3.8) is satisfied. Now

[23]Moreover, $\partial g(x)$ coincides with the so-called subdifferential of g in the sense of convex analysis.

let us prove that the generalized HJB equation (3.42) holds. In this particular example this equation can be written as

$$0 = \max\{xu + \alpha u + \beta \mid u \in [-1, 1], (\alpha, \beta) \in \partial V(x, t)\}. \tag{3.46}$$

If $x \neq 0$ or $t = T$ then $\partial V(x, t)$ is equal to the singleton $\{((T - t)\mathrm{sgn}(x), t - T - |x|)\}$ so that the right-hand side of the HJB equation (3.46) is given by

$$\max\{xu + (T - t)\mathrm{sgn}(x)u + t - T - |x| \mid u \in [-1, 1]\}.$$

If $x > 0$, the maximum is attained at $u = 1$ and is equal to 0. Analogously, if $x < 0$, the maximum is attained at $u = -1$ and is equal to 0. If $x = 0$ and $t = T$, any $u \in [-1, 1]$ attains the maximum, which is again equal to 0. Consequently, the HJB equation is satisfied for all $(x, t) \notin Z_V$. Now let us assume that $(x, t) \in Z_V$, that is, $x = 0$ and $t < T$. The generalized gradient of V is given by $\partial V(0, t) = [t - T, T - t] \times \{t - T\}$ so that the right-hand side of (3.46) is equal to

$$\max\{\alpha u + t - T \mid u \in [-1, 1], \alpha \in [t - T, T - t]\}.$$

The maximum is again equal to 0 and it is attained by both $(u, \alpha) = (1, T - t)$ and $(u, \alpha) = (-1, t - T)$. It follows that the generalized HJB equation holds and the set of maximizing triples (u, α, β) is given by

$$\Phi(x, t) = \begin{cases} \{(1, T - t, t - T - x)\} & \text{if } x > 0, \\ \{(-1, t - T, t - T + x)\} & \text{if } x < 0, \\ \{(u, 0, 0) | u \in [-1, 1]\} & \text{if } x = 0 \text{ and } t = T, \\ \{(-1, t - T, t - T), (1, T - t, t - T)\} & \text{if } x = 0 \text{ and } t < T. \end{cases}$$

Now consider the candidate paths $u_+(\cdot)$ and $u_-(\cdot)$ defined by $u_+(t) = 1$ and $u_-(t) = -1$ for all $t \in [0, T]$. The corresponding state trajectories are given by $x_+(t) = x_0 + t$ and $x_-(t) = x_0 - t$ where $x_0 = x(0)$ denotes the initial state at $t = 0$. It is easy to see that for $u(\cdot) = u_+(\cdot)$ and $(\alpha(t), \beta(t)) = (T - t, -T - x_0)$ the conditions of theorem 3.5 are satisfied, provided that $x_0 \geq 0$. Similarly, if $x_0 \leq 0$, the conditions are satisfied for $u(\cdot) = u_-(\cdot)$ and $(\alpha(t), \beta(t)) = (t - T, x_0 - T)$. We conclude that the control paths described in the discussion of this example in section 3.2 are indeed optimal paths. Theorem 3.6 is not applicable in this example because the maximized Hamiltonian function is given by $H^*(x, \lambda, t) = |x + \lambda|$, which is not a concave function of the state variable.

To conclude, we note that the results of this section can easily be modified to apply to problems with unbounded time horizon as in section 3.6. In theorem 3.5 the boundary condition (3.8) must be replaced by (3.29), (3.31), or (3.33), depending on the optimality criterion. In theorem 3.6 the transversality condition (3.45) must be replaced by (3.30), (3.32), or (3.34), again depending on the optimality criterion. The results stated in lemma 3.1 remain valid without change. The sufficient optimality conditions of theorem 3.4 are valid for nonsmooth problems, too, provided the HJB equation is replaced by its generalized form (3.42).

3.8 Further reading

There are a number of good textbooks on optimal control theory, written especially for graduate students and researchers in economics and management science: Chiang [18], Feichtinger and Hartl [85], Kamien and Schwartz [148], Léonard and Long [160], Petit [190], Seierstad and Sydsæter [212], Sethi and Thompson [216]. All cover optimal control theory and its applications in economics and management.

Optimal control problems on unbounded time domains are frequently used in economics and management and the subject is treated in all of the above references. More advanced material on infinite horizon optimal control theory is presented in Carlson and Haurie [15]. A good discussion of the various optimality criteria for infinite horizon optimal control problems (including those of definition 3.2) can be found in Stern [225].

The capital accumulation model of example 3.4 was originally developed by Ramsey [193] who solved it by means of the Calculus of Variations. A modern treatment of this model and many other economic growth models can be found in Barro and Sala-i-Martin [2].

A basic reference for the analysis and optimization of nonsmooth problems is Clarke [28]. The HJB equation for nondifferentiable problems is considered in Clarke and Vinter [29]. Sufficiency theorems based on the maximum principle for nonsmooth problems are discussed in Hartl [119].

One important remark concerns our treatment of control constraints in the approach based on Pontryagin's maximum principle. We have stated sufficient optimality conditions for problems involving very general constraints of the form $u(t) \in U(x(t), t)$. Although our results are correct for this general class of problems, the sufficient conditions can sometimes prove too strong to be applicable. For example, it could be the case that no continuous costate trajectory $\lambda(\cdot)$ exists which satisfies the conditions of theorem 3.2. In that case it may be possible to verify weaker conditions which are still sufficient for optimality. For theorems stating weaker sufficient conditions for problems in which the constraints take the form of inequalities $h(x(t), u(t), t) \geq 0$ we refer to Seierstad and Sydsæter [212] or Feichtinger and Hartl [85]. The general idea in these results is to form a Lagrangian function by adjoining the inequality constraints to the Hamiltonian function using appropriate multipliers. This is analogous to the Lagrangian method for solving static optimization problems with inequality restrictions. It turns out that constraints which involve control variables in a nondegenerate way are easier to handle than pure state constraints of the form $h(x(t), t) \geq 0$. For an extensive survey of the various approaches for dealing with pure state inequality constraints we refer to Hartl et al. [120].

3.9 Exercises

1. Consider a market for a durable good consisting of many consumers on the demand side and a single firm on the supply side. Let the total market potential (number of potential buyers) be constant and equal to M and denote by $x(t)$ the percentage of the market potential which has bought the product from the monopolist by time t. Furthermore, denote the advertising rate of the firm at time t by $u(t)$ and assume that advertising costs are given by the quadratic function $(1/2)u(t)^2$. Try to interpret the state dynamics

$$\dot{x}(t) = u(t)M[1 - x(t)], \; x(0) = 0.$$

The goal is to maximize market penetration by time T minus the advertising cost incurred up to time T, that is, the objective functional is

$$J(u(\cdot)) = -(1/2) \int_0^T u(t)^2 \, dt + x(T).$$

Show that any candidate $u(\cdot)$ satisfying the conditions of theorem 3.2 must be constant over time, i.e., $u(t) = \bar{u}$. Show that the optimal advertising level $\bar{u}$ satisfies $\bar{u}e^{\bar{u}MT} = M$.

2. Consider the problem of a firm that extracts a nonrenewable resource and sells it at the fixed price $p > 0$. Denote by $u(t)$ the extraction rate at time t and assume that it equals the sales rate (that is, there is no inventory). Moreover, denote by $x(t)$ the resource stock remaining by time t. Interpret the dynamics

$$\dot{x}(t) = -u(t)$$

and the constraints

$x(t) \geq 0\ , \ u(t) \geq 0.$

Assume that the terminal value of one unit of nonextracted resource is equal to $q \in [0, p]$ and that the cost of extraction is increasing with respect to the extraction rate u and decreasing with respect to the remaining stock x. Interpret these assumptions. Assume more specifically that the cost is given by $c(u, x) = \gamma u^2/(2x)$, where γ is a positive constant, and solve the problem of maximizing the profit

$$J(u(\cdot)) = \int_0^T e^{-rt}[pu(t) - c(u(t), x(t))]\,dt + e^{-rT}qx(T)$$

subject to the above constraints. How would the solution change if q were larger than p?

3. Consider the optimal control problem defined by $X = \mathbb{R}$, $U(x, t) = \mathbb{R}$, $F(x, u, t) = x - u^2/(1 + t)$, $S(x) = -x$, $f(x, u, t) = u$, finite time horizon T, and discount rate $r > 0$. Find a solution to the HJB equation which satisfies the boundary condition (3.8) and which is of the form $V(x, t) = A(t) + B(t)x$. Find an optimal control path.
4. Consider the control problem with a two-dimensional state variable $(x_1, x_2) \in X = \mathbb{R}^2$ and a single control variable $u \in \mathbb{R}$ defined by the following functions: $F(x_1, x_2, u, t) = -3x_1^2 + x_2 - u^2/2$, $S(x_1, x_2) = 4x_1 - x_1^2$, $f_1(x_1, x_2, u, t) = u$, $f_2(x_1, x_2, u, t) = u - x_1$. The initial values are $x_1(0) = x_{10}$ and $x_2(0) = x_{20}$, respectively, the finite time horizon is T, and the discount rate is $r = 1$. Show that $u(t) = \phi(t, x_1(t))$ with

$$\phi(t, x_1) = \frac{1}{2}\left[7e^{3(t-T)} + e^{t-T}\right] - 2x_1$$

is an optimal Markovian strategy, independently of the values of x_{10} and x_{20}. Find an open-loop representation of the control path defined by that strategy (in other words, find an optimal open-loop strategy). Define functions ψ_1 and ψ_2 by

$$\begin{aligned}\psi_1(t, x_1, x_2) &= \frac{1}{12}[18x_{10} - 12x_{20} - 3e^{-T} - 7e^{-3T}]\\ &\quad + \frac{7}{36}e^{-2t}[-30x_{10} + 5e^{-T} + 21e^{-3T}] - \frac{2}{9}e^{t-T} + \frac{7}{3}x_1 + x_2,\\ \psi_2(t, x_1) &= \frac{1}{6}e^{-2t}[-30x_{10} + 5e^{-T} + 21e^{-3T}] - \frac{1}{3}e^{t-T} + 3x_1.\end{aligned}$$

Does either of the two Markovian strategies $u(t) = \psi_1(t, x_1(t), x_2(t))$ and $u(t) = \psi_2(t, x_1(t))$ generate the optimal control path determined above?
5. Consider the optimal control problem with infinite time horizon $T = \infty$, a single state variable $x \in [0, 1]$, and two control variables u_1

$\in [0, 1]$ and $u_2 \in [0, 1]$. The utility function is given by $F(x, u_1, u_2, t) = x - u_1 x - u_2(1 - x)$ and the system dynamics by $f(x, u_1, u_2, t) = (u_1 + u_2)(1 - x)$. Let the discount rate be $r \in (0, 1)$ and the initial state $x_0 = 0$. Define two numbers τ and σ by

$$\tau = -\frac{2}{r+2}\ln\frac{r+3}{(r+1)^2} - \ln\frac{r+1}{r+3} \text{ and } \sigma = -\frac{1}{r+2}\ln\frac{r+3}{(r+1)^2} - \ln\frac{r+1}{r+3}$$

and the control path

$$(u_1(t), u_2(t)) = \begin{cases} (1, 0) & \text{if } t \in [0,\tau), \\ (1, 1) & \text{if } t \in [\tau,\sigma), \\ (0, 1) & \text{if } t \in [\sigma,\infty). \end{cases}$$

Show that this is a catching up optimal path.

6. Consider the finite horizon optimal control problem with state variable $x \in \mathbb{R}$, control variable $u \in \mathbb{R}$, utility function $F(x, u, t) = \min\{x, 2 - x\} - u^2/2$ and system dynamics $f(x, u, t) = 1 + u$. The discount rate r, the initial state x_0, and the scrap value function S are all equal to 0. Define $\tau(T) = [T - 1 + \sqrt{T^2 - 2T + 7}]/3$ and prove that $u(\cdot)$ defined in the following way is an optimal control path for this problem.
If $0 \leq T \leq \sqrt{3} - 1$ then $u(t) = T - t$ for all $t \in [0, T]$. If $\sqrt{3} - 1 < T \leq \sqrt{2} + 1$ then

$$u(t) = \begin{cases} 2\tau(T) - T - t & \text{if } t \in [0, \tau(T)], \\ t - T & \text{if } t \in (\tau(T), T]. \end{cases}$$

If $T > \sqrt{2} + 1$ then

$$u(t) = \begin{cases} \sqrt{2} - 1 - t & \text{if } t \in [0, \sqrt{2}], \\ -1 & \text{if } t \in (\sqrt{2}, T - 1], \\ t - T & \text{if } t \in (T - 1, T]. \end{cases}$$

7. Consider the infinite horizon optimal control problem

$$\text{Maximize} \int_0^\infty e^{-t}\sqrt{u(t)}dt$$

subject to $\dot{x}(t) = x(t) - u(t), u(t) \geq 0, x(t) \geq 0, x(0) = x_0 > 0$.

Derive the HJB equation for this problem and verify that the function $V(x, t; \alpha) = \alpha x + 1/(4\alpha)$ satisfies this equation whenever α is a positive real number. Show that the right-hand side of the HJB equation with $V = V(\cdot; \alpha)$ is maximized at the constant $u = 1/(4\alpha^2)$. Now assume that $\alpha > 1/(2\sqrt{x_0})$. Verify that the constant path $u(t) = 1/(4\alpha^2)$ is feasible for the problem. Compute the corresponding state trajectory. Can you verify the conditions of theorem 3.3 with the function $V(\cdot; \alpha)$ to show that this path is a catching up optimal path? If not, why not? Now choose $\alpha = 1/(2\sqrt{x_0})$ and show that the constant control path $u(t) = 1/(4\alpha^2)$ is generated by the nondegenerate Markovian strategy $u(t) = x(t)$. Can you verify the (catching up optimality) conditions of theorem 3.3 for this path using the value function $V(\cdot, \alpha)$? Try also to verify the (catching up optimality) conditions of the theorem using the optimal value function $V(x, t) = \sqrt{x}$.

4 Markovian equilibria with simultaneous play

We can now begin to deal with the main topic of this book: the analysis of differential games. In this chapter we shall see how the concept of Nash equilibrium introduced in chapter 2 can be applied in a dynamic setting. Each of the N players seeks to maximize his objective functional – the present value of utility derived over a finite or infinite time horizon – by designing a strategy for those variables which are under his control. His choice influences the evolution of the state of the game via a differential equation (the system dynamics) as well as the objective functionals of his opponents. Under the assumptions of the present chapter, we shall see that each player faces an optimal control problem of the form discussed in chapter 3. An important feature of each of these player-specific control problems is that the actions of the opponents become part of the definition of the problem. The most important assumptions of the present chapter are (i) that players make their choices simultaneously and (ii) that they represent the solutions to their control problems by Markovian strategies. We state conditions which can be used to verify that a given N-tuple of Markovian strategies constitutes a Nash equilibrium. We also discuss the important concepts of time consistency and subgame perfectness.

4.1 The Nash equilibrium

Consider a differential game which extends over the bounded time interval $[0, T]$ or the unbounded time interval $[0, \infty)$. To handle both cases simultaneously, we define a time interval $[0, T\rangle$ by $[0, T\rangle = [0, T]$, if $T < \infty$, and $[0, T\rangle = [0, \infty)$, if $T = \infty$. The state of the game at each instant $t \in [0, T\rangle$ is described by a vector $x(t) \in X$ where $X \subseteq \mathbb{R}^n$ is the state space of the game. The initial state of the game is a fixed constant $x_0 \in X$. There are N players $i = 1, 2, \ldots, N$. We shall denote player-specific variables, functions, and parameters by upper indices. At each instant of time $t \in [0, T)$, each player $i \in \{1, 2, \ldots, N\}$ chooses a control variable $u^i(t)$ from his set of feasible controls $U^i(x(t), u^{-i}(t), t) \subseteq \mathbb{R}^{m^i}$. In general, this set depends on time t, the current state $x(t)$, and the vector $u^{-i}(t)$ consisting of all other players' controls at time t, i.e.,

$$u^{-i}(t) = \big(u^1(t), u^2(t), \ldots, u^{i-1}(t), u^{i+1}(t), \ldots, u^N(t)\big).$$

The state of the game evolves according to the differential equation

$$\dot{x}(t) = f\big(x(t), u^1(t), u^2(t), \ldots, u^N(t), t\big), \qquad x(0) = x_0,$$

where the system dynamics f are defined on the set

$$\Omega = \big\{(x, u^1, u^2, \ldots, u^N, t) \mid x \in X, t \in [0, T), u^i \in U^i(x, u^{-i}, t),$$
$$i = 1, 2, \ldots, N\}$$

and have values in $\mathbb{R}^n$. Each player $i \in \{1, 2, \ldots, N\}$ seeks to maximize his objective functional[1]

$$J^i(u^i(\cdot)) = \int_0^T e^{-r^i t} F^i(x(t), u^1(t), u^2(t), \ldots, u^N(t), t)\, dt + e^{-r^i T} S^i(x(T)).$$

Here, $F^i : \Omega \mapsto \mathbb{R}$ is player i's utility function, r^i his individual rate of time preference, and $S^i : X \mapsto \mathbb{R}$ his scrap value function. In the case $T = \infty$ we assume that $S(x) = 0$ for all $x \in X$.

Recall from chapter 2 that a Nash equilibrium is an N-tuple of strategies $(\phi^1, \phi^2, \ldots, \phi^N)$ such that, given the opponents' equilibrium strategies, no player has an incentive to change his own strategy. In a differential game the following is true: if all opponents of player i use Markovian strategies $u^j(t) = \phi^j(x(t), t), j \neq i$, then player i faces a control problem of the form discussed in chapter 3. To see this, note that in this case player i's decision problem can be rewritten as

$$\begin{aligned} \text{Maximize } J^i_{\phi^{-i}}(u^i(\cdot)) &= \int_0^T e^{-r^i t} F^i_{\phi^{-i}}(x(t), u^i(t), t)\, dt + e^{-r^i T} S^i(x(T)) \\ \text{subject to } \dot{x}(t) &= f^i_{\phi^{-i}}(x(t), u^i(t), t), \\ x(0) &= x_0, \\ u^i(t) &\in U^i_{\phi^{-i}}(x(t), t), \end{aligned} \tag{4.1}$$

where

[1]We write player i's objective functional as a function of his own control path $u^i(\cdot)$ only, although it also depends on the control paths of the opponents. This is done in order to emphasize that the objective functional J^i is to be maximized only with respect to $u^i(\cdot)$.

$$F^i_{\phi^{-i}}(x, u^i, t) = F^i(x, \phi^1(x,t), \ldots, \phi^{i-1}(x,t), u^i, \phi^{i+1}(x,t), \ldots, \phi^N(x,t), t),$$
$$f^i_{\phi^{-i}}(x, u^i, t) = f(x, \phi^1(x,t), \ldots, \phi^{i-1}(x,t), u^i, \phi^{i+1}(x,t), \ldots, \phi^N(x,t), t),$$
$$U^i_{\phi^{-i}}(x, t) = U^i\big(x, \phi^1(x,t), \ldots, \phi^{i-1}(x,t), \phi^{i+1}(x,t), \ldots, \phi^N(x,t), t\big). \tag{4.2}$$

For any given $(N-1)$-tuple $\phi^{-i} = (\phi^1, \ldots, \phi^{i-1}, \phi^{i+1}, \ldots, \phi^N)$ of functions $\phi^j : X \times [0, T\rangle \mapsto \mathbb{R}^{m^j}, j \neq i$, problem (4.1) is in fact an optimal control model of the form described in section 3.1. We can now define a Markovian Nash equilibrium for the differential game.

Definition 4.1 The N-tuple $(\phi^1, \phi^2, \ldots, \phi^N)$ of functions $\phi^i : X \times [0, T\rangle \mapsto \mathbb{R}^{m^i}$, $i \in \{1, 2, \ldots, N\}$, is called a Markovian Nash equilibrium if, for each $i \in \{1, 2, \ldots, N\}$, an optimal control path $u^i(\cdot)$ of the problem (4.1) exists and is given by the Markovian strategy $u^i(t) = \phi^i(x(t), t)$.

The above definition shows that finding a Markovian Nash equilibrium of an N-player differential game amounts to finding Markovian strategies for the solutions of a system of N interdependent optimal control models. To find and characterize Markovian Nash equilibria of differential games, the methods described in chapter 3 are therefore of great importance.

If we replace the assumption that optimal paths are defined by Markovian strategies by the assumption that the optimal paths are given by open-loop strategies, then we obtain the following definition.

Definition 4.2 The N-tuple $(\phi^1, \phi^2, \ldots, \phi^N)$ of functions $\phi^i : [0, T\rangle \mapsto \mathbb{R}^{m^i}$, $i \in \{1, 2, \ldots, N\}$, is called an open-loop Nash equilibrium if, for each $i \in \{1, 2, \ldots, N\}$, an optimal control path $u^i(\cdot)$ of the problem (4.1) exists and is given by the open-loop strategy $u^i(t) = \phi^i(t)$.

If the horizon of the game is unbounded, i.e., if $T = \infty$, one has to be precise about what is meant by optimality in the above definitions. For example, one could consider open-loop Nash equilibria with optimality understood in the sense of overtaking optimality, Markovian Nash equilibria with sporadically catching up optimality, etc. (see section 3.6).

In section 3.5 we emphasized that a solution of an optimal control problem can be represented by different strategies (e.g., open-loop and nondegenerate Markovian), but that the different representations all correspond to the same control path. At first glance, this seems to make one of the above definitions of Nash equilibria superfluous. However, there is a crucial difference between a single decision maker control problem and a differential game: in a differential game, different representations of the same feasible control paths $u^{-i}(\cdot)$ lead to different optimization problems for player i. For example, if ψ^{-i} is an $(N-1)$-tuple of nondegenerate Markovian strategies defining the control paths $u^{-i}(\cdot)$, and ϕ^{-i} is an $(N-1)$-tuple of open-loop strategies for $u^{-i}(\cdot)$, then the functions $F^i_{\phi^{-i}}$, $f^i_{\phi^{-i}}$, and $U^i_{\phi^{-i}}$ in (4.1) are, in general, different from their counterparts $F^i_{\psi^{-i}}, f^i_{\psi^{-i}}$, and $U^i_{\psi^{-i}}$. It follows from this observation that the set of open-loop Nash equilibria of a particular differential game is typically different from the set of Markovian Nash equilibria. On the other hand, every open-loop strategy is by definition also a (degenerate) Markovian strategy, which means that every open-loop Nash equilibrium is also a Markovian Nash equilibrium. Thus, the set of open-loop Nash equilibria of a particular game is a subset of the set of all Markovian Nash equilibria. In general, it is a proper subset.

Before we illustrate the derivation of Markovian Nash equilibria, let us point out that the assumptions concerning the strategies in definitions 4.1 and 4.2 are not the only ones that make sense. One could, for example, consider equilibria in which some of the players represent their optimal control paths in open-loop form while others choose nondegenerate Markovian strategies. We shall not discuss such situations in detail, but recall from section 3.5 that the choice of the strategy reflects the informational assumptions of the model.[2] By computing a nondegenerate Markovian Nash equilibrium one makes the assumption that the state variable can be observed and that the players condition their actions on these observations. On the other hand, if a player uses an open-loop strategy, he either cannot observe the state variable or he chooses to commit to a fixed time function. To summarize, the choice to solve a differential game for an open-loop equilibrium or for a Markovian equilibrium (or for equilibria in which some players use open-loop strategies while others employ non-degenerate Markovian strategies) is part of the modelling stage and one should try to analyse that equilibrium which describes best the actual situation at hand (see also chapter 2).

Example 4.1 Consider a differential game with $N=2$ players and finite horizon T. To save on notation, denote the control variables of the two players by u and v instead of u^1 and u^2. The state space is $X=[0,\infty)$, the initial state is a fixed number $x_0>0$, and the set of feasible controls is $U^1=[0,\infty)$ for player 1 and $U^2=[0,1]$ for player 2. The objective functionals are

[2]See example 4.1 for a simple differential game in which Nash equilibria with asymmetric information can be analysed.

$$J^1(u(\cdot)) = \int_0^T e^{-rt}\Big[v(t) - x(t) - \frac{\alpha}{2}u(t)^2\Big]dt$$

and

$$J^2(v(\cdot)) = \int_0^T e^{-rt}[v(t) - x(t)]\,dt.$$

The system dynamics are given by

$$\dot{x}(t) = 1 + v(t) - u(t)\sqrt{x(t)}.$$

Let us first try to find an open-loop Nash equilibrium of the above game, that is, a pair (ϕ, ψ) where $\phi : [0, T] \mapsto [0, \infty)$ and $\psi : [0, T] \mapsto [0, 1]$ are the strategies for player 1 and player 2, respectively. If player 2 chooses to play $v(t) = \psi(t)$ then player 1's problem can be written as

$$\begin{aligned} &\text{Maximize } \int_0^T e^{-rt}[\psi(t) - x(t) - \frac{\alpha}{2}u(t)^2]\,dt \\ &\text{subject to } \dot{x}(t) = 1 + \psi(t) - u(t)\sqrt{x(t)}, \\ &\qquad x(0) = x_0, \\ &\qquad u(t) \geq 0. \end{aligned} \tag{4.3}$$

Since $\psi(\cdot)$ is assumed by player 1 to be a fixed function, the maximization of the integral in (4.3) is equivalent to the maximization of $\int_0^T e^{-rt}[-\, x(t) - (\alpha/2)u(t)^2]\,dt$ so that problem (4.3) is equivalent to the problem discussed in example 3.3. with $\beta(t) = 1 + \psi(t) = 1 + v(t)$. We have seen in section 3.4 that this problem has the optimal open-loop strategy $u(t) = \phi(t) = -\tilde{\lambda}(t)\sqrt{\tilde{x}(t)}/\alpha$, where $\tilde{\lambda}(\cdot)$ is the unique solution of

$$\dot{\lambda}(t) = 1 + r\lambda(t) - \frac{1}{2\alpha}\lambda(t)^2, \qquad \lambda(T) = 0 \tag{4.4}$$

and

$$\tilde{x}(t) = e^{\int_0^t \tilde{\lambda}(\tau)/\alpha\, d\tau}\left\{\int_0^t [1 + v(s)]e^{-\int_0^s \tilde{\lambda}(\tau)/\alpha\, d\tau}\, ds + x_0\right\}.$$

We have also seen (cf. equation (3.21)) that

$$\tilde{\lambda}(t) = \frac{2\Big[1 - e^{C(t-T)}\Big]}{(r - C)e^{C(t-T)} - (r + C)}, \tag{4.5}$$

where $C = \sqrt{r^2 + 2/\alpha}$. Finally, the state trajectory generated by this solution is given by $x(t) = \tilde{x}(t)$. Note that the formula for $\tilde{x}(t)$ now depends on player 2's control path $v(\cdot)$, still to be determined.

Now consider player 2's control problem. If player 1 chooses $u(t) = \phi(t)$, player 2's problem can be written as

$$\begin{aligned} &\text{Maximize } \int_0^T e^{-rt}[v(t) - x(t)]\,dt \\ &\text{subject to } \dot{x}(t) = 1 + v(t) - \phi(t)\sqrt{x(t)}, \\ &\qquad x(0) = x_0, \\ &\qquad v(t) \in [0, 1]. \end{aligned} \tag{4.6}$$

Denoting by μ the costate variable of player 2, the Hamiltonian function for this problem is given by $H(x, v, \mu, t) = v - x + \mu[1 + v - \phi(t)\sqrt{x}]$. Maximization with respect to $v \in [0, 1]$ yields $v = 0$ if $\mu < -1$ and $v = 1$ if $\mu > -1$. If $\mu = -1$ then $H(x, v, \mu, t) = -1 - x + \phi(t)\sqrt{x}$, independently of v. These properties imply that the maximized Hamiltonian function is given by

$$H^*(x, \mu, t) = \begin{cases} -x + \mu[1 - \phi(t)\sqrt{x}] & \text{for } \mu \leq -1, \\ 1 - x + \mu[2 - \phi(t)\sqrt{x}] & \text{for } \mu \geq -1. \end{cases}$$

The adjoint equation and transversality condition for player 2's problem are

$$\dot{\mu}(t) = 1 + r\mu(t) + \frac{1}{2\sqrt{x(t)}}\mu(t)\phi(t), \qquad \mu(T) = 0.$$

Using $\phi(t) = -\lambda(t)\sqrt{x(t)}/\alpha$ this can be written as

$$\dot{\mu}(t) = 1 + r\mu(t) - \frac{1}{2\alpha}\mu(t)\lambda(t), \qquad \mu(T) = 0. \tag{4.7}$$

The boundary value problem consisting of (4.4) and (4.7) has a unique solution which is given by $\mu(t) = \lambda(t) = \tilde{\lambda}(t)$ with $\tilde{\lambda}(t)$ from (4.5). The function $\tilde{\lambda}(\cdot)$ is easily seen to be nonpositive and strictly increasing on $[0, T]$. It depends on the parameters α, r, and T whether $\tilde{\lambda}(t) > -1$ for all $t \in [0, T]$ or whether $\tilde{\lambda}(t)$ can be smaller than -1 for some $t \in [0, T]$. Because of the monotonicity of $\tilde{\lambda}(\cdot)$, however, we know that in the latter case there exists a number $\tau \in (0, T)$ such that $\tilde{\lambda}(t) < -1$ for $t \in [0, \tau)$ and $\tilde{\lambda}(t) > -1$ for all $t \in (\tau, T]$. Careful analysis of (4.5) reveals that such a number τ exists if and only if $r + C < 2$ and $T > [\ln(2 - r + C) - \ln(2 - r - C)]/C$, in which case τ is given by

$$\tau = T + [\ln(2 - r - C) - \ln(2 - r + C)]/C.$$

In all other cases let us formally set $\tau = 0$. We summarize our results as follows. There exists a candidate for an open-loop Nash equilibrium, given by

$$u(t) = \phi(t) = -\tilde{\lambda}(t)\sqrt{\tilde{x}(t)}/\alpha$$

$$v(t) = \psi(t) = \begin{cases} 0 & \text{for } t \in [0, \tau), \\ 1 & \text{for } t \in [\tau, T], \end{cases}$$

where $\tilde{\lambda}(t)$, $\tilde{x}(t)$, and τ are specified as above.

From the discussion of example 3.3 we know that $u(\cdot)$ is indeed an optimal solution of player 1's problem. To verify that the above candidate is an open-loop Nash equilibrium it suffices therefore to prove that $v(\cdot)$ is an optimal control path in player 2's problem. This, however, follows from theorem 3.2 by noting that $\mu(t) = \tilde{\lambda}(t) \leq 0$ for all t, which shows that the maximized Hamiltonian function $H^*(x, \mu(t), t)$ of player 2's problem is a concave function with respect to x. This concludes the derivation of an open-loop Nash equilibrium for this example.

We have already noted that every open-loop Nash equilibrium is also a degenerate Markovian Nash equilibrium. In the next step, we want to find a nondegenerate Markovian Nash equilibrium of this differential game. We start by conjecturing that the Markovian representation of the optimal control path for player 2 is again degenerate in the sense that it does not depend on the current state. In this case we would have $v(t) = \psi(t)$ (as in the open-loop Nash equilibrium from above but possibly with a different function ψ) so that player 1's optimal control problem is again equivalent to example 3.3 with $\beta(t) = 1 + \psi(t) = 1 + v(t)$. In sections 3.4 and 3.5 we saw that a nondegenerate Markovian representation of the optimal control path of this problem is $u(t) = \phi(x(t), t)$, with $\phi(x, t) = -\tilde{\lambda}(t)\sqrt{x}/\alpha$. Here, $\tilde{\lambda}(t)$ is again determined by (4.5). If we substitute this solution into player 2's optimal control problem we get

$$\begin{aligned} &\text{Maximize} \int_0^T e^{-rt}[v(t) - x(t)]\,dt \\ &\text{subject to } \dot{x}(t) = 1 + v(t) + \tilde{\lambda}(t)x(t)/\alpha, \\ &\qquad x(0) = x_0, \\ &\qquad v(t) \in [0, 1]. \end{aligned} \tag{4.8}$$

Note how this problem differs from (4.6), which was player 2's optimization problem in the open-loop Nash game. For example, the system dynamics are now linear with respect to the state variable, whereas they were nonlinear in (4.6). Proceeding in exactly the same way as in the open-loop case we obtain from the maximum condition that $v(t) = 0$ if $\tilde{\mu}(t) < -1$ and $v(t) = 1$ if $\tilde{\mu}(t) > -1$. The costate, now denoted by $\tilde{\mu}$ to distinguish it from the costate μ in the open-loop case, satisfies the adjoint equation and transversality condition

$$\dot{\tilde{\mu}}(t) = 1 + r\tilde{\mu}(t) - \frac{1}{\alpha}\tilde{\mu}(t)\lambda(t), \qquad \tilde{\mu}(T) = 0. \tag{4.9}$$

This equation differs from (4.7) in that the term $\tilde{\mu}(t)\lambda(t)$ is multiplied by $1/\alpha$ as opposed to $1/(2\alpha)$. The unique solution of the boundary value problem consisting of (4.4) and (4.9) is more complicated than the solution of the corresponding problem (4.4) and (4.7) so that we omit an explicit formula.[3] However, it is easy to see that $\tilde{\mu}(t) < 0$ must hold for all $t \in [0, T)$ and that

$$\begin{aligned}\dot{\tilde{\mu}}(t)\Big|_{\mu(t)=\tilde{\mu}(t)<0} &= 1 + r\mu(t) - \frac{1}{\alpha}\mu(t)\lambda(t)\\ &< 1 + r\mu(t) - \frac{1}{2\alpha}\mu(t)\lambda(t) = \dot{\mu}(t)\Big|_{\mu(t)=\tilde{\mu}(t)<0}.\end{aligned}$$

This implies that the trajectory $\tilde{\mu}(\cdot)$ is less steep than the trajectory $\mu(\cdot)$ from the open-loop case. Furthermore, since both trajectories terminate in $\mu(T) = \tilde{\mu}(T) = 0$ it must hold that $\tilde{\mu}(t) > \mu(t)$ for all $t \in [0, T)$. As in the open-loop equilibrium it could be that $\tilde{\mu}(t) > -1$ for all $t \in [0, T]$ or that there exists a single switching point $\sigma \in (0, T)$ such that $\tilde{\mu}(t) < -1$ for $t \in [0, \sigma)$ and $\tilde{\mu}(t) > -1$ for $t \in (\sigma, T]$. Since $\tilde{\mu}(t) > \mu(t)$ for all t this switching point must be smaller than τ. However, there is also a third possibility which has no counterpart in the open-loop case. A numerical example in which this possibility occurs is depicted in figure 4.1, where the parameter values are $\alpha = 2$, $r = 1/4$, and $T = 10$. The trajectory $\tilde{\mu}(\cdot)$ starts at $t = 0$ at a level greater than -1, then falls below -1, and finally rises again to satisfy the terminal condition $\tilde{\mu}(10) = 0$. Consequently there are two switching points σ_1 and σ_2, with $0 < \sigma_1 < \sigma_2 < 10$, such that

$$v(t) = \psi(t) = \begin{cases} 1 & \text{for } t \in [0, \sigma_1), \\ 0 & \text{for } t \in [\sigma_1, \sigma_2), \\ 1 & \text{for } t \in [\sigma_2, T] \end{cases}$$

is player 2's optimal control path. It depends on the parameters α, r, and T which of the three cases (no switching point, one switching point, or

[3]Since an explicit formula for $\lambda(t)$ is given by (4.5), equation (4.9) is a nonautonomous linear differential equation which can be solved explicitly by, say, the variation of constants method.

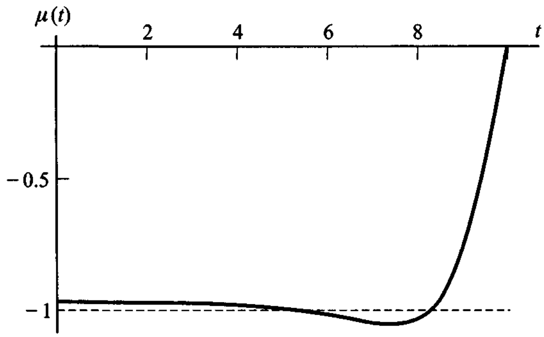

Figure 4.1 Solution of (4.9) for $\alpha = 2$, $r = \frac{1}{4}$, and $T = 10$

two switching points) occurs. In any case, the strategies $v(t) = \psi(t)$ for player 2 and $\phi(x, t) = -\tilde{\lambda}(t)\sqrt{x}/\alpha$ for player 1 constitute a candidate for an equilibrium. By noting that player 2's maximized Hamiltonian function is linear (and hence concave) with respect to the state variable we see that this candidate is indeed a nondegenerate Markovian Nash equilibrium of the game. Since $\psi(t)$ does not depend on $x(t)$ this equilibrium could also be regarded as a Nash equilibrium of a different game in which player 1 has full Markovian information whereas player 2 has only open-loop information.

4.2 Equilibrium conditions

In a Markovian Nash equilibrium of a differential game, each player solves an optimal control problem of the form discussed in chapter 3 given the fixed (equilibrium) strategies of his opponents. Therefore, we can use the conditions developed in chapter 3 to verify whether an N-tuple $(\phi^1, \phi^2, \ldots, \phi^N)$ is a Markovian Nash equilibrium. In this section we summarize these conditions for general Markovian Nash equilibria of the N-player game under consideration. We do not treat the special case of open-loop Nash equilibria separately, but in the ensuing discussion of the conditions we shall point out some of the simplifications which arise in the open-loop case. Throughout this section we use the notation from section 4.1 (see, in particular, equation (4.2)).

Theorem 4.1 *Let $(\phi^1, \phi^2, \ldots, \phi^N)$ be a given N-tuple of functions $\phi^i : X \times [0, T) \mapsto \mathbb{R}^{m^i}$ and make the following assumptions:*

(i) there exists a unique absolutely continuous solution $x : [0, T) \mapsto X$ of the initial value problem

$$\dot{x}(t) = f(x(t), \phi^1(x(t), t), \phi^2(x(t), t), \ldots, \phi^N(x(t), t), t), \; x(0) = x_0,$$

(ii) for all $i \in \{1, 2, \ldots, N\}$ there exists a continuously differentiable function $V^i : X \times [0, T\rangle \mapsto \mathbb{R}$ such that the HJB equations

$$r^i V^i(x, t) - V^i_t(x, t) = \max\left\{F^i_{\phi^{-i}}(x, u^i, t) + V^i_x(x, t) f^i_{\phi^{-i}}(x, u^i, t) \mid u^i \in U^i_{\phi^{-i}}(x, t)\right\} \tag{4.10}$$

are satisfied for all $(x, t) \in X \times [0, T\rangle$,
(iii) if $T < \infty$ then $V^i(x, T) = S^i(x)$ for all $i \in \{1, 2, \ldots, N\}$ and all $x \in X$,
(iv) if $T = \infty$ then for all $i \in \{1, 2, \ldots, N\}$ either V^i is a bounded function and $r^i > 0$ or V^i is bounded below, $r^i > 0$, and $\limsup_{t\to\infty} e^{-r^i t} V^i(x(t), t) \leq 0$.
Denote by $\Phi^i(x, t)$ the set of all $u^i \in U^i_{\phi^{-i}}(x, t)$ which maximize the right-hand side of (4.10). If $\phi^i(x(t), t) \in \Phi^i(x(t), t)$ holds for all $i \in \{1, 2, \ldots, N\}$ and almost all $t \in [0, T\rangle$ then $(\phi^1, \phi^2, \ldots, \phi^N)$ is a Markovian Nash equilibrium. (If $T = \infty$, optimality is understood in the sense of catching up optimality.)

Proof If $T < \infty$ the result follows from application of theorem 3.1 to problem (4.1). If $T = \infty$ the result follows from application of theorem 3.3 and lemma 3.1 to (4.1). ■

In the case of an unbounded time interval we have confined the theorem to the catching up optimality criterion. Furthermore, we have only stated the strong forms of the limiting transversality condition given in lemma 3.1. The reader should not find it difficult to modify condition (iv) of theorem 4.1 in order to cope with stronger optimality criteria or weaker transversality conditions. Essentially the same remark also applies to the following theorem.

Theorem 4.2 *Let an N-tuple $(\phi^1, \phi^2, \ldots, \phi^N)$ of functions $\phi^i : X \times [0, T\rangle \mapsto \mathbb{R}^{m^i}$ be given and let assumption (i) from theorem 4.1 be satisfied. Define for all $i \in \{1, 2, \ldots, N\}$ the Hamiltonians $H^i_{\phi^{-i}} : X \times \mathbb{R}^{m^i} \times \mathbb{R}^n \times [0, T\rangle \mapsto \mathbb{R}$ by*

$$H^i_{\phi^{-i}}(x, u^i, \lambda^i, t) = F^i_{\phi^{-i}}(x, u^i, t) + \lambda^i f^i_{\phi^{-i}}(x, u^i, t)$$

and the maximized Hamiltonians $H^{i*}_{\phi^{-i}} : X \times \mathbb{R}^n \times [0, T\rangle \mapsto \mathbb{R}$ by

$$H^{i*}_{\phi^{-i}}(x, \lambda^i, t) = \max\left\{H^i_{\phi^{-i}}(x, u^i, \lambda^i, t) \mid u^i \in U^i_{\phi^{-i}}(x, t)\right\}.$$

Assume that the state space X is convex, the scrap value functions S^i are continuously differentiable and concave, and that there exist N absolutely continuous functions $\lambda^i : [0, T\rangle \mapsto \mathbb{R}^n$ such that

(i) the maximum condition $H^i_{\phi^{-i}}(x(t), \phi^i(x(t), t), \lambda^i(t), t) = H^{i*}_{\phi^{-i}}(x(t), \lambda^i(t), t)$ *holds for all $i \in \{1, 2, \ldots, N\}$ and almost all $t \in [0, T\rangle$,*
(ii) the adjoint equation $\dot{\lambda}^i(t) = r^i\lambda^i(t) - (\partial/\partial x)H^{i*}_{\phi^{-i}}(x(t), \lambda^i(t), t)$ *holds for all $i \in \{1, 2, \ldots, N\}$ and for almost all $t \in [0, T\rangle$,*
(iii) if $T < \infty$ then $\lambda^i(T) = S^i_x(x(T))$ holds for all $i \in \{1, 2, \ldots, N\}$,
(iv) if $T = \infty$ then either $\lim_{t\to\infty} e^{-r^i t}\lambda^i(t)\tilde{x}(t) = 0$ holds for all $i \in \{1, 2, \ldots, N\}$ and all feasible state trajectories $\tilde{x}(\cdot)$, or there exists a vector $a \in \mathbb{R}^n$ such that $x \geq a$ for all $x \in X$, $\lambda^i(t) \geq 0$ for all $i \in \{1, 2, \ldots, N\}$ and all sufficiently large t, and $\limsup_{t\to\infty} e^{-r^i t}\lambda^i(t)[x(t) - a] \leq 0$,
(v) the function $x \mapsto H^{i*}_{\phi^{-i}}(x, \lambda^i(t), t)$ is continuously differentiable and concave for all $i \in \{1, 2, \ldots, N\}$ and all $t \in [0, T\rangle$.

Then $(\phi^1, \phi^2, \ldots, \phi^N)$ is a Markovian Nash equilibrium. (If $T = \infty$, optimality is understood in the sense of catching up optimality.)

Proof If $T < \infty$ the result follows from application of theorem 3.2 to problem (4.1). If $T = \infty$ the result follows from application of theorem 3.3 and lemma 3.1 to (4.1). ■

We now discuss the equilibrium conditions stated in the above two theorems with the goal of pointing out certain tricks concerning their application. We emphasize that these tricks should only be regarded as heuristics and not be interpreted as proven results. We do not state any formal theorems and frequently assume (without explicitly mentioning it) that all functions are sufficiently smooth so that the derivatives used in the formulas exist and are continuous.

Our first remark is that one can often take advantage of symmetries. For example, if the game is completely symmetric in the sense that all players have the same utility function, the same scrap value function, the same sets of feasible controls, and the same discount rate, and if the state equation is symmetric with respect to the players' controls, then one can try to find a symmetric Nash equilibrium, that is, an equilibrium $(\phi^1, \phi^2, \ldots, \phi^N)$ with $\phi^i = \phi^j$ for all $i, j \in \{1, 2, \ldots, N\}$. This reduces the dimension of the problem considerably, because in that case the optimal value functions V^i are also the same for all players $i \in \{1, 2, \ldots, N\}$. Note, however, that there is no result which states that a symmetric game admits (only) symmetric Nash equilibria. As a matter of fact, this statement is in general false as can be seen from simple examples. However, in many economic situations which can be modelled as a symmetric differential game it is more natural to look for symmetric equilibria than for asymmetric ones.

Next consider the conditions in theorem 4.1 which are based on the HJB equation. It has already been mentioned in chapter 3 that solving an HJB equation for an optimal control problem is usually quite difficult and that closed-form solutions are only known for relatively few models. This remark is even more relevant for HJB equations arising in differential games since these equations depend on the strategies of the opponents, which are unknown unless one has already solved their HJB equation. One of the main obstacles to solving (4.10) is that the equations are highly nonlinear, implicit partial differential equations for the function V^i. To explain this further, assume that the maximization on the right-hand side of (4.10) yields a unique solution. Of course, this solution depends on x, on t, and on the gradient vector $V^i_x(x, t)$ of the unknown function V^i. Formally, we can write the optimal u^i as $u^i = g^i(x, t, V^i_x(x, t))$. Substituting this into (4.10) we obtain

$$\begin{aligned} r^i V^i(x, t) - V^i_t(x, t) = F^i_{\phi^{-i}}(x, g^i(x, t, V^i_x(x, t)), t) \\ + V^i_x(x, t) f^i_{\phi^{-i}}(x, g^i(x, t, V^i_x(x, t)), t). \end{aligned} \tag{4.11}$$

The derivatives $V^i_x(x, t)$ occur on the right-hand side of this equation as arguments of the (usually nonlinear) functions $F^i_{\phi^{-i}}$, $f^i_{\phi^{-i}}$, and g^i. This makes the solution of (4.11) extremely difficult. Quite often it is easier to analyse a partial differential equation for the strategies $\phi^i(x, t)$ which can be derived from (4.11). To this end first solve $\phi^i(x, t) = u^i = g^i(x, t, V^i_x(x, t))$ for $V^i_x(x, t)$. This is especially easy if player i has only a single control variable (i.e., $m^i = 1$), the maximum on the right-hand side of (4.10) is attained in the interior of $U^i_{\phi^{-i}}(x, t)$, and the functions $F^i_{\phi^{-i}}$ and $f^i_{\phi^{-i}}$ are differentiable with respect to u^i, because in that case we must have

$$\frac{\partial}{\partial u^i}\left[F^i_{\phi^{-i}}(x, u^i, t) + V^i_x(x, t) f^i_{\phi^{-i}}(x, u^i, t)\right]\Big|_{u^i=\phi^i(x,t)} = 0.$$

This equation leads to

$$V^i_x(x, t) = -\left[\frac{\partial}{\partial u^i} F^i_{\phi^{-i}}(x, u^i, t) \Big/ \frac{\partial}{\partial u^i} f^i_{\phi^{-i}}(x, u^i, t)\right]\Big|_{u^i=\phi^i(x,t)}.$$

But even in the case of boundary solutions or multi-dimensional control spaces one can often derive an equation of the form

$$V^i_x(x, t) = G^i(x, t, \phi^i(x, t)). \tag{4.12}$$

Because $u^i = \phi^i(x, t)$ is assumed to maximize the right-hand side of (4.10), we must have

$$r^i V^i(x, t) - V_t^i(x, t) = F^i_{\phi^{-i}}(x, \phi^i(x, t), t) + G^i(x, t, \phi^i(x, t)) f^i_{\phi^{-i}}(x, \phi^i(x, t), t).$$

Since this equation has to hold for all $(x, t) \in X \times [0, T\rangle$, the derivative with respect to x of the left-hand side must be equal to the derivative with respect to x of the right-hand side, i.e.,

$$\begin{aligned}
&r^i V_x^i(x, t) - V_{tx}^i(x, t) \\
&= \frac{\partial}{\partial x} F^i_{\phi^{-i}}(x, \phi^i(x, t), t) + \frac{\partial}{\partial u^i} F^i_{\phi^{-i}}(x, \phi^i(x, t), t) \phi_x^i(x, t) \\
&+ f^i_{\phi^{-i}}(x, \phi^i(x, t), t) \left[\frac{\partial}{\partial x} G^i(x, t, \phi^i(x, t)) + \frac{\partial}{\partial \phi^i} G^i(x, t, \phi^i(x, t)) \phi_x^i(x, t) \right] \\
&+ G^i(x, t, \phi^i(x, t)) \left[\frac{\partial}{\partial x} f^i_{\phi^{-i}}(x, \phi^i(x, t), t) + \frac{\partial}{\partial u^i} f^i_{\phi^{-i}}(x, \phi^i(x, t), t) \phi_x^i(x, t) \right].
\end{aligned} \tag{4.13}$$

The term $V_x^i(x, t)$ on the left-hand side of (4.13) can be replaced by $G^i(x, t, \phi^i(x, t))$, because of (4.12). Similarly, by differentiating (4.12) with respect to t we obtain

$$V_{xt}^i(x, t) = V_{tx}^i(x, t) = \frac{\partial}{\partial t} G^i(x, t, \phi^i(x, t)) + \frac{\partial}{\partial \phi^i} G^i(x, t, \phi^i(x, t)) \phi_t^i(x, t),$$

which can be substituted for $V_{tx}^i(x, t)$ on the left-hand side of (4.13). Putting all this together we obtain

$$\begin{aligned}
&r^i G^i(x, t, \phi^i(x, t)) - \frac{\partial}{\partial t} G^i(x, t, \phi^i(x, t)) - \frac{\partial}{\partial \phi^i} G^i(x, t, \phi^i(x, t)) \phi_t^i(x, t) \\
&= \frac{\partial}{\partial x} F^i_{\phi^{-i}}(x, \phi^i(x, t), t) + \frac{\partial}{\partial u^i} F^i_{\phi^{-i}}(x, \phi^i(x, t), t) \phi_x^i(x, t) \\
&+ f^i_{\phi^{-i}}(x, \phi^i(x, t), t) \left[\frac{\partial}{\partial x} G^i(x, t, \phi^i(x, t)) + \frac{\partial}{\partial \phi^i} G^i(x, t, \phi^i(x, t)) \phi_x^i(x, t) \right] \\
&+ G^i(x, t, \phi^i(x, t)) \left[\frac{\partial}{\partial x} f^i_{\phi^{-i}}(x, \phi^i(x, t), t) + \frac{\partial}{\partial u^i} f^i_{\phi^{-i}}(x, \phi^i(x, t), t) \phi_x^i(x, t) \right].
\end{aligned} \tag{4.14}$$

Note that (4.14) is linear in the partial derivatives $\phi_x^i(x, t)$, and $\phi_t^i(x, t)$, which can make the equation easier to analyse than the equation for V^i

(4.11). A boundary condition for the partial differential equation (4.14) can be derived from the transversality condition. For example, in the finite horizon case we can differentiate with respect to x in condition (iii) of theorem 4.1, which, together with (4.12), implies

$$G^i(x, T, \phi^i(x, T)) = V^i_x(x, T) = S^i_x(x). \tag{4.15}$$

Let us point out once more that the derivation of (4.14) and (4.15) relies on many implicit smoothness assumptions as well as on the assumption of unique solvability of the maximization problem on the right-hand side of the HJB equation. All these implicit assumptions have to be checked if using this approach. Once a solution for ϕ^i has been found from (4.14) and (4.15), one can try to recover a solution for V^i from (4.12) and the transversality condition (iii) of theorem 4.1.[4]

Our next remark concerns games in which the time horizon is $T = \infty$ and the functions F^i, f, and U^i do not explicitly depend on the time variable t. Such a game is labelled autonomous. Since the fundamentals of this game do not change over time, and since at time t the remaining decision interval $[t, \infty)$ has the same (infinite) length no matter at which time t we look at the system, it is reasonable to consider stationary Markovian Nash equilibria. These are Nash equilibria $(\phi^1, \phi^2, \ldots, \phi^N)$ where the strategies are time independent functions $\phi^i : X \mapsto U^i$, $i \in \{1, 2, \ldots, N\}$. For such an equilibrium, the value functions V^i must also be time independent, that is, $V^i : X \mapsto \mathbb{R}$. The term $-V^i_t(x, t)$ on the left-hand side of the HJB equation (4.10) is therefore equal to 0 and the structure of the equation becomes somewhat simpler. If the state space X is one-dimensional in an autonomous game, the HJB equation is an ordinary differential equation. Solutions of such equations are usually much easier to find than solutions of partial differential equations. It is common practice to restrict the analysis of an autonomous game defined on the infinite time interval $[0, \infty)$ to stationary Markovian Nash equilibria as defined above. It has to be emphasized, however, that such a game may have nonstationary Markovian Nash equilibria as well. An example will be given in exercise 5 at the end of the chapter. This concludes our discussion of equilibrium conditions based on the HJB equation.

We next consider equilibrium conditions derived from Pontryagin's maximum principle. Our main point here is that the conditions of theorem 4.2 are considerably simplified when $(\phi^1, \phi^2, \ldots, \phi^N)$ are open-loop strategies and the feasible control sets $U^i(x, u^{-i}, t)$, $i = 1, 2, \ldots, N$,

[4]Exercise 4 at the end of the chapter illustrates the approach. See the references mentioned in section 4.4 for further examples of its application.

do not depend on the state x. To see this, first note that under these assumptions the feasible control sets $U^i_{\phi^{-i}}(x,t)$ of problem (4.1) are also independent of x for all $i \in \{1,2,\dots,N\}$. Therefore, we can replace the maximized Hamiltonian function in the adjoint equation from condition (ii) in theorem 4.2 with the ordinary Hamiltonian function (see the last statement of theorem 3.2), which yields

$$\dot{\lambda}^i(t) = r^i\lambda^i(t) - \frac{\partial}{\partial x}H^i_{\phi^{-i}}(x(t), \phi^i(t), \lambda^i(t), t).$$

Using the definition of the function $H^i_{\phi^{-i}}$, equation (4.2), and the fact that open-loop strategies ϕ^j do not depend on x for all $j \in \{1,2,\dots,N\}$, the above equation can be rewritten as

$$\begin{aligned}\dot{\lambda}^i(t) = r^i\lambda^i(t) &- F^i_x(x(t), \phi^1(t), \phi^2(t), \dots, \phi^N(t), t)\\ &- \lambda^i(t)f_x(x(t), \phi^1(t), \phi^2(t), \dots, \phi^N(t), t).\end{aligned}$$

Together with the transversality condition from condition (iii) (or (iv)) of theorem 4.2 and the state equation

$$\dot{x}(t) = f(x(t), \phi^1(t), \phi^2(t), \dots, \phi^N(t), t), \quad x(0) = x_0$$

this constitutes a two-point boundary value problem which at least can be solved by numerical techniques. In principle, the equilibrium conditions for general Markovian Nash equilibria are also a two-point boundary value problem. The structure of that problem, however, is more complicated because it also involves the partial derivatives $\phi^j_x(x(t),t)$, $j \in \{1,2,\dots,N\}$.

We conclude this discussion of theorems 4.1 and 4.2 by recalling that in both cases we have made smoothness assumptions for the functions V^i, H^{i*}, and S^i, respectively. It is, however, possible to use the weaker optimality conditions presented in section 3.7 to formulate equilibrium conditions for differential games.

4.3 Time consistency and subgame perfectness

The discussion in section 4.1 showed that, within the framework of differential games, Nash equilibria can be defined in many different ways, depending on the assumptions imposed on the information available to the players. We have emphasized that the informational assumptions should be made on the same grounds as other model assumptions. Solving a game, for instance, first for an open-loop Nash equilibrium and then for a nondegenerate Markovian Nash equilibrium should not be regarded as finding two solutions of a single model but rather as finding the solutions of two different models. Which of the two models is the better depends on the specific situation that one wants to depict, and there is no reason to believe that one type of equilibrium is generally better than another one. Nevertheless, it is worthwhile to evaluate different informational assumptions on the basis of their properties and implications. In the present section, we discuss two important properties: time consistency and subgame perfectness. We continue to restrict ourselves to Markovian strategies.

Let us start by introducing some additional notation. Denote by $\Gamma(x_0, 0)$ the game studied in section 4.1 where player $i \in \{1, 2, \ldots, N\}$ seeks to maximize the objective functional

$$\int_0^T e^{-r^i t} F^i(x(t), u^1(t), u^2(t), \ldots, u^N(t), t)\, dt + e^{-r^i T} S^i(x(T))$$

subject to the constraints $u^i(t) \in U^i(x(t), u^{-i}(t), t)$ and the system dynamics

$$\dot{x}(t) = f\big(x(t), u^1(t), u^2(t), \ldots, u^N(t), t\big), \qquad x(0) = x_0.$$

For each pair $(x, t) \in X \times [0, T\rangle$ we define a subgame $\Gamma(x, t)$ by replacing the objective functional for player i and the system dynamics by

$$\int_t^T e^{-r^i(s-t)} F^i(x(s), u^1(s), u^2(s), \ldots, u^N(s), s)\, ds + e^{-r^i(T-t)} S^i(x(T))$$

and

$$\dot{x}(s) = f\big(x(s), u^1(s), u^2(s), \ldots, u^N(s), s\big), \qquad x(t) = x,$$

respectively. Hence $\Gamma(x, t)$ is a differential game defined on the time interval $[t, T\rangle$ with initial condition $x(t) = x$.

Definition 4.3 Let $(\phi^1, \phi^2, \ldots, \phi^N)$ be a Markovian Nash equilibrium for the game $\Gamma(x_0, 0)$ and denote by $x(\cdot)$ the unique state trajectory generated in this equilibrium. We call the equilibrium time consistent if, for each $t \in [0, T\rangle$, the subgame $\Gamma(x(t), t)$ admits a Markovian Nash equilibrium $(\psi^1, \psi^2, \ldots, \psi^N)$ such that $\psi^i(y, s) = \phi^i(y, s)$ holds for all $i \in \{1, 2, \ldots, N\}$ and all $(y, s) \in X \times [t, T\rangle$.

To interpret the property of time consistency, assume that the N players are playing the Markovian Nash equilibrium $(\phi^1, \phi^2, \ldots, \phi^N)$ for the game $\Gamma(x_0, 0)$ and recall that ϕ^i is a mapping from $X \times [0, T\rangle$ to $\mathbb{R}^{m^i}$. This means that each player $i \in \{1, 2, \ldots, N\}$ determines his action $u^i(s)$ at time $s \in [0, T\rangle$ according to the rule $u^i(s) = \phi^i(x(s), s)$ where $x(s)$ is the state variable he observes at time s. Although the control value $u^i(s)$ may (and usually does) change as time s evolves, the rule ϕ^i does not change. Now assume that at some time $t \in [0, T\rangle$ the players are allowed to reconsider their strategy choices. At that time the clock shows t and the current state is $x(t)$, so that the N players find themselves at the start of the subgame $\Gamma(x(t), t)$. If the restriction of the Nash equilibrium $(\phi^1, \phi^2, \dots, \phi^N)$ of $\Gamma(x_0, 0)$ to the relevant domain $X \times [t, T\rangle$ constitutes a Markovian Nash equilibrium for the subgame $\Gamma(x(t), t)$, then the original equilibrium is time consistent. In other words, a Markovian Nash equilibrium is time consistent if it is also a Markovian Nash equilibrium of every subgame along the original equilibrium trajectory $x(\cdot)$. The following result is true.

Theorem 4.3 *Every Markovian Nash equilibrium of a differential game is time consistent.*

Proof Let $(\phi^1, \phi^2, \dots, \phi^N)$ be a Markovian Nash equilibrium of $\Gamma(x_0, 0)$ and denote by $x(\cdot)$ the state trajectory and by $u^1(\cdot)$, $u^2(\cdot)$, ..., $u^N(\cdot)$ the control paths of the N players corresponding to these strategies. If the equilibrium is not time consistent, then there must exist a time $t \in [0, T\rangle$ and a player $i \in \{1, 2, \dots, N\}$ such that player i can improve his objective functional in the subgame $\Gamma(x(t), t)$, given that his opponents stick to their strategies ϕ^{-i}. Assume that $\tilde{\phi}^i$ is a Markovian strategy which leads to such an improvement and consider the strategy $\pi^i : X \times [0, t\rangle \mapsto \mathbb{R}^{m^i}$ defined as follows:

$$\pi^i(x, s) = \begin{cases} \phi^i(x, s) & \text{if } s \in [0, t), \\ \tilde{\phi}^i(x, s) & \text{if } s \in [t, T\rangle. \end{cases}$$

Obviously, this is also a Markovian strategy. Denote the state trajectory and the control paths generated by the strategies $(\phi^1, \phi^2, \dots, \phi^{i-1}, \pi^i, \phi^{i+1}, \dots, \phi^N)$ by $y(\cdot)$ and $v^j(\cdot), j \in \{1, 2, \dots, N\}$. It is clear that $x(s) = y(s)$ and $v^j(s) = u^j(s)$ holds for all $j \in \{1, 2, \dots, N\}$ and all $s \in [0, t)$. Therefore, we have

$$J^i(v^i(\cdot)) = \int_0^t e^{-r^i s} F^i(x(s), u^1(s), u^2(s), \dots, u^N(s), s)\, ds \\ + \int_t^T e^{-r^i s} F^i(y(s), v^1(s), v^2(s), \dots, v^N(s), s)\, ds + e^{-r^i T} S^i(y(T)). \tag{4.16}$$

Moreover, because $\tilde{\phi}^i$ was assumed to be a better strategy than ϕ^i for player i in the game $\Gamma(x(t), t)$ it must hold that

$$\int_t^T e^{-r^i s} F^i(y(s), v^1(s), v^2(s), \ldots, v^N(s), s)\, ds + e^{-r^i T} S^i(y(T))$$
$$> \int_t^T e^{-r^i s} F^i(x(s), u^1(s), u^2(s), \ldots, u^N(s), s)\, ds + e^{-r^i T} S^i(x(T)).$$

This inequality, together with (4.16), implies that $J^i(v^i(\cdot)) > J^i(u^i(\cdot))$, which is a contradiction to the assumption that $(\phi^1, \phi^2, \ldots, \phi^N)$ is a Markovian Nash equilibrium of the game $\Gamma(x_0, 0)$. The contradiction proves the claim. ■

Time consistency could be seen as a minimal requirement for the credibility of an equilibrium strategy. If player i had an incentive to deviate from his strategy ϕ^i during the time interval $[0, T\rangle$ then the other players would not believe his announcement of ϕ^i in the first place. Consequently, they would compute their own strategies by taking into account the expected future deviation of player i which, in general, would lead to strategies different from $\phi^j, j \neq i$. Theorem 4.3 ensures that such a situation cannot occur as long as we restrict ourselves to Markovian Nash equilibria.[5]

In chapter 5 we shall encounter examples of equilibria which fail to be time consistent.

In order better to understand the issue of credibility, and to motivate the notion of subgame perfectness in differential games, consider the following, extremely simple, example.

Example 4.2 There are N identical players who try to maximize the objective functional

$$J^i(u^i(\cdot)) = -\int_0^T (u^i(t))^2\, dt - x(T)^2,$$

$i \in \{1, 2, \ldots, N\}$, subject to the system dynamics $\dot{x}(t) = \sum_{j=1}^N u^j(t)$ and the initial condition $x(0) = 0$. The time horizon T is finite. The state space is $X = \mathbb{R}$ and the feasible control sets are $U^i(x, t) = \mathbb{R}$ for all players $i \in \{1, 2, \ldots, N\}$ and all $(x, t) \in \mathbb{R} \times [0, T]$.

[5] See theorem 6.1 below for a generalization of this result to games with non-Markovian strategies.

First observe that $J^i(u^i(\cdot)) \leq 0$ holds for all feasible control paths so that any feasible path which leads to $J^i(u^i(\cdot)) = 0$ is automatically an optimal path. Now consider the N-tuple of strategies $(\phi^1, \phi^2, \ldots, \phi^N)$ where $\phi^j(x, t) = x$ for all $j \in \{1, 2, \ldots, N\}$. If all players $j \neq i$ use the strategies ϕ^j then player i faces the system dynamics $\dot{x}(t) = u^i(t) + (N-1)x(t)$. Since the initial state is $x_0 = 0$ it follows immediately that choosing the control path $u^i(t) = 0$ for all $t \in [0, T]$ leads to the state trajectory $x(t) = 0$ for all $t \in [0, T]$ and therefore to $J^i(u^i(\cdot)) = 0$. This implies that $u^i(t) = 0$ for all $t \in [0, T]$ is an optimal control path for player i's problem. Moreover, this path can be represented by the strategy $u^i(t) = \phi^i(x(t), t) = x(t)$ since the state trajectory is also identically equal to 0. We have therefore proved that $(\phi^1, \phi^2, \ldots, \phi^N)$ with $\phi^j(x, t) = x$ for all $j \in \{1, 2, \ldots, N\}$ is a Markovian Nash equilibrium.[6] We know from theorem 4.3 that this Nash equilibrium is time consistent so that no player $i \in \{1, 2, \ldots, N\}$ has an incentive to deviate from ϕ^i during the interval $[0, T]$. However, the equilibrium strategies ϕ^i are not credible in the following sense. If the state would satisfy $x(t) \neq 0$ at some time instant t (in the Markovian Nash equilibrium described above, this does not happen!), then all players sticking to ϕ^i would have to choose nonzero controls $u^i(t) = \phi^i(x(t), t) = x(t)$, which would drive the state even further away from 0. This is not in the best interest of any player. For example, it would be better for each player to choose $u^i(t) = 0$ to avoid the cost associated with a nonzero control value and, in addition, to reduce the speed at which the state diverges from 0. We summarize by saying that although the actions induced by the strategies ϕ^i are credible along the equilibrium trajectory $x(\cdot)$ they are not credible as specifications of optimal behaviour off the equilibrium path.

The above example suggests a stronger, and probably more useful credibility criterion than time consistency, namely, that the strategies ϕ^i, $i \in \{1, 2, \ldots, N\}$ should represent optimal behaviour not only along the equilibrium state trajectory but also off this trajectory. This can be formalized as follows.

Definition 4.4 Let $(\phi^1, \phi^2, \ldots, \phi^N)$ be a Markovian Nash equilibrium for the game $\Gamma(x_0, 0)$. We call the equilibrium subgame perfect if, for each $(x, t) \in X \times [0, T)$, the subgame $\Gamma(x, t)$ admits a Markovian Nash equilibrium $(\psi^1, \psi^2, \ldots, \psi^N)$ such that $\psi^i(y, s) = \phi^i(y, s)$ holds for all $i \in \{1, 2,$

[6] The argument can easily be generalized to show that any N-tuple of strategies $(\phi^1, \phi^2, \ldots, \phi^N)$, such that $\phi^j(0, t) = 0$ holds for all $j \in \{1, 2, \ldots, N\}$ and for all $t \in [0, T]$, is a Nash equilibrium.

$\ldots, N\}$ and all $(y, s) \in X \times [t, T\rangle$. A Markovian Nash equilibrium which is subgame perfect is also called a Markov perfect Nash equilibrium.

In contrast to time consistency, subgame perfectness not only requires that the restriction of ϕ to $X \times [t, T\rangle$ is a Markovian Nash equilibrium of the subgame $\Gamma(x(t), t)$, where $x(\cdot)$ is the state trajectory generated by the equilibrium strategies, but that it is a Markovian Nash equilibrium for all subgames $\Gamma(x, t)$, $(x, t) \in X \times [0, T\rangle$. Of course, subgame perfectness of a Markovian Nash equilibrium implies its time consistency.

The following result shows that a slight strengthening of the conditions of theorem 4.1 guarantees not only the Nash equilibrium property but also the subgame perfectness of the equilibrium.

Theorem 4.4 *Let $(\phi^1, \phi^2, \ldots, \phi^N)$ be a given N-tuple of functions $\phi^i : X \times [0, T\rangle \mapsto \mathbb{R}^{m^i}$ and make the following assumptions:*

(i) for every pair $(y, s) \in X \times [0, T\rangle$ there exists a unique absolutely continuous solution $x_{y,s} : [s, T\rangle \mapsto X$ of the initial value problem

$$\dot{x}(t) = f(x(t), \phi^1(x(t), t), \phi^2(x(t), t), \ldots, \phi^N(x(t), t), t), \quad x(s) = y,$$

(ii) conditions (ii)–(iv) of theorem 4.1 are satisfied with the additional requirement that in (iv), when V^i is not bounded above, $\limsup_{t\to\infty} e^{-r^i t} V^i(x_{y,s}(t), t) \leq 0$ must hold for all $(y, s) \in X \times [0, T\rangle$ (here $x_{y,s}(\cdot)$ is defined in condition (i) above).

Denote by $\Phi^i(x, t)$ the set of all $u^i \in U^i_{\phi^{-i}}(x, t)$ which maximize the right-hand side of (4.10). If $\phi^i(x, t) \in \Phi^i(x, t)$ holds for all $i \in \{1, 2, \ldots, N\}$ and all $(x, t) \in X \times [0, T\rangle$ then $(\phi^1, \phi^2, \ldots, \phi^N)$ is a Markov perfect Nash equilibrium. (If $T = \infty$, optimality is understood in the sense of catching up optimality.)

Proof The proof is a straightforward application of theorem 3.1 in the finite horizon case, and theorem 3.3 and lemma 3.1 in the infinite horizon case. This time, however, we apply these results not only to the individual optimal control problems of the game $\Gamma(x_0, 0)$ but also to the individual optimal control problems of $\Gamma(x, t)$ for all $(x, t) \in X \times [0, T\rangle$. ■

The main difference between the assumptions of theorem 4.1 and theorem 4.4 is that in the latter, the condition $\phi^i(x, t) \in \Phi^i(x, t)$ has to be satisfied for all $(x, t) \in X \times [0, T\rangle$ while in the former it was only required to hold for all (x, t) satisfying $x = x(t)$. Let us now reconsider example 4.2 and try to find a Markov perfect Nash equilibrium.

Example 4.2 (continued) Since the game is symmetric with respect to all players, we look for a symmetric equilibrium $(\phi^1, \phi^2, \ldots, \phi^N)$ in which $\phi^i = \phi^j = \phi$. This leads to identical optimal value functions so we may simply write V instead of V^i. Let us start by assuming that V is continuously differentiable so that the HJB equation (4.10) can be written as

$$-V_t(x,t) = \max\{-(u^i)^2 + V_x(x,t)[(N-1)\phi(x,t) + u^i] \mid u^i \in \mathbb{R}\} \tag{4.17}$$

for all $(x, t) \in \mathbb{R} \times [0, T]$. The transversality condition requires $V(x, T) = -x^2$ for all $x \in \mathbb{R}$. Since the model is of the linear quadratic variety (see section 7.1) we try to find a solution of the above partial differential equation which has the form $V(x, t) = A(t) + B(t)x + C(t)x^2$. The transversality condition implies

$$A(T) = B(T) = 0 \qquad \text{and} \qquad C(T) = -1. \tag{4.18}$$

Substituting the quadratic form of $V(x, t)$ into (4.17) we obtain

$$-\dot{A}(t) - \dot{B}(t)x - \dot{C}(t)x^2 = \max\{-(u^i)^2 + [B(t) + 2C(t)x]$$
$$[(N-1)\phi(x,t) + u^i] \mid u^i \in \mathbb{R}\}.$$

The unique maximum on the right-hand side of this equation is attained at $u^i = \phi(x, t) = B(t)/2 + C(t)x$. Therefore, the above HJB equation can be written as

$$-\dot{A}(t) - \dot{B}(t)x - \dot{C}(t)x^2 = (2N-1)[B(t)^2/4 + B(t)C(t)x + C(t)^2x^2],$$

which can hold for all $x \in \mathbb{R}$ only if the differential equations $\dot{A}(t) = -(2N-1)B(t)^2/4$, $\dot{B}(t) = -(2N-1)B(t)C(t)$, and $\dot{C}(t) = -(2N-1)C(t)^2$ are satisfied for all $t \in [0, T]$. Together with the terminal conditions from (4.18) this implies $A(t) = B(t) = 0$ and $C(t) = [(2N-1)(t-T) - 1]^{-1}$. Therefore, we obtain the candidate $\phi(x, t) = B(t)/2 + C(t)x = x/[(2N-1)(t-T) - 1]$. The conditions of theorem 4.4 are satisfied with $V^i(x, t) = V(x, t) = x^2/[(2N-1)(t-T) - 1]$ so that this candidate is indeed a Markov perfect Nash equilibrium.

An important remark has to be made at this point. The notion of subgame perfectness was originally introduced into game theory for games in extensive form in which the exact timing of the actions by all players, and the information that is available to each player at each instant of time, are completely specified (see chapter 2). It depends crucially on the information structure which equilibria of an extensive form game are subgame perfect. In general, one can say that a coarser information structure allows more subgame perfect equilibria. The situation is exactly the same in differential games. Note that we have not considered an extensive form representation for differential games but we have specified one particular information structure through the family of possible subgames $\{\Gamma(x, t) \mid (x, t) \in X \times [0, T\rangle\}$. This specification is the most relevant for the applications of differential game theory discussed in this book but it is worth emphasizing that other information structures, and hence other families of subgames, and hence other subgame perfect equilibria, could also be plausible; see chapter 6 for a discussion of such information structures.

In particular, we mention a slightly weaker form of subgame perfectness than the one introduced in definition 4.4. Recall that we motivated subgame perfectness by the argument that an equilibrium strategy ϕ^i should describe a credible rule to determine the actions of player i not only along the equilibrium path $x(\cdot)$ but also for all other pairs $(x, t) \in X \times [0, T\rangle$. This argument is not convincing if there are pairs $(x, t) \in X \times [0, T\rangle$ which cannot be reached along a feasible state trajectory starting from the initial condition $x(0) = x_0$. For example, if the state dynamics are $\dot{x}(t) = \sum_{i=1}^{N} u^i(t)$, as in example 4.2, and the feasible control sets are given by $U^i(x, t) = [0, 1]$, then every feasible state trajectory $x(\cdot)$ must satisfy $0 \leq x(t) \leq Nt$. If the strategy ϕ^i generates controls which are not credible only for pairs (x, t) with $x < 0$ or $x > Nt$, then the equilibrium might still be considered as credible because those pairs (x, t) cannot be reached anyway. To formalize this idea, define the set $A \subseteq X \times [0, T\rangle$ of attainable pairs by the following condition: $(x, t) \in A$ if and only if there exists a feasible state trajectory $x(\cdot)$ such that $(x(t), t) = (x, t)$. The weaker form of subgame perfectness is then defined as follows.

Definition 4.5 Let $(\phi^1, \phi^2, \ldots, \phi^N)$ be a Markovian Nash equilibrium for the game $\Gamma(x_0, 0)$. The equilibrium is said to be weakly subgame perfect if for each pair $(x, t) \in X \times [0, T\rangle \cap A$, the subgame $\Gamma(x, t)$ admits a Markovian Nash equilibrium $(\psi^1, \psi^2, \ldots, \psi^N)$ such that $\psi^i(y, s) = \phi^i(y, s)$ holds for all $i \in \{1, 2, \ldots, N\}$ and all $(y, s) \in X \times [t, T\rangle$.

We close this section with a brief remark concerning subgame perfection of stationary Markovian Nash equilibria in autonomous games defined on the time interval $[0, \infty)$. Since for such a game the subgame $\Gamma(x, t)$ is equivalent (in fact, identical) to $\Gamma(x, 0)$, it follows from the definition of subgame perfectness that any stationary Markovian Nash equilibrium is subgame perfect, provided it is independent of the initial state x_0. If the game is nonautonomous, this conclusion may not hold (see exercise 6).

4.4 Further reading

Markovian Nash equilibria are those equilibria which are most often studied in the differential games literature. Related material appears in all textbooks on differential games, including Başar and Olsder [4], Case [16], Mehlmann [175, 176], and Petit [190]. Some books on economic and management science applications of optimal control theory, like Feichtinger and Hartl [85], Kamien and Schwartz [148], or Sethi and Thompson [216] have chapters on differential games in which the Markovian Nash equilibrium concept is discussed. Other references are Clemhout and Wan [32] and chapter 13 of the game theory book by Fudenberg and Tirole [105].

There is a point we should note concerning the concept of Nash equilibrium. The alert reader will have noted that the situation dealt with in this chapter is somewhat more complicated than the normal form games discussed in chapter 2. The reason is primarily that we allow state constraints and control constraints which depend on the opponents' actions. This implies that the set of feasible actions for player i at a certain time t depends, in general, on the strategies of all opponents. Such a model is usually called a generalized game in the sense of Debreu [40]. One particular problem that arises in this framework will be discussed in section 12.1.2 below: the distinction between weakly and strictly feasible replies.

The method that we described in section 4.2 whereby the HJB equation is differentiated with respect to the state variable in order to get an auxiliary differential equation which is sometimes easier to analyse than the HJB equation itself has been used by several authors, e.g., Dockner and Long [56], Sorger [222], Tsutsui and Mino [233].

The concept of time consistency has received considerable attention in the macroeconomic literature, e.g., Kydland and Prescott [155]. A good discussion of the subject and its relation to dynamic game theory can be found in chapter 3 of Fudenberg and Tirole [105], in Petit [190], or in Pohjola [192]. Subgame perfectness was introduced into the game theory literature by Selten [214]. As already emphasized in section 4.3, subgame perfectness is usually discussed in the framework of extensive form games. The concept of Markov perfectness is used to capture essentially the same idea as subgame perfectness without explicitly using an extensive form of the differential game. Instead, a family of subgames $\Gamma(x, t)$ is postulated directly. Nevertheless, the term subgame perfectness is also often used as a synonym of Markov perfectness. Some authors use still another terminology, namely weak and strong time consistency, for what we called time consistency and Markov perfectness, respectively (see Başar and Olsder [4]).

We have seen that the use of Markov strategies allows players to be equipped with different periods of commitment. While in case of an open-loop equilibrium the period of commitment coincides with the duration of the game, it vanishes in case of a Markov perfect equilibrium. Reinganum and Stokey [202] analyse the relationship between commitment and the use of various Markov strategies.

4.5 Exercises

1. Consider the following two-player differential game with state space $X = \mathbb{R}$ and finite time horizon T. Both players have the same time preference rate $r = 0$, the utility functions are given by $F^1(x, u^1, u^2, t) = x - u^1u^2$ and $F^2(x, u^1, u^2, t) = (1 + u^1)x - u^2$, and the scrap value functions S^1 and S^2 are identically equal to 0. Finally, the system dynamics are given by $f(x, u^1, u^2, t) = u^1 + u^2$. Each player i can choose his controls in the fixed set $U^i = [0, 1]$. Derive an open-loop Nash equilibrium.
2. Consider a symmetric N-player differential game with finite time horizon T and state space $X = [0, \infty)$. Player i's utility function is $F^i(x, u^1, u^2, \ldots, u^N, t) = -x - (u^i)^2/2$ and the system dynamics are $f(x, u^1, u^2, \ldots, u^N, t) = -\sqrt{x}\left(\sum_{i=1}^N u^i\right)$. The scrap value functions are identically equal to 0 and all players have the same discount rate r and the same feasible control set $U^i = [0, \infty)$. Prove that $(\phi^1, \phi^2, \ldots, \phi^N)$ is a nondegenerate Markovian Nash equilibrium, where $\phi^i(t, x) = A(t)\sqrt{x}$ and $A : [0, T] \mapsto \mathbb{R}$ is the unique solution to the boundary value problem

$$\dot{A}(t) = (2N - 1)A(t)^2/2 + rA(t) - 1\,, \quad A(T) = 0.$$

Furthermore, show that $(\psi^1, \psi^2, \ldots, \psi^N)$ is an open-loop Nash equilibrium where $\psi^i(t) = B(t)\sqrt{x(t)}$, $x(t) = x_0 \exp\left\{-N\int_0^t B(s)\,ds\right\}$, and $B : [0, T] \mapsto \mathbb{R}$ is the unique solution to the boundary value problem

$$\dot{B}(t) = NB(t)^2/2 + rB(t) - 1\,, \quad B(T) = 0.$$

3. Consider a multi-player version of exercise 2 in section 3.9. There are N identical firms extracting a common nonrenewable resource. We denote by $x(t)$ the remaining resource stock by time t, by $p > 0$ the constant market price of one unit of the resource, and by $u^i(t)$ the extraction rate (sales rate) of firm $i \in \{1, 2, \ldots, N\}$ at time t. Assuming

that all players have the same cost function $c(u^i, x) = \gamma(u^i)^2/(2x)$ and the same discount rate r, the game is given by the objective functionals

$$J^i(u^i(\cdot)) = \int_0^T e^{-rt}[pu^i(t) - c(u^i(t), x(t))]\, dt + e^{-rT} qx(T)$$

and the system dynamics

$$\dot{x}(t) = -\sum_{j=1}^N u^j(t).$$

Find a Markov perfect Nash equilibrium for this game under the assumption $q \in [0, p]$.

4. Consider the symmetric N-player differential game with time horizon T, discount rates $r^i = r$, utility functions $F^i(x, u^1, u^2, \ldots, u^N, t) = \sqrt{u^i}$, scrap value functions $S^i(x) = 0$, and system dynamics $f(x, u^1, u^2, \ldots, u^N, t) = K(x) - \sum_{i=1}^N u^i$. Assume that there exists a symmetric Markovian Nash equilibrium $(\phi^1, \phi^2, \ldots, \phi^N)$ with $\phi^i = \phi^j$ for all $i, j \in \{1, 2, \ldots, N\}$. Derive equations (4.11) and (4.14) for this example.
5. Consider the symmetric two-player, infinite horizon differential game with $X = \mathbb{R}$, $F^i(x, u^1, u^2, t) = x - \alpha u^i$, $f(x, u^1, u^2, t) = u^1 u^2$, and $U^i = [0, 1]$. Assume that the common discount rate satisfies $0 < r < 1/\alpha$. Prove that (ϕ^1, ϕ^2) with $\phi^i(x, t) = 1$ is a stationary Markovian Nash equilibrium. Furthermore, show that (ψ^1, ψ^2) with

$$\psi^i(x, t) = \begin{cases} 0 & \text{if } t < \tau, \\ 1 & \text{if } t \geq \tau, \end{cases}$$

is a nonstationary Markovian Nash equilibrium for any $\tau > 0$. Can you think of other nonstationary Markovian Nash equilibria? Which of these equilibria are subgame perfect?
6. Consider the symmetric two-player infinite horizon differential game with $X = [0, \infty)$, $F^i(x, u^1, u^2, t) = x - u^i e^{2t-x}$, $f(x, u^1, u^2, t) = u^1 + u^2$, and $U^i = [0, 1]$. Assume that the common discount rate satisfies $0 < r < 1$. Prove that (ϕ^1, ϕ^2) with $\phi^i(x, t) = 1$ is a Markovian Nash equilibrium independently of the value of the initial state x_0. Show that this equilibrium is not subgame perfect.

5 Differential games with hierarchical play

The preceding chapter dealt with differential games in which all players make their moves simultaneously. We now turn to a class of differential games in which some players have priority of moves over other players. To simplify matters, we focus mostly on the case where there are only two players. The player who has the right to move first is called the leader and the other player is called the follower. A well-known example of this type of hierarchical-moves games is the Stackelberg model of duopoly, which is often contrasted with the Cournot model of duopoly.

The plan of this chapter is as follows. In section 5.1 we review the one-shot Cournot duopoly game and the corresponding one-shot Stackelberg game. We also present a modified version of the one-shot Stackelberg game as a quick means of raising the issue of time inconsistency (sometimes referred to as dynamic inconsistency) in Stackelberg games, which we further expound in the rest of the chapter.

In sections 5.2 and 5.3 we define the concepts of open-loop Stackelberg equilibrium and (nondegenerate) Markovian Stackelberg equilibrium for differential games. We show in section 5.2 that open-loop Stackelberg equilibria are, in general, not time consistent. There are, of course, exceptions to this rule which we also consider. In section 5.3 we turn to the analysis of nondegenerate Markovian Stackelberg equilibria. In general, it is difficult to find such equilibria. However, we are able to provide some rules of thumb which work in a number of situations.

5.1 Cournot and Stackelberg equilibria in one-shot games

In this section we review in more detail the one-shot Cournot duopoly model and the associated Stackelberg leadership model, both of which were introduced in chapter 2.

5.1.1 The one-shot Cournot duopoly game

Consider two firms that produce a homogeneous good. Their outputs are $Q_1 \geq 0$ and $Q_2 \geq 0$, respectively. Let $Q = Q_1 + Q_2$ denote industry output. It is assumed, for simplicity, that the market price is given by

$$P(Q) = \begin{cases} 1 - Q & \text{if } Q \in [0, 1], \\ 0 & \text{if } Q \geq 1. \end{cases}$$

Firm i's unit cost is denoted by c_i, where $0 \leq c_i < 1$ for $i = 1, 2$. In the Cournot game, each firm i must choose its output Q_i without knowing the output Q_j which is chosen by its rival $j \neq i$. Both firms must act simultaneously, and collusion is not allowed. The game is played only once, i.e., this is not a dynamic game. Given Q_j, firm i's profit is $\pi_i(Q_i, Q_j) = [P(Q_i + Q_j) - c_i]Q_i$. A Cournot equilibrium is a pair (Q_i^*, Q_j^*) such that

$$\pi_1(Q_1^*, Q_2^*) \geq \pi_1(Q_1, Q_2^*) \qquad \text{for all } Q_1 \geq 0$$

and

$$\pi_2(Q_2^*, Q_1^*) \geq \pi_2(Q_2, Q_1^*) \qquad \text{for all } Q_2 \geq 0.$$

It is easy to verify that a Cournot equilibrium (Q_1^*, Q_2^*) exists and is unique (see below). One may argue that the pair (Q_1^*, Q_2^*) is a good prediction of the outcome of the game. Thus, if both firms were to consult a game theorist, she would recommend firm 1 (firm 2) to produce Q_1^* (Q_2^*) on the ground that it should expect its opponent, firm 2 (firm 1), to produce Q_2^* (Q_1^*). Each firm's equilibrium output is its unique best choice, given that the other firm is expected to produce its equilibrium output.

To find the Cournot equilibrium for the game described above, it is convenient to define firm j's best reply function, or reaction function, $R_j(\cdot)$. For any $Q_i \geq 0$ it specifies the output level $R_j(Q_i)$ that maximizes firm j's profit function $Q_j \mapsto \pi_j(Q_j, Q_i)$. It is easy to see that

$$R_j(Q_i) = \max\{0, (1 - Q_i - c_j)/2\}.$$

The Cournot equilibrium is determined as the unique intersection of the graphs of the best reply functions $R_1(\cdot)$ and $R_2(\cdot)$. It can be shown that both Q_1^* and Q_2^* are positive if

$$1 > 2c_1 - c_2 \quad \text{and} \quad 1 > 2c_2 - c_1. \tag{5.1}$$

In what follows, we assume that (5.1) holds. The Cournot equilibrium outputs are then given by

$$Q_1^* = \frac{1 - 2c_1 + c_2}{3} \quad \text{and} \quad Q_2^* = \frac{1 - 2c_2 + c_1}{3}.$$

5.1.2 The one-shot Stackelberg duopoly game

Now let us modify the rules of the game. We require firm 2 to make its move before firm 1, while retaining the assumption that each firm produces only once. The first mover is said to be the leader and the second mover is called the follower. We offer two different interpretations of this modified game. The first one is that firm 2 produces first and firm 1 makes its output decision Q_1 after it observes the output Q_2 of firm 2. We assume that both outputs are sold in the same period and, thus, that the realized price is $P(Q_1 + Q_2)$. The second interpretation is that the leader is able to commit itself to an output level, Q_2, and this committed level is announced truthfully to the follower, who subsequently makes its output decision, believing that firm 2 will honour its commitment. We also assume that the leader must honour its commitment. Both firms carry out their production simultaneously. Again the realized price is $P(Q_1 + Q_2)$.

The second interpretation has the pedagogical advantage of allowing us to proceed further by adding a modification to the Stackelberg game and pose an important question about credibility and time inconsistency in the modified Stackelberg game. This will become clear in what follows.

Let us begin with an analysis of the Stackelberg equilibrium under the first interpretation. Since each firm makes only one move, this is still a one-shot game. Firm 2 (the leader) knows that firm 1, having observed Q_2, will choose $Q_1 = R_1(Q_2)$, where $R_1(\cdot)$ is the best reply function introduced in the preceding subsection; any output level different from $R_1(Q_2)$ would not maximize $\pi_1(\cdot, Q_2)$. In other words, among all possible decision rules $f_1(\cdot)$ that determine Q_1 as a function of the observed output level Q_2, the rule $R_1(\cdot)$ is firm 1's rational choice.

To find the output of the leader, we maximize the profit $P(R_1(Q_2) + Q_2)Q_2 - c_2Q_2$ over all $Q_2 \geq 0$. Given assumption (5.1), it is easy to see that the equilibrium output of the leader, Q_2^L, is

$$Q_2^L = \frac{1 + c_1 - 2c_2}{2} > 0,$$

and that the output of the follower, $Q_1^F = R_1(Q_2^L)$, is also positive. It is not difficult to verify that the leader's output is greater than what it would produce in a Cournot equilibrium ($Q_2^L > Q_2^*$) and that the follower's output is smaller than its output in the Cournot equilibrium ($Q_1^F < Q_1^*$). The profit of firm 2, when it is the Stackelberg leader, is greater than what it would obtain in a Cournot duopoly game with simultaneous play. It is always true that a Stackelberg leader can achieve a payoff that is at least as great as what it could achieve if the two players were to move simultaneously, because the leader can always choose its Cournot output.

To summarize, a Stackelberg equilibrium is a pair of outputs (Q_1^F, Q_2^L) such that $Q_1^F = R_1(Q_2^L)$ and $Q_2^L \in \text{argmax}\{\pi_2(Q_2, R_1(Q_2)) \mid Q_2 \geq 0\}$. This definition corresponds to the method of solving the game by backward induction, where a game tree is folded back towards its root.

It is important to note that we may also think of the Stackelberg equilibrium described above as a pair of strategies $(R_1(\cdot), Q_2^L)$ such that, given Q_2^L (which is firm 2's strategy), the rule $R_1(\cdot)$ maximizes firm 1's profit and, given that firm 1 uses this rule, the output Q_2^L maximizes firm 2's profit. In this sense, a Stackelberg equilibrium is a Nash equilibrium of a duopoly game where the strategy space for firm 1 is the set of functions $f_1(\cdot)$ mapping the set of nonnegative real numbers (firm 2's output) to the set of nonnegative real numbers (firm 1's output), and the strategy space for firm 2 is the set of nonnegative real numbers. This definition is appropriate when the game is presented in the normal form (see chapter 2).

Even though it is true that a Stackelberg equilibrium is a Nash equilibrium in a well defined sense, we avoid this terminology and instead follow the traditional usage by distinguishing Stackelberg equilibria in a game with hierarchical moves from Nash equilibria of a corresponding game with simultaneous moves.

5.1.3 Introduction to time inconsistency of Stackelberg equilibria

We now turn to the second interpretation of the Stackelberg duopoly game. Clearly, given that the leader must honour its commitment, the announcement of Q_2^L prior to the simultaneous production of the two outputs has the same force as the actual production of Q_2^L before the output decision by firm 1. Both interpretations give rise to the same equilibrium. However, with the second interpretation we can now modify the Stackelberg game by introducing a dynamic element in this one-shot game. The modification can be explained by considering the following situation. Suppose the leader has made the commitment to produce the output level Q_2^L and the follower believes this commitment, but when the time comes for both firms to produce simultaneously, the leader is released from its commitment without the follower's knowledge. Would the leader have an interest in producing less than Q_2^L? The answer is yes because, having induced the follower to make the decision to produce the output level Q_1^F, the leader would earn more profit by actually producing the amount $R_2(Q_1^F)$, which is below the committed quantity Q_2^L.[1] This modification serves to raise the issue of time inconsistency in Stackelberg games (see section 4.3, where time consistency is discussed for Markovian Nash equilibria).

The idea of time inconsistency has played a major role in the theory of economic policies. According to this theory, if an optimizing government makes a commitment to carry out a number of actions over a number of periods, and if other economic agents (consumers and producers) believe this commitment and choose their actions under this belief, then at some instant of time in the future the government would want to deviate from its commitment. Thus, government policies generated by solving an optimal control problem based on commitment are not time consistent. Furthermore, the theory asserts that, since consumers and producers are smart, they know of this time inconsistency property and, therefore, they will not believe any commitment made by a government that determines its policies in this way.

The issue of time inconsistency is obviously important. In those cases where a leader cannot credibly make a commitment, the concept of a Stackelberg equilibrium, which is based on commitment, seems inappropriate. In the following sections we discuss Stackelberg equilibria in the framework of differential games. We begin with the case where the players use open-loop strategies and then consider the more general case of Markovian strategies. We shall see that, in differential games, Stackelberg equilibria may be time consistent or not, depending on the model under consideration.

5.2 Open-loop Stackelberg equilibria

Consider a differential game with two players, which we refer to as L and F for leader and follower, respectively. To begin with, the time horizon, T, is fixed and finite. The case of an infinite horizon will be discussed later. Let x denote the vector of state variables, u^L the vector of control variables of the leader, and u^F the vector of control variables of the follower. Assume $x \in \mathbb{R}^n$, $u^L \in \mathbb{R}^{m^L}$, and $u^F \in \mathbb{R}^{m^F}$. The evolution of the state variables is given by

$$\dot{x}_i(t) = f_i(x(t), u^F(t), u^L(t), t), \qquad x_i(0) = x_{i0}, \qquad i = 1, 2, \ldots, n. \tag{5.2}$$

We assume that x_{i0} is given and that $x_i(T)$ is free.

[1]Of course, if the follower anticipates that the leader will be released from its commitment, then the follower will not believe that the leader will produce Q_2^L and in this case the only possible equilibrium is the Cournot equilibrium.

At time 0, the leader announces the control path $u^L(\cdot)$. The follower, taking this path as given, chooses his control path $u^F(\cdot)$ so as to maximize his integral of discounted net benefit, or utility,

$$J^F = \int_0^T e^{-r^F t} v^F(x(t), u^F(t), u^L(t), t)\, dt,$$

where $r^F > 0$ is the follower's rate of discount and v^F is his instantaneous utility function. Denoting by λ the vector of costate variables for this maximization problem, the follower's Hamiltonian is

$$H^F(x, u^F, \lambda, t) = v^F(x, u^F, u^L(t), t) + \sum_{i=1}^{n} \lambda_i f_i(x, u^F, u^L(t), t).$$

In what follows we assume for simplicity that either the controls are unconstrained or, if they are constrained (typically by nonnegativity constraints), assumptions have been made that rule out boundary solutions.

Given the time path $u^L(\cdot)$, the optimality conditions for the follower's problem are

$$\frac{\partial v^F(x(t), u^F(t), u^L(t), t)}{\partial u_j^F} + \sum_{k=1}^{n} \lambda_k \frac{\partial f_k(x(t), u^F(t), u^L(t), t)}{\partial u_j^F} = 0, \tag{5.3}$$

for $j = 1, 2, \ldots, m^F$, and

$$\dot{\lambda}_i(t) = r^F \lambda_i(t) - \frac{\partial v^F(x(t), u^F(t), u^L(t), t)}{\partial x_i} - \sum_{k=1}^{n} \lambda_k(t) \frac{\partial f_k(x(t), u^F(t), u^L(t), t)}{\partial x_i}, \tag{5.4}$$

for $i = 1, 2, \ldots, n$. Furthermore, since the time horizon T is finite and there is no salvage value, we have the transversality conditions

$$\lambda_i(T) = 0, \qquad i = 1, 2, \ldots, n. \tag{5.5}$$

Let us assume that the Hamiltonian H^F is jointly concave in the variables x and u^F. Then the above conditions are sufficient for the optimality of $u^F(\cdot)$. If H^F is strictly concave in u^F then condition (5.3) uniquely determines the value of each control variable $u_j^F(t)$ as a function of $x(t)$, $\lambda(t)$, $u^L(t)$, and t. That is, we can write

$$u_j^F(t) = g_j(x(t), \lambda(t), u^L(t), t), \qquad j = 1, \ldots, m^F, \tag{5.6}$$

or, in vector notation, $u^F(t) = g(x(t), \lambda(t), u^L(t), t)$. Substituting (5.6) into (5.4) we obtain

$$\dot{\lambda}_i(t) = r^F \lambda_i(t) - \frac{\partial v^F(x(t), g(x(t), \lambda(t), u^L(t), t), u^L(t), t)}{\partial x_i}$$
$$- \sum_{k=1}^{n} \lambda_k(t) \frac{\partial f_k(x(t), g(x(t), \lambda(t), u^L(t), t), u^L(t), t)}{\partial x_i}, \quad i = 1, 2, \ldots, n. \tag{5.7}$$

Equations (5.2) and (5.5)–(5.7) characterize the follower's best response to the leader's control path $u^L(\cdot)$. The leader, knowing the follower's best response to each path $u^L(\cdot)$, then proceeds to choose a path $u^{L*}(\cdot)$ so as to maximize the integral of her discounted utility.

Before presenting a technique of finding a solution to the leader's problem, we would like to draw the reader's attention to an important technical detail. Suppose that, for each time path $u^L(\cdot)$ that the leader announces, there is a corresponding best response by the follower that satisfies conditions (5.2) and (5.5)–(5.7). Denote the state and costate variables corresponding to this best response by $x^*(\cdot)$ and $\lambda^*(\cdot)$ and note that they satisfy the conditions $x_i^*(0) = x_{i0}$ and $\lambda_i^*(T) = 0$ for all $i = 1, 2, \ldots, n$. What do we know about the initial values $\lambda_i^*(0)$? Do they depend on the leader's announced time path $u^L(\cdot)$ or not? The answer depends on the specific structure of the problem at hand. We now give two examples to illustrate this point.

Example 5.1 In this example there is only one state variable, x. It denotes the stock of pollution that affects the welfare of two countries. One country is the leader in the game, whereas the other is the follower. The control variables u^F and u^L are the follower's and the leader's consumption rates, which are assumed to be proportional to their emission rates. The follower's utility function is assumed to be

$$v^F(x, u^F) = u^F - \frac{(u^F)^2}{2} - \frac{x^2}{2}.$$

The transition equation is $\dot{x}(t) = u^F(t) + u^L(t)$. Assume that $r^F = 0$ and that T is fixed and finite. The follower's Hamiltonian is

$$H^F(x, u^F, \lambda, t) = u^F - \frac{(u^F)^2}{2} - \frac{x^2}{2} + \lambda[u^F + u^L(t)]$$

and we obtain the optimality conditions

$$1 - u^F(t) + \lambda(t) = 0,$$
$$\dot{\lambda}(t) = x(t),$$
$$\dot{x}(t) = 1 + \lambda(t) + u^L(t).$$

The two differential equations together with the boundary conditions $x(0) = x_0$ and $\lambda(T) = 0$ have a unique solution $(\lambda^*(\cdot), x^*(\cdot))$, and it can be verified that $\lambda^*(0)$ depends on the control path $u^L(\cdot)$ of the leader (see appendix). This means that the follower's control variable $u^F(t)$ at time t depends also on future values of $u^L(\cdot)$, that is, on values $u^L(s)$ with $s > t$.

Example 5.2 Consider the same situation as in example 5.1, except that the follower's utility function is

$$v^F(x, u^F) = u^F - \frac{(u^F)^2}{2} - x.$$

In this case the adjoint equation is $\dot{\lambda}(t) = 1$ and, since $\lambda(T) = 0$, we must have $\lambda^*(t) = t - T$. It follows that the leader's choice of $u^L(\cdot)$ has no influence on $\lambda^*(0)$. In fact, in this example, the leader's choice has no influence on the follower's best response $u^F(\cdot)$ at all.

The above examples serve to motivate the following definition.

Definition 5.1 The initial value of the follower's costate variable λ_i is said to be noncontrollable if $\lambda_i^*(0)$ is independent of the leader's control path $u^L(\cdot)$. Otherwise, it is said to be controllable.

We now turn to the leader's problem. She knows the follower's best response to each control path $u^L(\cdot)$. Her optimization problem is to choose a control path $u^L(\cdot)$ to maximize the integral of discounted net benefit (or utility)

$$J^L = \int_0^T e^{-r^L t} v^L(x(t), u^F(t), u^L(t), t)\, dt,$$

where $u^F(t) = g(x(t), \lambda(t), u^L(t), t)$. The maximization is subject to (5.2), (5.5), and (5.7). In this optimization problem, the costate variables λ_i, $i = 1, 2, \ldots, n$, of the follower's optimization problem are treated as state variables in the leader's optimization problem (in addition to the original state variables x_i, $i = 1, 2, \ldots, n$). Note that, while the initial value $x_i(0)$ is fixed at x_{i0}, the initial value $\lambda_i(0)$ is fixed if and only if it is noncontrollable.

The Hamiltonian function for the leader is

$$H^L(x, \lambda, u^L, \psi, \pi, t)$$
$$= v^L(x, g(x, \lambda, u^L, t), u^L, t) + \sum_{i=1}^{n} \psi_i f_i(x, g(x, \lambda, u^L, t), u^L, t) +$$
$$\sum_{i=1}^{n} \pi_i k_i(x, \lambda, u^L, t),$$

where $k_i(x, \lambda, u^L, t)$ denotes the right-hand side of (5.7). The variables ψ_i and π_i are the costate variables associated with the state variables x_i and λ_i, respectively. We then have the optimality conditions

$$\frac{\partial H^L(x(t), \lambda(t), u^L(t), \psi(t), \pi(t), t)}{\partial u_j^L} = 0, \tag{5.8}$$

$$\dot{\psi}_i(t) = r^L \psi_i(t) - \frac{\partial H^L(x(t), \lambda(t), u^L(t), \psi(t), \pi(t), t)}{\partial x_i}, \tag{5.9}$$

$$\dot{\pi}_i(t) = r^L \pi_i(t) - \frac{\partial H^L(x(t), \lambda(t), u^L(t), \psi(t), \pi(t), t)}{\partial \lambda_i}, \tag{5.10}$$

$$\psi_i(T) = 0, \tag{5.11}$$

for all $i \in \{1, 2, \ldots, n\}$ and all $j \in \{1, 2, \ldots, m^L\}$. In addition, if the initial value of the state variable λ_i is controllable, then we have the transversality condition (see (3.17))

$$\pi_i(0) = 0. \tag{5.12}$$

If the Hamiltonian H^L is jointly concave in the state variables x_i and λ_i, $i = 1, 2, \ldots, n$, and the control variables u_j^L, $j = 1, 2, \ldots, m^L$, then the conditions (5.2) and (5.7)–(5.12) are sufficient for the optimality of $u^L(\cdot)$.

Example 5.1 (continued) The leader seeks to maximize

$$\int_0^T \{u^L(t) - (1/2)[u^L(t)^2 + x(t)^2]\}\, dt$$

subject to

$$\dot{x}(t) = 1 + \lambda(t) + u^L(t),$$
$$\dot{\lambda}(t) = x(t),$$

and the boundary conditions

$$x(0) = x_0 \quad \text{and} \quad \lambda(T) = 0. \tag{5.13}$$

The Hamiltonian function for this problem is

$$H^L(x, \lambda, u^L, \psi, \pi) = u^L - (1/2)[(u^L)^2 + x^2] + \psi(1 + \lambda + u^L) + \pi x,$$

from which we obtain

$$1 - u^L(t) + \psi(t) = 0,$$
$$\dot{\psi}(t) = x(t) - \pi(t),$$
$$\dot{\pi}(t) = -\psi(t).$$

Since $x(T)$ and $\lambda(0)$ are free, we have the transversality conditions

$$\psi(T) = 0 \quad \text{and} \quad \pi(0) = 0. \tag{5.14}$$

Substituting $u^L(t) = 1 + \psi(t)$ into the equation for $\dot{x}(t)$, we end up with a system $\dot{z}(t) = Bz(t) + k$ where $z = (x, \lambda, \psi, \pi)'$, $k = (2, 0, 0, 0)'$, and

$$B = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

This system and the four boundary conditions given by (5.13) and (5.14) uniquely determine the solution. An important property of this solution is that it is time inconsistent. To see this, notice that at some time $t_1 > 0$ we have $\pi(t_1) \neq 0$. If the leader can replan at time t_1 in the sense that from time t_1 onwards she is no longer bound to the initially announced control path $u^L(\cdot)$, she will choose a new solution which satisfies $\pi(t_1) = 0$. This implies that she will announce a new control path $u^L(\cdot)$ that is different from the original one. The intuition behind this result is as follows. The leader, by announcing at time $t = 0$ a whole control path $u^L(\cdot)$, seeks to influence the follower's choice of $u^F(\cdot)$ in a way that is favourable to the leader. At time t_1, when the follower has carried out his actions for all $t \leq t_1$, the leader no longer has an incentive to keep her promises.

Example 5.2 (continued) We have seen that in this example the leader cannot influence the follower's action. On the other hand, the follower's action does affect the leader's payoff via the effect of u^F on x. Assume that the leader's utility function is

$$v^L(u^L, x) = u^L - \frac{1}{2}(u^L)^2 - \frac{1}{2}x^2.$$

When the discount rate r^L is equal to 0, the leader seeks to maximize

$$\int_0^T v^L(u^L(t), x(t))\, dt$$

subject to

$$\dot{x}(t) = 1 + \lambda(t) + u^L(t),$$
$$\dot{\lambda}(t) = 1,$$

and the boundary conditions $x(0) = x_0$ and $\lambda(T) = 0$. Hence $\lambda(t) = t - T$ and $\dot{x}(t) = 1 + t - T + u^L(t)$. The leader's Hamiltonian is

$$H^L(x, u^L, \psi, t) = u^L - \frac{1}{2}(u^L)^2 - \frac{1}{2}x^2 + \psi(1 + t - T + u^L)$$

and it is straightforward to find the solution of the leader's optimal control problem.

Notice that in this example it is not true that $\lambda(0)$ can be chosen by the leader. Note also that, since $\lambda(\cdot)$ is independent of the actions of the leader, we did not treat λ as an explicit state variable but simply substituted the known solution $\lambda(t) = t - T$ into the model equations. Alternatively, one could treat λ as a state variable. In this case the Hamiltonian would be

$$\tilde{H}^L(x, \lambda, u^L, \psi, \pi) = u^L - \frac{1}{2}(u^L)^2 - \frac{1}{2}x^2 + \psi(1 + \lambda + u^L) + \pi$$

but one would obtain the same solution. We do not have $\pi(0) = 0$ in this case.

We now turn to the case where the time horizon is infinite. Condition (5.5) is now replaced by the limiting transversality condition

$$\lim_{t\to\infty} e^{-r^F t}\lambda(t)x(t) = 0. \tag{5.15}$$

Again it is important to determine in each application whether $\lambda(0)$ is controllable. If it is controllable, then $\pi(0) = 0$, otherwise there is no presumption that $\pi(0) = 0$. We consider below two examples of infinite horizon Stackelberg games involving two players.

Example 5.3 This example is a model of pollution control which resembles example 5.1. This time, however, we allow general utility functions and assume that the stock of pollution is subject to natural decay. Moreover, the time horizon is infinite. Let $P(t)$ denote the stock of pollution at time t that affects the two communities L and F. Each community $i \in \{L, F\}$ produces a consumption good whose output at time t is denoted by $Q^i(t)$. Production gives rise to the emission $E^i(t)$ of a pollutant. Assume for simplicity that $E^i(t) = Q^i(t)$. The evolution of the stock of pollution is described by

$$\dot{P}(t) = Q^F(t) + Q^L(t) - kP(t), \tag{5.16}$$

where $k > 0$ is the rate of decay.

The objective function of community $i \in \{F, L\}$ is

$$\int_0^\infty e^{-rt}[U_i(Q^i(t)) - D_i(P(t))]\,dt,$$

where $U_i(Q^i(t))$ represents the utility of consumption and $D_i(P(t))$ is the damage caused by pollution. We assume that $U_i(\cdot)$ is increasing and strictly concave and that $D_i(\cdot)$ is increasing and strictly convex.

At the beginning of the game, the leader announces the control path $Q^L(\cdot)$. The follower, taking this time path as given, seeks to maximize his objective function by choosing the control path $Q^F(\cdot)$.

The follower's Hamiltonian function is

$$H^F(P, Q^F, \lambda, t) = U_F(Q^F) - D_F(P) + \lambda[Q^F + Q^L(t) - kP].$$

This gives the optimality conditions

$$U'_F(Q^F(t)) + \lambda(t) = 0,$$
$$\dot{\lambda}(t) = (r + k)\lambda(t) + D'_F(P(t)).$$

The first condition yields $Q^F(t)$ as a function of $\lambda(t)$. Formally, we have $Q^F(t) = f(\lambda(t))$ where $f(\lambda) = (U'_F)^{-1}(-\lambda)$. From the implicit function theorem one can see that $f'(\lambda) = -1/U''_F(f(\lambda)) > 0$. It follows that, if we can find functions $\lambda(\cdot)$ and $P(\cdot)$ that satisfy the differential equations

$$\dot{P}(t) = f(\lambda(t)) + Q^L(t) - kP(t), \tag{5.17}$$

$$\dot{\lambda}(t) = (r + k)\lambda(t) + D'_F(P(t)), \tag{5.18}$$

and the boundary conditions $P(0) = P_0$ and

$$\lim_{t\to\infty} e^{-rt}\lambda(t)P(t) = 0, \tag{5.19}$$

then the optimal open-loop strategy of the follower is given by $Q^F(\cdot) = f(\lambda(\cdot))$.

Note that $\lambda(0)$ is not exogenously given while $P(0)$ is. In the case, where both $U_F(\cdot)$ and $D_F(\cdot)$ are quadratic functions, it can be verified that $\lambda(0)$ is controllable.

The leader's maximization problem can now be formulated. The leader takes the differential equations (5.17) and (5.18) as constraints. This means that λ is treated as a state variable in the leader's optimization problem. The Hamiltonian for this problem is

$$H^L(P, \lambda, Q^L, \psi, \pi) = U_L(Q^L) - D_L(P) + \psi[f(\lambda) + Q^L - kP] + \pi[(r + k)\lambda + D'_F(P)].$$

This gives us the optimality conditions

$$U'_L(Q^L(t)) + \psi(t) = 0,$$
$$\dot{\psi}(t) = (r+k)\psi(t) - \pi(t)D''_F(P(t)) + D'_L(P(t)),$$
$$\dot{\pi}(t) = -k\pi(t) - \psi(t)f'(\lambda(t)).$$

To simplify the computation, let us assume that $U_i(Q^i) = AQ^i - (1/2)(Q^i)^2$ and $D_i(P) = (s/2)P^2$ for $i \in \{L, F\}$, where s is a strictly positive parameter. Then $Q^F(t) = f(\lambda(t)) = A + \lambda(t)$ and $Q^L(t) = A + \psi(t)$. We then have the differential equations

$$\dot{P}(t) = 2A + \lambda(t) + \psi(t) - kP(t),$$
$$\dot{\lambda}(t) = (r+k)\lambda(t) + sP(t),$$
$$\dot{\psi}(t) = (r+k)\psi(t) - s\pi(t) + sP(t),$$
$$\dot{\pi}(t) = -k\pi(t) - \psi(t).$$

Since the initial value $\lambda(0)$ is controllable we have the associated transversality condition $\pi(0) = 0$. The system of four differential equations has the unique steady state solution

$$P_\infty = \frac{2A}{k + s(r+k)^{-1} + s(r+k+sk^{-1})^{-1}} > 0,$$
$$\lambda_\infty = \frac{-sP_\infty}{r+k} < 0, \quad \psi_\infty = \frac{-sP_\infty}{r+k+sk^{-1}} < 0, \quad \pi_\infty = -k^{-1}\psi_\infty > 0.$$

The Jacobian matrix of the system of linear differential equations is

$$J = \begin{pmatrix} -k & 1 & 1 & 0 \\ s & k+r & 0 & 0 \\ s & 0 & k+r & -s \\ 0 & 0 & -1 & -k \end{pmatrix}.$$

The four characteristic roots of this matrix are

$$\frac{r}{2} \pm \left(\frac{r^2}{4} - \frac{w}{2} \pm \frac{1}{2}\Delta^{1/2}\right)^{1/2},$$

where $w = -2k^2 - 2rk - 3s < 0$ and $\Delta = w^2 - 4\det J$. A sufficient condition for having exactly two roots with negative real parts is that $w < 0$ and $\det J > 0$.[2] These conditions are satisfied in our example. It follows that the steady state is stable in the saddle point sense. The saddle point

[2] See section 5.4 for references to results of this type.

property means that, given the initial stock of pollution, $P(0) = P_0$, and given $\pi(0) = 0$, which is one of the transversality conditions, we can find initial values $\lambda(0)$ and $\psi(0)$ such that the system converges to the steady state $(P_\infty, \lambda_\infty, \psi_\infty, \pi_\infty)$ described above as time approaches infinity.

This example also displays the property of time inconsistency. We touched briefly on this issue in the discussion of example 5.1, and the reader can also consult section 4.3. The essence is that, if the leader is allowed to determine her optimal control path for the problem which starts at $t_1 > 0$, then this path does not coincide with the remaining part of the optimal control path calculated at time 0. In the present pollution control problem, from a formal point of view, the time inconsistency can be seen as follows. Suppose we have solved the problem for the leader. Let the optimal solution be $(P^*(\cdot), \lambda^*(\cdot), \psi^*(\cdot), \pi^*(\cdot))$, with $P^*(0) = P_0$ and $\pi^*(0) = 0$. The function $\psi^*(\cdot)$ determines the time path of the leader's output announced by her at the outset of the game. (Recall that $U'_L(Q^{L*}(t)) = -\psi^*(t)$ and, hence, $Q^{L*}(t) = (U'_L)^{-1}(-\psi^*(t))$.) Let the game proceed, with the leader sticking to her announced output path. Then, at some time $t_1 > 0$, we observe the values $(P^*(t_1), \lambda^*(t_1), \psi^*(t_1), \pi^*(t_1))$ with $\pi^*(t_1) \neq 0$. Suppose that at time t_1 we allow the leader to revise her announced output path in the sense that the revised output $Q^L(t)$ need not be identical to $Q^{L*}(t)$ for $t \geq t_1$. At this new starting time t_1, the leader faces the new initial stock $P(t_1) = P^*(t_1)$, which she must take as given. However, she does not have to take $\lambda(t_1) = \lambda^*(t_1)$ as a given initial condition. Instead, she is free to choose a new $\lambda(t_1) \neq \lambda^*(t_1)$. For this reason, the optimal value of the associated costate variable is $\pi(t_1) = 0$ (this is the transversality condition at the new initial time t_1). Thus $\pi(t_1) \neq \pi^*(t_1)$, implying that the new solution is not a continuation of the solution of the original problem.

We now offer an example of an infinite horizon differential game in which $\lambda(0)$ is noncontrollable and, hence, the open-loop Stackelberg equilibrium is time consistent.

Example 5.4 This example deals with the issues of taxation and provision of a government service. The follower is a representative consumer who owns a stock of capital that provides a continuous flow of income. We denote by $k(t)$ the capital stock at time t and assume that each unit of capital yields $A > 0$ units of income. The leader is the government that taxes the consumer's income flow at the rate $\theta(t)$ where $0 \leq \theta(t) \leq 1$. The flow of tax revenue, $\theta(t)Ak(t)$, is used to provide a flow of government service $g(t)$, which the consumer assumes to be independent of her consumption and capital accumulation decisions. The follower takes also the time path $\theta(\cdot)$ as given and chooses the path of consumption $c(\cdot)$ to maximize her integral of discounted utility

$$\int_0^\infty e^{-\rho t}[\ln c(t) + \ln g(t)]\,dt$$

subject to the constraints

$$\dot{k}(t) = [1 - \theta(t)]Ak(t) - c(t) \tag{5.20}$$

and $k(0) = k_0 > 0$. Here, $\rho > 0$ is the discount rate. We also impose the condition $\lim_{t\to\infty} k(t) \geq 0$.

The follower's Hamiltonian is

$$H^F(k, c, \lambda, t) = \ln c + \ln g(t) + \lambda\{[1 - \theta(t)]Ak - c\}$$

from which we get the optimality conditions

$$\frac{1}{c(t)} = \lambda(t), \tag{5.21}$$

$$\dot{\lambda}(t) = \lambda(t)\{\rho - [1 - \theta(t)]A\}. \tag{5.22}$$

The limiting transversality condition is

$$\lim_{t\to\infty} e^{-\rho t}\lambda(t)k(t) = 0. \tag{5.23}$$

We now show that $\lambda(0)$ is not controllable by the leader. To do this, use (5.20) and (5.21) to get $\lambda(t)\dot{k}(t) = [1 - \theta(t)]Ak(t)\lambda(t) - 1$. Together with (5.22) this implies

$$\lambda(t)\dot{k}(t) + \dot{\lambda}(t)k(t) = -1 + \rho k(t)\lambda(t).$$

Observing that the left-hand side of this equation is $d[k(t)\lambda(t)]/dt$ we obtain

$$k(t)\lambda(t) = 1/\rho + Qe^{\rho t},$$

where Q is an integration constant. Using (5.23) we get $Q = 0$. It follows that

$$\lambda(t) = 1/[\rho k(t)] \tag{5.24}$$

and, therefore, $\lambda(0) = 1/(\rho k_0)$, independently of the leader's control path $\theta(\cdot)$. Note in particular that $\lambda(0)$ is not equal to zero. From (5.20), (5.21), and (5.24) we have

$$c(t) = \rho k(t) \tag{5.25}$$

and

$$\dot{k}(t) = [1 - \theta(t)]Ak(t) - \rho k(t),$$

from which we get

$$k(t) = k_0 \exp\left\{\int_0^t [(1-\theta(s))A - \rho]\,ds\right\}.$$

This equation, together with (5.25), implies that the consumption rate at time t does not depend on the tax rates that apply after t. Hence the follower does not condition her action at time t on the leader's actions after time t (i.e., the consumer's strategy is nonanticipating). This is a special property of this example that is due to the assumptions that utility is logarithmic and the gross return to capital is constant.

We now turn to the leader's problem. The government, acting as the leader, announces a time path for the tax rate $\theta(\cdot)$. The tax revenue is used to provide a government service $g(t) = \theta(t)Ak(t)$. The utility function of the government is assumed to be $\ln c + \ln g$. Since $c(t) = 1/\lambda(t)$, the leader's objective is to maximize

$$\int_0^\infty e^{-\rho t}\{\ln[1/\lambda(t)] + \ln[\theta(t)Ak(t)]\}\,dt$$

subject to

$$\dot{k}(t) = [1-\theta(t)]Ak(t) - 1/\lambda(t), \tag{5.26}$$

$$\dot{\lambda}(t) = \lambda(t)\{\rho - [1-\theta(t)]A\}, \tag{5.27}$$

and the initial conditions $k(0) = k_0$ and $\lambda(0) = 1/(\rho k_0)$.

The leader treats k and λ as state variables. Let ψ and π be the associated costate variables in the leader's Hamiltonian. We obtain the optimality conditions

$$\theta(t) = \frac{1}{[k(t)\psi(t) - \pi(t)\lambda(t)]A}, \tag{5.28}$$

$$\dot{\psi}(t) = \psi(t)\{\rho - [1-\theta(t)]A\} - \frac{1}{k(t)}, \tag{5.29}$$

$$\dot{\pi}(t) = \pi(t)[1-\theta(t)]A + \frac{1}{\lambda(t)} - \frac{\psi(t)}{\lambda(t)^2}. \tag{5.30}$$

In the appendix it is shown that the system (5.26)–(5.30) has the solution

$$k(t)\psi(t) = G - t, \tag{5.31}$$

$$\pi(t)\lambda(t) = Qe^{\rho t} - (2/\rho) + G - t, \tag{5.32}$$

$$\theta(t) = \frac{\rho}{A(2 - Q\rho e^{\rho t})}, \tag{5.33}$$

where G is an integration constant and $Q = \lim_{t\to\infty} \pi(t)\lambda(t)e^{-\rho t}$. If we choose $Q = 0$, then we obtain the constant tax rate $\theta(t) = \rho/(2A)$ and the solution is time consistent. Here, we must assume that $\rho \leq 2A$ if we insist that $\theta \leq 1$. We would like to stress that this time consistency property is quite unusual in open-loop Stackelberg games. In this example it is due to the fact that the consumer's consumption is dependent only on her current stock of capital. This is in turn due to the combination of logarithmic utility and linear return to capital.

Notice that, in this example, the time consistent solution $\theta(t) = \rho/(2A)$ does not imply that $\pi(t) = 0$ identically. To see this, consider equation (5.30). If $\pi(t_1) = 0$ at some $t_1 \geq 0$ then

$$\dot{\pi}(t_1) = \frac{1}{\lambda(t_1)} - \frac{\psi(t_1)}{\lambda(t_1)^2} = \rho k(t_1)[1 - \rho k(t_1)\psi(t_1)] = \rho k(t_1)[1 - \rho(G - t_1)]$$

which equals zero only if $G = t_1 + 1/\rho$.

So far, we have considered only examples where there is a single follower and a leader. It is useful to turn now to a variant of the open-loop Stackelberg leadership game. More specifically, we refer to situations where there are a large number of 'small' followers, such that the actions of any single follower have no impact on the leader's payoff but, in the aggregate, the actions of all followers do have a noticeable impact on the leader's payoff. The usual interpretation is that the leader is a government and the followers are price-taking firms and consumers. Instead of formulating a fairly general model to depict this type of situation, we simply work with a few examples.

Example 5.5 This example depicts an oil-importing Stackelberg leader. Consider a world consisting of two countries, which we call country H (the home country) and country F (the foreign country). In each country there exists a continuum $[0, N]$ of identical consumers, each having a quasi-linear utility function $u(c(t), q(t)) = v(c(t)) + q(t)$, where $c(t)$ is the representative consumer's individual consumption of oil at time t and $q(t)$ is her consumption of the numeraire good at time t. Assume that $v(0) = 0$, $v'(0) = a > 0$, $v'(\bar{c}) = 0$ for some $\bar{c} > 0$, and $v''(c) < 0$ for all $c > 0$. Let $p(t) > 0$ denote the price of oil in terms of the numeraire good at time t. Each individual receives an endowment of the numeraire good at the rate $y(t)$ at time t. This endowment flow is taken as exogenous. Assume for simplicity that oil, once extracted, is nonstorable.

In country F consumers take the oil price as given. Assume that the rate of interest r is constant over time and is equal to the rate of discount ρ. The representative consumer in country F chooses the control paths $c(\cdot)$ and $q(\cdot)$ to maximize the integral of her discounted utility

$$\int_0^\infty e^{-rt}[v(c(t)) + q(t)]\,dt$$

subject to the budget constraint

$$\int_0^\infty e^{-rt}[p(t)c(t) + q(t)]\,dt = \int_0^\infty e^{-rt}y(t)\,dt.$$

Substituting the budget constraint into the objective function, we get

$$\int_0^\infty e^{-rt}[v(c(t)) - p(t)c(t) + y(t)]\,dt, \tag{5.34}$$

which shows that the control variable $q(t)$ has been eliminated. We explicitly impose the constraint that $c(t) \geq 0$. Then the maximization of the objective function (5.34) gives the first order optimality conditions

$$v'(c(t)) - p(t) \leq 0, \quad c(t) \geq 0, \quad c(t)[v'(c(t)) - p(t)] = 0.$$

In what follows we assume that $v(c) = ac - (b/2)c^2$, where a and b are positive constants. Then the consumer's demand function is $c(t) = \max\{0, [a - p(t)]/b\}$ and the aggregate flow demand for oil by consumers in country F is

$$C_F(t) = N \max\{0, [a - p(t)]/b\}.$$

Without loss of generality we normalize by setting $N = 1$ in what follows. In country F, there is also a continuum $[0, M]$ of identical oil producers, each endowed with an initial stock of oil $s(0) = s_0$. Every oil producer takes the oil price path as given and seeks to find a time path of oil extraction $E(\cdot)$ to maximize the discounted stream of profit

$$\int_0^\infty e^{-rt}p(t)E(t)\,dt$$

subject to $E(t) \geq 0$ and $\int_0^\infty E(t)\,dt \leq s_0$. The extraction cost is assumed to be zero. The profits are distributed as dividends to consumers in country F. These dividends are included in the income flow $y(\cdot)$ but, owing to the quasi-linear utility function, they do not affect the consumers' demand for oil.

Clearly the profit maximization problem has a solution with positive extraction at each $t \in [t_a, t_b]$ if and only if $e^{-rt}p(t) = e^{-rt_a}p(t_a)$ holds for $t \in [t_a, t_b]$ and $e^{-rt}p(t) \leq e^{-rt_a}p(t_a)$ otherwise.[3] The aggregate initial stock of oil is $S_0 = Ms_0$. Again, we normalize by setting $M = 1$.

[3]This is a simple version of the well known Hotelling rule, which says that net price, i.e., price minus marginal cost, must rise at the rate of interest along any positive extraction path if firms are perfectly competitive. See chapter 12 for a derivation and discussion of Hotelling's rule.

Now consider country H. Assume that this economy has no oil producers and that it is inhabited by a continuum of consumers who are, in terms of preferences and endowment, identical to those in country F (except that they do not own any oil firms). We study two scenarios. In the first scenario consumers in both economies are price takers. In the second scenario the consumers in country H form a coalition that collectively commits itself to a time path of oil consumption. The intention is that, by restricting their demand for oil, they are able to drive down the price of oil, thus achieving a gain in the terms of trade (they export the numeraire good to finance their imports of oil from country F's oil producers). However they must find an optimal tradeoff: the gain from a lower price is achieved at the cost of restricting oil consumption. The second scenario is in fact the equilibrium of a dynamic game with an open-loop Stackelberg leader (the coalition of consumers in country H). We will also discuss the problem of time inconsistency of this equilibrium.

Scenario 1 (*Competitive Equilibrium*): The demand curve in country $i \in \{F, H\}$ is

$$C_i(t) = [a - p(t)]/b \quad \text{if } p(t) \in [0, a].$$

Let T be the time when $p(\cdot)$ reaches its 'choke value' a, that is, $p(T) = a$. After T, the consumption of oil is zero. The world stock of oil is exhausted at T. For $t \in [0, T]$, we have $p(t)e^{-rt} = p(T)e^{-rT}$, which is the Hotelling rule mentioned in footnote 3. Total demand must equal total supply so that

$$\int_0^T [C_H(t) + C_F(t)]\, dt = \frac{2}{b} \int_0^T [a - p(t)]\, dt = S_0.$$

This equation determines the exhaustion time T. Upon integration, we get

$$rT - 1 + e^{-rT} = \frac{rbS_0}{2a}. \tag{5.35}$$

This implies that $T = \phi(rbS_0/(2a))$, where $\phi : [0, \infty) \mapsto [0, \infty)$ is a function satisfying $\phi(0) = 0$ and $\phi'(z) > 0$ for all $z \geq 0$. This shows that the exhaustion time is an increasing function of the initial stock S_0. The resulting equilibrium price at $t = 0$ is $p(0) = ae^{-r\phi(rbS_0/(2a))}$. The integral of discounted utility of consumers can then be computed.

Notice that in this equilibrium each country consumes, by symmetry, half of the world's oil stock, i.e.,

$$\int_0^T C_H(t)\,dt = \int_0^T C_F(t)\,dt = S_0/2.$$

Scenario 2 (*Stackelberg Leadership*): Now assume that the consumers of country H (the home country) form a coalition. They commit themselves to a time path of consumption (i.e., imports) of oil, denoted by $C_H(\cdot)$. Let T_H denote the time beyond which their consumption of oil becomes zero and remains zero ever after, that is

$$T_H = \sup\{t \geq 0 \mid C_H(t) > 0\}.$$

Let Z_0 denote the (implied) total accumulated consumption of oil by the coalition, that is

$$Z_0 = \int_0^{T_H} C_H(t)\,dt.$$

The price the home country must pay at time t is the price at which firms are willing to supply oil. This price is $p(t)$ and, given that the cost of extraction is zero, Hotelling's rule implies that $p(t) = p(0)e^{rt}$. The initial price $p(0)$ must be such that the market clears, that is, it must generate a price path that equates total oil demand with total oil supply. Formally, this can be expressed as

$$\int_0^{T_F} C_F(t)\,dt + \int_0^{T_H} C_H(t)\,dt = S_0,$$

where the oil demand in the foreign country (country F) is given by

$$C_F(t) = [a - p(t)]/b = [a - ae^{r(t-T_F)}]/b \qquad \text{for } t \in [0, T_F]$$

and T_F is the time at which $p(T_F) = a$. Recalling (5.35), it is easy to see that the time T_F satisfies the equation

$$rT_F - 1 + e^{-rT_F} = \frac{rb}{a}[S_0 - Z_0],$$

from which we obtain $T_F = \phi(rb(S_0 - Z_0)/a)$. Here $\phi(\cdot)$ is the same function that we used in the discussion of scenario 1. In particular, it holds that $\phi(0) = 0$ and $\phi'(z) > 0$ for all z. It follows that the market price path $p(\cdot)$ is uniquely determined once we know the accumulated demand for oil by the home country, Z_0. More specifically, given Z_0 the equilibrium price path satisfies

$$p(0) = p(t)e^{-rt} = p(T_F)e^{-rT_F} = ae^{-r\phi(rb(S_0-Z_0)/a)}.$$

The integral of discounted utility of the coalition of consumers in the home country is

$$\int_0^\infty e^{-rt}[v(C_H(t)) - p(t)C_H(t) + y(t)]\,dt$$
$$= \int_0^\infty e^{-rt}[v(C_H(t)) + y(t)]\,dt - Z_0 a e^{-r\phi(rb(S_0-Z_0)/a)}.$$

This integral is to be maximized with respect to the committed consumption path $C_H(\cdot)$ and the total consumption Z_0, subject to $\int_0^{T_H} C_H(t)\,dt = Z_0$. It will be verified that at the optimum $Z_0 < S_0/2$, indicating that the home country imports less oil than under the competitive equilibrium of scenario 1. Consequently, the resulting equilibrium price $p(0)$ is lower than the competitive equilibrium price obtained in scenario 1.

To see this, we define the state variable $Z(t)$ as the home country's cumulative consumption of oil from time t onwards. This implies that $Z(0) = Z_0$ and $\dot{Z}(t) = -C_H(t)$. We can now write the Hamiltonian for the home country's problem as

$$H(Z, C_H, \pi, t) = v(C_H) + y(t) - \pi C_H.$$

Then we have the conditions

$$v'(C_H(t)) - \pi(t) = 0 \quad \text{if } C_H(t) > 0,$$
$$\dot{\pi}(t) = r\pi(t).$$

The transversality condition which determines the optimal value for Z_0 is

$$\pi(0) = p(0)\left[1 + \frac{r^2 b Z_0}{a}\phi'(rb(S_0 - Z_0)/a)\right],$$

where $p(0) = ae^{-r\phi(rb(S_0-Z_0)/a)}$.[4] From these optimality conditions one can easily derive

$$a - bC_H(0) - p(0)\left[1 + \frac{r^2 b Z_0}{a}\phi'(rb(S_0 - Z_0)/a)\right] = 0. \tag{5.36}$$

Recall that T_H is the time at which the home country's consumption of oil falls to zero. Then for all $t \le T_H$ we must have

$$[a - bC_H(t)]e^{-rt} = [a - bC_H(T_H)]e^{-rT_H} = ae^{-rT_H}. \tag{5.37}$$

This equation and the condition $Z_0 = \int_0^{T_H} C_H(t)\,dt$ yield

$$rT_H - 1 + e^{-rT_H} = rbZ_0/a$$

[4]See (3.18) for a derivation of this kind of transversality condition.

and, hence, $T_H = \phi(rbZ_0/a)$. Substituting this and (5.37) into (5.36) we obtain an equation that determines the optimal value of Z_0. This equation is given by

$$ae^{-r\phi(rbZ_0/a)} - ae^{-r\phi(rb(S_0-Z_0)/a)}\left[1 + \frac{r^2bZ_0}{a}\phi'(rb(S_0 - Z_0)/a)\right] = 0.$$

Denoting the left-hand side of this equation by $G(Z_0)$ it is easy to see that $G(0) > 0$ and $G(S_0/2) < 0$. It follows that $G(Z_0) = 0$ at some $Z_0 < S_0/2$. The intuition behind this result is that the coalition commits to a lower cumulative demand, so that the net supply of oil to consumers of country F, $S_0 - Z_0$, is greater than $S_0/2$. This drives down the world price for oil at $t = 0$. Thus the coalition of consumers in country H pays a lower price for their oil imports.

Let us now discuss the issue of time inconsistency in this scenario. From the above analysis we obtain $T_H < T_F$ because $Z_0 < S_0 - Z_0$. It follows that, at the time when the oil consumption in country H becomes zero, the world stock of oil is still positive. Clearly, if the coalition of consumers in country H is allowed to renege at time T_H from their earlier commitment that they would stop importing oil at T_H, then they would find it in their interest to start importing oil again. This would cause the price of oil to jump up discontinuously at the time the commitment is cancelled. Producers would be caught by surprise, because they would be expecting the oil price to rise continuously at the rate r.

The game described above is not a plausible story, because we would expect the producers, being rational, to anticipate that the commitment will not be honoured. It would seem that a reasonable equilibrium concept should display the property that no one will be caught by surprise. This argument puts serious doubt on the concept of an open-loop Stackelberg equilibrium. In the next section, we shall therefore consider nondegenerate Markovian Stackelberg equilibria. Unfortunately, it will be seen that this leads, in general, to intractable problems and that it resolves the time inconsistency only under quite restrictive assumptions about the model and the class of admissible strategies. Before doing so, we consider one more example of an open-loop Stackelberg equilibrium with many followers. This one deals with redistributive capital taxation.

Example 5.6 Consider an economy with a continuum of identical capitalists over an interval $[0, N]$ and a continuum of workers over an interval $[0, M]$. Assume that capitalists save optimally and that workers do not save. We model the government as a leader and the capitalists as followers. The workers are not players in the game. The task of the government is to find a (time-dependent) capital income tax rate (which can be negative) and transfer the tax revenue to the workers as lump sum transfers so as to achieve maximum welfare. If the capital income tax rate is negative, this means the government subsidizes capital income and the subsidy is financed by a lump sum tax on the workers.

Let $a(t)$ denote the stock of capital owned by the representative capitalist at time t. We restrict attention to a symmetric equilibrium, so that the economy's stock of capital is $K(t) = Na(t)$. Each capitalist decides on his time path $a(\cdot)$ whereby he takes the aggregate time path $K(\cdot)$ as given. Each worker supplies a fixed flow of labour, normalized at unity, regardless of the wage rate. The economy's capital–labour ratio at time t is $k(t) = K(t)/M$. Output per worker at time t is $y(t) = f(k(t))$, where $f(0) = 0, f'(k) > 0$, and $f''(k) < 0$ for all k.

In what follows, we set $M = N = 1$ without loss of generality. There is no depreciation of the capital stock. Under perfect competition, the rental rate at time t is $f'(k(t))$. Capitalists can either rent out their capital, earning $f'(k(t))$ per unit, or lend their capital at the interest rate $r(t) = f'(k(t))$. They are indifferent between these two ways of earning income. (We assume a perfect capital market.) The government imposes a tax rate $\theta(t)$ on rental income and on interest income. The net income per unit of capital of the representative capitalist is therefore $z(t) = [1 - \theta(t)]f'(k(t)) = [1 - \theta(t)]r(t)$, which is the after-tax interest rate. Each capitalist, taking as given the time path $z(\cdot)$, seeks to maximize the integral of discounted utility of consumption

$$\int_0^\infty e^{-\rho t} u(c(t))\, dt$$

subject to the intertemporal budget constraint

$$\int_0^\infty c(t) e^{-\int_0^t z(s)\, ds}\, dt = a(0).$$

If we define

$$a(t) = \int_t^\infty c(\tau) e^{-\int_t^\tau z(s)\, ds}\, d\tau,$$

then the intertemporal budget constraint can also be written as

$$\dot{a}(t) = z(t)a(t) - c(t) \tag{5.38}$$

where $a(0)$ is given and

$$\lim_{t\to\infty} a(t)e^{-\int_0^t z(s)\,ds} = 0. \tag{5.39}$$

Equation (5.39) acts as a terminal state constraint on $a(t)$.

It is easy to show that the optimal consumption path must satisfy the condition

$$\frac{\dot{c}(t)}{c(t)} = \frac{1}{\gamma}[z(t) - \rho],$$

where $\gamma = -cu''(c)/u'(c)$ is called the elasticity of marginal utility (or, in the context of choice under uncertainty, the coefficient of relative risk aversion). We assume γ to be a positive constant (independent of c). Integration yields

$$c(t) = c(0)\exp\left\{\int_0^t [z(s) - \rho]/\gamma\,ds\right\}. \tag{5.40}$$

Substitute this into (5.38) and integrate to get

$$a(0) = c(0)\int_0^\infty \exp\left\{\int_0^t z(s)(1/\gamma - 1) - \rho/\gamma\,ds\right\}dt, \tag{5.41}$$

where we have made use of (5.39). Thus, the optimal consumption at time $t = 0$ depends on the whole time path $z(\cdot)$ as well as on the initial stock.[5]

We obtain from (5.40) and (5.41) the representative capitalist's reaction function

$$c(t) = \frac{a(0)\exp\{\int_0^t [z(s) - \rho]/\gamma\,ds\}}{\int_0^\infty \exp\{\int_0^t z(s)(1/\gamma - 1) - \rho/\gamma\,ds\}\,dt}. \tag{5.42}$$

If $\gamma = 1$, then (5.42) simplifies to

$$c(t) = \rho a(0)\exp\left\{\int_0^t [z(s) - \rho]\,ds\right\}, \tag{5.43}$$

which shows that, in this special case, future values of $z(\cdot)$ have no impact on current consumption.

The government must find a time path $z(\cdot)$ that maximizes the integral of the discounted flow of a weighted sum of the workers' utility, $v(f(k(t)) - z(t)k(t))$, and the capitalists' utility, $u(c(t))$. Note that $f(k(t)) - z(t)k(t)$ denotes the representative worker's consumption. Thus, the government's objective is

[5]Note, however, that in the special case where $\gamma = 1$ (which holds if $u(c) = \ln(c)$) equation (5.41) reduces to $c(0) = \rho a(0)$ such that the optimal consumption at time $t = 0$ is independent of $z(\cdot)$.

$$\int_0^\infty e^{-\rho t}[\alpha v(f(k(t)) - z(t)k(t)) + u(c(t))]\,dt,$$

where $\alpha > 0$ is the weight given to the workers' utility. The maximization is subject to

$$\dot{k}(t) = z(t)k(t) - c(t), \qquad k(0) = k_0,$$

and (5.42). The latter constraint can be replaced by

$$\frac{\dot{c}(t)}{c(t)} = \frac{1}{\gamma}[z(t) - \rho].$$

The condition

$$\lim_{t\to\infty} k(t)e^{-\int_0^t z(s)\,ds} = 0$$

must also be satisfied. (Recall the corresponding conditions (5.38) and (5.39).)

The Hamiltonian for the government's problem is

$$H(k, c, z, \psi, \pi) = \alpha v(f(k) - zk) + u(c) + \psi(zk - c) + \pi\left[\frac{c}{\gamma}(z - \rho)\right],$$

from which we derive the conditions

$$\begin{aligned}
&-\alpha k(t)v'(f(k(t)) - z(t)k(t)) + \psi(t)k(t) + \pi(t)c(t)/\gamma = 0,\\
&\dot{\psi}(t) = [\rho - z(t)]\psi(t) - \alpha[f'(k(t)) - z(t)]v'(f(k(t)) - z(t)k(t)), \qquad (5.44)\\
&\dot{\pi}(t) = \pi(t)\{\rho - [z(t) - \rho]/\gamma\} - u'(c(t)) + \psi(t).
\end{aligned}$$

Let us focus on a steady state solution with $\dot{c}(t) = \dot{k}(t) = \dot{\psi}(t) = \dot{\pi}(t) = 0$. Putting $\dot{c}(t) = 0$ implies $z(t) = \rho$ while $\dot{\psi}(t) = 0$ implies $f'(k(t)) = \rho$. It follows that the optimal tax on capital income is zero in the steady state and the steady state capital stock is the so-called modified golden rule level $\hat{k}$ defined by $f'(\hat{k}) = \rho$. It can be shown that the optimal policy in the present example is time inconsistent except for the special case where $u(c) = \ln c$, for in this case the capitalists' consumption at time t, $c(t)$, is independent of future values of the tax rate. Note, however, that if $u(c) = \ln c$, then $\gamma = 1$ and hence the steady state must satisfy both $f'(\hat{k}) = \rho$ and $\alpha v'(f(\hat{k}) - \rho\hat{k}) = 1/(\rho\hat{k})$. This is in general not possible, because α and $v(\cdot)$ are arbitrary. We conclude, therefore, that if $u(c) = \ln c$ then a steady state with $\dot{c}(t) = \dot{k}(t) = \dot{\psi}(t) = \dot{\pi}(t) = 0$ generically does not exist. We shall return to this problem in the next section.

The result that the optimal tax on capital income is zero in the steady state hinges upon the open-loop formulation. As will be seen in the next section, this property need no longer hold in a nondegenerate Markovian Stackelberg equilibrium.

5.3 Nondegenerate Markovian Stackelberg equilibria

In the previous section we introduced the concept of open-loop Stackelberg equilibrium and showed that it is, in general, not time consistent. Thus, it is not a plausible equilibrium concept in situations where economic agents cannot credibly commit to a fixed control path. In the present section we investigate whether this is only due to the assumption of open-loop strategies or whether it is caused by the hierarchical game structure per se. To this end, we develop the concept of nondegenerate Markovian Stackelberg equilibrium. It will be seen that the analysis of such an equilibrium in a differential game may lead to considerable technical difficulties but that it may produce a time consistent outcome. Unfortunately, the latter is only true if we severely restrict the set of Markovian strategies that are available to the players so that, in general, Markovian Stackelberg equilibria are hardly more plausible than open-loop Stackelberg equilibria.

We consider for simplicity an environment where the utility functions and the transition equations do not contain t explicitly (we allow time to appear only in the discount factor $e^{-r^i t}$). We also assume that the time horizon is infinite, so that we can find equilibrium strategies that are independent of time. (See chapter 4 for a discussion of stationary Markovian strategies.) Suppose player 2, the leader, can announce to the follower the policy rule that she (the leader) will use throughout the game. Let this policy rule be denoted by $u^L(t) = \phi^L(x(t))$. The follower, taking this rule as given, seeks to maximize his payoff. The optimal controls chosen by the follower must satisfy the HJB equation

$$r^F V^F(x) = \max_{u^F} \left\{ v^F(x, u^F, \phi^L(x)) + \sum_{i=1}^{n} \frac{\partial V^F(x)}{\partial x_i} f_i(x, u^F, \phi^L(x)) \right\}.$$

In principle, this yields the follower's reaction function of the form

$$u^F(t) = \phi^F(x(t), \phi^L(\cdot)).$$

The leader, knowing this reaction function, then chooses among all possible rules $\phi^L(\cdot)$ one that maximizes her objective function. However, since $\phi^L(\cdot)$ can be any function, it is not clear how such an optimal rule can be obtained in practice. Note that the leader's problem is not a standard optimal control problem. The simplest way to resolve this difficulty is to restrict the space of functions from which the leader can choose the strategy $\phi^L(\cdot)$. We now explain this approach in more detail.

One possible restriction is that $\phi^L(\cdot)$ can only be a linear affine function of the state variable.[6] Let

$$\phi^L(x) = a + bx$$

where a and b are real numbers, not functions of t, because we consider stationary strategies. In this case, the follower's reaction function is of the form

$$u^F(t) = \phi^F(x(t), a, b)$$

and the leader's optimization problem amounts to choosing a and b. This optimization problem is well defined and solvable. There are, however, several objections to this approach. In particular, one could argue that, by restricting the leader's strategy space to a parametric family of functions, the parameters of which (here a and b) the leader has to fix at the outset, one is really considering an open-loop strategy for the problem in which the leader has the controls a and b instead of a nondegenerate Markovian strategy for the problem in which the leader has the control variable u^L. Thus, one might say that by using this approach one solves a game with Markovian strategies for the follower and open-loop strategies for the leader, but not a game in which both players have Markovian state information. Moreover, the leader does not only have to fix a and b at the outset but she must also choose these decision variables as constants. As we shall see below, this additional restriction is responsible for the time consistency of the resulting Stackelberg equilibrium. Despite these objections, the approach can make sense and is applied in the literature. We now present some examples.

Example 5.7 This example describes the joint exploitation of a pesticide. A pesticide has the property that its effectiveness declines with the accumulated number of doses. This reflects the ability of the pest population to develop resistance to the pesticide, and this resistance increases with repeated exposure. Let $E(t) \geq 0$ denote the effectiveness of the pesticide at time t. We treat $E(t)$ as the state variable of the model. We propose a differential game between two firms that use the pesticide. Let $y_i(t) \geq 0$ denote firm i's rate of application of the pesticide at time t. We assume that the decline in effectiveness is equal to the total application by both firms. Therefore, we have

[6]Here we consider only problems with one state variable though, in principle, the approach can be extended to problems with several state variables.

$$\dot{E}(t) = \begin{cases} -y_1(t) - y_2(t) & \text{if } E(t) > 0, \\ 0 & \text{if } E(t) = 0. \end{cases} \tag{5.45}$$

Assume that the doses can be produced costlessly, and that the profit rate of firm i at time t is $\pi_i(t) = [y_i(t)E(t)]^\alpha$ where $0 < \alpha < 1/2$. Each firm seeks to maximize the integral of discounted profit

$$\int_0^\infty e^{-rt}\pi_i(t)\,dt$$

subject to (5.45).

Let us begin by considering the case where both firms make their decisions simultaneously (i.e., there is no Stackelberg leader). We seek a stationary Markov perfect equilibrium, i.e., a pair of strategies $y_i(t) = \phi_i(E(t))$, $i = 1, 2$, that are best replies to each other in every possible subgame.

The HJB equation for firm i is

$$rV_i(E) = \max\{(y_iE)^\alpha + V_i'(E)[-y_i - \phi_j(E)] \mid y_i \geq 0\}.$$

It is easy to show that the pair of strategies (ϕ_1, ϕ_2) defined by

$$\phi_i(E) = \frac{rE}{2 - 4\alpha}, \qquad i = 1, 2 \tag{5.46}$$

is a stationary Markov perfect Nash equilibrium and that the corresponding value functions are $V_i(E) = KE^{2\alpha}$, $i = 1, 2$, where

$$K = \frac{1}{2^\alpha}\left(\frac{1 - 2\alpha}{r}\right)^{1-\alpha} > 0.$$

Thus, the equilibrium is symmetric.

Now assume instead that firm 2 is the Stackelberg leader. It can announce in advance its strategy, denoted by $y_2(t) = \phi_2(E(t))$, and commit itself to that announced strategy. Firm 1, the follower, takes $\phi_2(\cdot)$ as given and finds its best reply to $\phi_2(\cdot)$. Denote this best reply by $y_1(t) = R_1(E(t), \phi_2(\cdot))$. Firm 2, knowing the reaction function R_1, then chooses among all permissible $\phi_2(\cdot)$ the one that maximizes its integral of discounted profit.

Let us restrict the space of strategies that the leader can choose from by assuming that only linear strategies $\phi_2(E) = bE$ with $b > 0$ are allowed. Then, for any given $b > 0$ announced by the leader, the follower (firm 1) seeks a solution to its HJB equation

$$rW_1(E) = \max\{(y_1E)^\alpha + W_1'(E)(-y_1 - bE) \mid y_1 \geq 0\}.$$

It can be shown that the follower's best reply is

$$y_1(t) = R_1(E(t), b) = [2K_1(b)]^{1/(\alpha-1)}E(t) \tag{5.47}$$

where

$$K_1(b) = \frac{1}{2^\alpha}\left(\frac{1-\alpha}{r+2b\alpha}\right)^{1-\alpha}.$$

The optimal value function for the follower is

$$W_1(E) = K_1(b)E^{2\alpha}.$$

These equations imply that the greater is the leader's b, the lower is the value of the stock E to the follower, and the faster is the follower's rate of depletion of the stock.

We now turn to the leader's problem. Its objective is to choose a constant b that maximizes

$$\int_0^\infty e^{-rt}[bE(t)^2]^\alpha\, dt \tag{5.48}$$

subject to

$$\dot{E}(t) = -bE(t) - [2K_1(b)]^{1/(\alpha-1)}E(t). \tag{5.49}$$

It is a routine matter to find the optimal value of b. Notice that, if at some date $t_1 > 0$ the leader is allowed to determine a new b, the new optimal b will be the same as the one it chose at $t = 0$. This can be seen, for example, by solving the linear differential equation (5.49) with the initial condition $E(t_1) = E_1$, substituting the result into (5.48), and evaluating the integral over the interval $[t_1, \infty)$. The resulting expression is maximized at the same b for all values $t_1 \geq 0$ and $E_1 \geq 0$. This verifies that the Markovian Stackelberg equilibrium in this particular model displays the property of time consistency. The very same calculation, however, also makes clear that the time consistency property no longer holds if (i) the model is not stationary (that is, the leader's utility function or the system dynamics depend explicitly on t) or (ii) the time horizon is finite.

The next example investigates the question of how to tax a polluting monopolist.

Example 5.8 A monopolist produces a consumption good. His output at time t is denoted by $Q(t)$. The production process emits a flow of pollution $E(t) = Q(t)$ (cf. example 5.3). Let $S(t)$ denote the stock of pollution. We assume that

$$\dot{S}(t) = Q(t) - \delta S(t), \tag{5.50}$$

where $\delta > 0$ is the rate of decay of the pollution stock. The inverse demand function for the consumption good is $P(Q) = M - bQ$ where $M > 0$ and $b > 0$. The unit cost of production is $c > 0$. The government imposes a tax τ per unit of output. We assume that the tax rate is dependent on the pollution stock and we impose the restriction that it is a linear affine function of S, namely

$$\tau(S) = \eta + \alpha S. \tag{5.51}$$

This is a model where the government is the Stackelberg leader and the monopolist is the follower. The follower takes the leader's tax strategy (5.51) as given and seeks to maximize the value of the integral of the discounted stream of profit,

$$\int_0^\infty e^{-rt}[MQ(t) - bQ(t)^2 - cQ(t) - \tau(S(t))Q(t)]\,dt,$$

subject to (5.50) and the initial condition $S(0) = S_0$. Assume for the moment that $\alpha \geq 0$ and that $A := M - c \geq \eta$. Given η and α, the monopolist's problem yields a unique solution. This solution implies a time path $S(\cdot)$ that converges to a steady state stock S_∞ given by

$$S_\infty = S_\infty(\alpha, \eta) := \frac{(A - \eta)(r + \delta)}{2\delta b(r + \delta) + \alpha r + 2\alpha\delta}.$$

The state trajectory $S(\cdot)$ is given by

$$S(t) = S_\infty + (S_0 - S_\infty)e^{\beta t}, \tag{5.52}$$

where

$$\beta = \beta(\alpha) := \frac{r - \sqrt{r^2 + 4\delta r + 4\delta^2 + (4\delta + 2r)\alpha/b}}{2} < 0.$$

The negative root is chosen to ensure convergence to the steady state. Because of $Q(t) = \dot{S}(t) + \delta S(t)$ it follows from (5.52) that the monopolist follows the output strategy $Q(t) = \tilde{Q}(S(t), \alpha, \eta)$, where

$$\tilde{Q}(S, \alpha, \eta) = [\delta + \beta(\alpha)][S - S_\infty(\alpha, \eta)] + \delta S_\infty(\alpha, \eta). \tag{5.53}$$

Turning now to the government, we assume that it seeks to maximize the integral of a weighted sum of discounted flows of consumers' surplus net of pollution damage, producer's surplus, and tax revenue. Consumers' surplus at time t is $U(Q(t)) - P(Q(t))Q(t)$ where

$$U(Q) = \int_0^Q P(\tilde{Q})\,d\tilde{Q}.$$

Pollution damage at time t is assumed to depend on the stock of pollution in the form $\gamma S(t)^2/2$, where γ is a positive parameter. Producers' surplus at time t is $P(Q(t))Q(t) - cQ(t) - \tau(S(t))Q(t)$ and tax revenue at time t is $\tau(S(t))Q(t)$. Assume for simplicity that in the government's objective functional, consumers' and producers' surplus, pollution damages, and tax revenue all appear with the same weight. It follows that the government tries to maximize

$$\int_0^\infty e^{-rt}\left[U(Q(t)) - cQ(t) - \gamma S(t)^2/2\right] dt \tag{5.54}$$

subject to the monopolist reaction function (5.53), the system dynamics (5.50), and the initial condition $S(0) = S_0$. Notice that the tax revenue does not appear in the integrand because it cancels out when added to the producers' surplus. In this optimization problem, the government chooses the parameters α and η which determine the tax rule $\tau(S) = \eta + \alpha S$.

Let us digress for a moment and ask ourselves the following question. If the government could directly control the production, what would be the optimal time path of the pollution stock? It is easy to show that in this case the steady state stock of pollution would be $\hat{S}_\infty$, where

$$\hat{S}_\infty = \frac{A(r+\delta)}{(r+\delta)b\delta + \gamma}$$

and the optimal feedback control rule would be $Q(t) = \hat{Q}(S(t))$, where

$$\hat{Q}(S) = \rho S + (\delta - \rho)\hat{S}_\infty \tag{5.55}$$

and ρ is the negative root of the quadratic equation $\rho^2 - (r + 2\delta)\rho - \gamma/b = 0$.

Returning to the leader's problem (5.54), after some tedious manipulations, it can be shown that the optimal values for α and η are

$$\alpha^* = \frac{2\gamma}{r + 2\delta} \quad \text{and} \quad \eta^* = \frac{[\gamma - b(r+\delta)\delta - (r+2\delta)\alpha^*]\hat{S}_\infty}{r+\delta}.$$

It turns out that with these values the monopolist's output path coincides with the socially optimal output path given above. To see this, it suffices to verify that $S_\infty(\alpha^*, \eta^*) = \hat{S}_\infty$ and that $\beta(\alpha^*) + \delta = \rho$, so that the strategies (5.53) and (5.55) are identical. In view of this result, we conclude that for the linear-quadratic pollution model under consideration, our restriction of the space of tax strategies (i.e., the requirement that the tax rate be a linear affine function of S) is not really restrictive.

The final example of this chapter is actually a continuation of example 5.6 from the preceding section.

Example 5.6 (continued) Let us reconsider the problem of redistributive capital income taxation but, for simplicity, assume that $\gamma = 1$. Recall that if the representative capitalist uses the open-loop strategy stated in (5.43), which is equivalent to $\dot{c}(t) = c(t)[z(t) - \rho]$, then in the steady state the optimal tax on capital income is zero.

Consider now the case where the representative capitalist uses a non-degenerate Markovian strategy. When $\gamma = 1$, we have shown that $c(0) = \rho a(0)$ and, hence, $c(t) = \rho a(t)$ for all t. A more formal way of seeing this is via the use of the HJB equation.

If $\gamma = 1$ then we have $u(c) = \ln c$ and the HJB equation for the representative capitalist's problem is

$$\rho V(t, a) - V_t(t, a) = \max\{\ln c + V_a(t, a)[z(t)a - c] \mid c > 0\}.$$

It can be verified that the function

$$V(t, a) = \int_t^\infty \left\{ e^{-\rho(s-t)} \int_t^s [z(\tau) - \rho]\, d\tau \right\} ds + \frac{\ln(\rho a)}{\rho}$$

satisfies the above functional equation and that the optimal control rule is $c(t) = \rho a(t)$. Thus, the capitalists' current consumption is independent of current and future tax rates. Their current income (and, hence, their savings), on the other hand, are dependent on the current tax rate via $\dot{a}(t) = z(t)a(t) - c(t) = [z(t) - \rho]a(t)$.

The government then seeks to maximize

$$\int_0^\infty e^{-\rho t}[\alpha v(f(k(t)) - z(t)k(t)) + \ln(\rho k(t))]\, dt$$

over $z(\cdot)$ subject to

$$\dot{k}(t) = [z(t) - \rho]k(t).$$

This is a simple optimal control problem with a solution that converges to a steady state k^*. A Markovian control rule $z(t) = \phi(k(t))$ can thus be obtained, at least in principle. At the steady state, we have $z(t) = \rho$, and k^* must satisfy the equation

$$f'(k^*) - 2\rho + \frac{1}{\alpha v'(f(k^*) - \rho k^*)k^*} = 0. \tag{5.56}$$

The steady state tax rate on capital income is θ^* where

$$(1 - \theta^*)f'(k^*) = z(k^*) = \rho.$$

It follows from these two equations that, in general, $\theta^* \neq 0$. In particular, if $v(\cdot) = \ln(\cdot)$ and $f(k) = k^\beta$ then $\theta^* = 0$ if and only if $1 + \alpha = 1/\beta$. This contrasts sharply with the result for the case where capitalists use open-loop strategies (see the discussion of example 5.6 in the previous section). The intuition behind this difference in tax treatment between the open-loop Stackelberg equilibrium and the nondegenerate Markovian Stackelberg equilibrium is as follows. In the open-loop case, consumers are happy with any suggested time path of consumption, as long as their intertemporal budget constraint is satisfied and the rate of change in consumption is proportional to the difference between the after-tax rate of interest and the rate of utility discount (with the factor of proportionality being $1/\gamma$). They do not condition their current consumption on the current size of the economy's capital stock even though, along the equilibrium path, there is a monotone relationship between these two variables. In contrast, when capitalists formulate nondegenerate Markovian consumption strategies, which must be optimal not only along the equilibrium path but also off the equilibrium path, the government has less freedom.

It is interesting to note that, in this example when $\gamma = 1$, the steady state of the Markovian Stackelberg equilibrium turns out to be a kind of steady state of the corresponding open-loop game discussed in the previous section. To see this, consider the constant path $k(t) = k^*$, $c(t) = \rho k^*$, and $z(t) = \rho$ with k^* defined by (5.56). Referring to the optimality conditions in (5.44) we see that, if we set $\psi(0) + \rho\pi(0) = \alpha v'(f(k^*) - \rho k^*)$, then $\dot{\psi}(t) + \rho\dot{\pi}(t) = 0$ even though $\dot{\psi}(t) = -\alpha[f'(k^*) - \rho]v'(f(k^*) - \rho k^*) \neq 0$ and $\dot{\pi}(t) = \alpha v'(f(k^*) - \rho k^*) - 1/(\rho k^*)$.

5.4 Further reading

For a more complete analysis of the open-loop model of pollution control with a Stackelberg leader, see Long [163]. For Markovian Nash equilibria of a related game, see Dockner and Long [56], from which a nondegenerate Markovian Stackelberg equilibrium with linear affine strategies can be derived using the methods explained above (the reader is warned that the calculation of such an equilibrium is quite tedious).

Our definition of controllability (see definition 5.1) is inspired by a recent note by Xie [245], who deals with a class of problems of which example 5.4 is a special case.

The Markovian Stackelberg solution discussed in this chapter is sometimes referred to as global Stackelberg solution (cf. Başar and Olsder [4]). The difficulty with this approach is that the equilibrium cannot be characterized in terms of the solutions of standard optimal control problems.

The approach we have chosen requires specification of the functional form of the strategy of the leader up to finitely many parameters and then letting the leader solve an optimization problem on the parameter space. There exists, however, an alternative approach in which the leader only faces a stagewise advantage over the follower. This Stackelberg solution is sometimes referred to as feedback Stackelberg solution (cf. Başar and Olsder [4]). It requires that both the leader and the follower solve stagewise dynamic programming problems.

For the results which we used in example 5.3 on characteristic roots of a Hamiltonian system with two state variables, see Dockner [42], Dockner and Feichtinger [48], and Kemp et al. [151], where bifurcation theory is also discussed.

Example 5.5 of an open-loop Stackelberg leader with a continuum of competitive followers (the model on oil import restriction) is inspired by the work on an optimal tariff on oil imports by Maskin and Newbery [169] and Kemp and Long [149]. The issue of time inconsistency raised in these two papers has also been discussed by Kydland and Prescott [155].

The Stackelberg leadership model of utilization of a pesticide in example 5.7 is built on a symmetric Nash equilibrium treatment by Cornes et al. [34].

Example 5.8, dealing with the taxation of a polluting monopolist, is based on the work of Benchekroun and Long [5]. For a model of Stackelberg leadership in an oligopoly with learning-by-doing, see Benchekroun et al. [6].

Nondegenerate Markovian equilibria in differential games of capitalism can also be analysed diagrammatically; see Shimomura [217].

5.5 Exercises

1. Analyse the game from example 5.5 (optimal restriction of oil imports) under the assumption that $v(c) = c^{1-b}/(1-b)$ where $0 < b < 1$.
2. (Duopoly with learning-by-doing.) Let $Q_i(t)$ denote firm i's output at time t. The market price is $P(t) = a - b[Q_1(t) + Q_2(t)]$, where a and b are positive constants. Firm 2 is a mature firm, its unit production cost is zero. Firm 1's unit production cost is $c(K(t)) = \bar{c} - \gamma K(t)$, where $K(t)$ denotes its knowledge capital and $\bar{c}$ is a positive constant. Assume that $\dot{K}(t) = Q_1(t) - \delta K(t)$ where $\delta > 0$. Find the open-loop Nash equilibrium, the Markov perfect Nash equilibrium, and a nondegenerate Markovian Stackelberg equilibrium if firm 2 is the leader and can only adopt a linear strategy of the form $Q_2(t) = e + fK(t)$.

3. Generalize example 5.4 to the case where the utility function of the follower is $u(c) = (c^{1-\sigma} - 1)/(1 - \sigma)$ and her gross return per unit of capital is $Ak^{\sigma-1}$, where $0 < \sigma \leq 1$ and $A > 0$. In other words, the state equation is $\dot{k}(t) = [1 - \theta(t)]Ak(t)^{\sigma} - c(t)$.
4. Show that the pair of strategies given by (5.46) constitute a Markov perfect Nash equilibrium of the pesticide game in example 5.7.
5. Show that (5.47) is the follower's best reply in the pesticide game in example 5.7.
6. Find the optimal value of b for the leader's problem (5.48) in example 5.7.
7. Find an open-loop Stackelberg equilibrium in the following model of capital and labour income taxation. There is a continuum of identical consumers, each with the utility function $U(C, L) = \ln(C - L)$ where C is consumption and L is the labour supply. (Implicit in this formulation is the constraint that $L \leq bC$ where b is a coefficient set equal to unity by choice of units; the constraint reflects the fact that work uses up calories which the human body generates by burning food intake.) The representative firm has the production function $Q = K^{\beta}L^{1-\beta}$. Wage W and rental R are the marginal products of labour and capital, respectively. R is also the interest rate. The representative consumer owns capital and government bonds. The amount of government bonds held, denoted by B, can be negative, in which case individuals owe debts to the government. Each consumer takes the time paths $W(\cdot)$, $R(\cdot)$, and the tax rates $\theta_L(\cdot)$ and $\theta_K(\cdot)$ as given. The representative consumer seeks to maximize

$$\int_0^{\infty} e^{-\rho t} \ln[C(t) - L(t)]\, dt$$

subject to

$$\dot{A}(t) = [1 - \theta_K(t)]R(t)A(t) + [1 - \theta_L(t)]W(t)L(t) - C(t),$$

where $A(t) = K(t) + B(t)$ is the total wealth at time t. The initial wealth $A(0) = A_0$ is given and we impose the transversality condition

$$\lim_{t\to\infty} e^{-\rho t}\lambda(t)A(t) = 0$$

where λ is the costate variable associated with A.

The government provides a service flow G which is constant over time. The government's budget deficit is financed by new bonds, that is,

$$\dot{B}(t) = G + R(t)B(t) - [\theta_K(t)R(t)A(t) + \theta_L(t)W(t)L(t)]. \tag{5.57}$$

The government, acting as the Stackelberg leader, chooses the time paths $\theta_K(\cdot)$ and $\theta_L(\cdot)$ to maximize

$$\int_0^\infty e^{-\rho t}\{\ln[C(t) - L(t)] + \ln G\}\, dt$$

subject to (5.57), the initial condition $B(0) = B_0$, and $\lim_{t\to\infty} e^{-\rho t} B(t) = 0$.

Appendix

Controllability in example 5.1

Let $y(t) = (\lambda(t), x(t))'$ and $h(t) = (0, 1 + u^L(t))'$, where the prime indicates transposition. Then the follower's system of differential equations can be written in the matrix form $\dot{y}(t) = Ay(t) + h(t)$ where

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This matrix equation has the solution[7]

$$y(t) = e^{At} y(0) + e^{At} \int_0^t e^{-As} h(s)\, ds, \tag{5.58}$$

where

$$e^{At} = \begin{pmatrix} (e^t + e^{-t})/2 & (e^t - e^{-t})/2 \\ (e^t - e^{-t})/2 & (e^t + e^{-t})/2 \end{pmatrix}$$

and e^{-As} is the inverse of e^{As}. Setting $t = T$ in (5.58) and recalling that $\lambda(T) = 0$ and $x(0) = x_0$, we obtain two equations that determine $\lambda(0)$ and $x(T)$. One of them is

$$\begin{aligned} 0 = (e^T + e^{-T}) &\left\{ \frac{\lambda(0)}{2} + \frac{1}{4} \int_0^T [1 + u^L(s)](e^{-s} - e^s)\, ds \right\} \\ + (e^T - e^{-T}) &\left\{ \frac{x_0}{2} + \frac{1}{4} \int_0^T [1 + u^L(s)](e^{-s} + e^s)\, ds \right\}. \end{aligned}$$

This shows that $\lambda(0)$ depends on $u^L(\cdot)$.

[7] See, for example, pages 40–48 in Brock and Malliaris [12] for solution methods.

Derivation of equations (5.31)–(5.33) in example 5.4

Substitute $\lambda(t) = 1/[\rho k(t)]$ into (5.26). Then the resulting equation and (5.29) give $d[\psi(t)k(t)]/dt = -1$, which proves (5.31). From (5.31) and (5.24) we get $\psi(t) = (G - t)\rho\lambda(t)$. Substitute this into (5.30) and use the result, together with (5.27), to get (5.32). Finally, equations (5.28), (5.31), and (5.32) give (5.33).

6 Trigger strategy equilibria

A Nash equilibrium of a game describes, by definition, a situation in which no player can improve his objective functional by a unilateral deviation from the equilibrium strategy. It has to be emphasized, however, that joint deviations by more than one player could lead to such improvements. In particular, Nash equilibria are usually not Pareto efficient. This observation raises the question of whether there exist efficient Nash equilibria at all and whether there are any general methods for constructing such equilibria. In the present chapter we present one such method which is based on the use of so-called trigger strategies.

Trigger strategies are non-Markovian, that is, they determine the control variable at time t as a function of the entire history of the endogenous variables up to time t and not just as a function of the state at time t. In the first section of this chapter we discuss a few concepts which are important for the analysis of non-Markovian equilibria. In the second section we introduce the main building blocks for trigger strategies: target paths, threats, and punishment. The basic idea of the scenario under consideration is that the agents agree to follow a certain target path and sustain their agreement by threatening to punish any defector. We start by discussing trigger strategies in the framework of an infinite horizon differential game and assume that there is a fixed positive delay δ between the deviation of a defector from the target path and the start of punishment by his opponents. We discuss the credibility of threats and its relation to the property of subgame perfection in a separate section. Finally, we consider the case where the time delay δ is infinitesimally small and the game may have a finite time horizon.

6.1 Non-Markovian strategies

So far we have assumed that players use Markovian strategies, that is, the control variables at time t depend only on time t and on the state of the system at time t, $x(t)$. These strategies are those which are most often used in applied differential game models. However, equilibria which are based on alternative informational assumptions may have quite interesting properties, too. In the present section we deal with non-Markovian equilibria and, in particular, with so-called history-dependent equilibria. We do not attempt to present a complete and rigorous theory but restrict ourselves to sketching the basic concepts.

We consider an N-player differential game with objective functionals

$$\int_0^T e^{-r^i s} F^i(x(s), u^1(s), u^2(s), \ldots, u^N(s), s)\, ds + e^{-r^i T} S^i(x(T)),$$
$$i = 1, 2, \ldots, N, \tag{6.1}$$

system dynamics

$$\dot{x}(s) = f(x(s), u^1(s), u^2(s), \ldots, u^N(s), s), \; x(0) = x_0, \tag{6.2}$$

and control constraints

$$u^i(s) \in U^i(x(s), u^{-i}(s), s) \subseteq \mathbb{R}^{m^i}, \; i = 1, 2, \ldots, N, \tag{6.3}$$

as in chapter 4. The terminal date of the game, T, may be finite or infinite. In the latter case we assume that the scrap value functions S^i, $i \in \{1, 2, \ldots, N\}$, are identically equal to 0. As in previous chapters we shall use the notation $[0, T\rangle$ for the time domain of the game, that is, $[0, T\rangle = [0, T]$ if T is finite and $[0, T\rangle = [0, \infty)$ if $T = \infty$.

Let us denote by $\mathcal{U}$ the set of all N-tuples of feasible control paths for this game. In other words, $\mathcal{U}$ is the set of all N-tuples $u(\cdot) = (u^1(\cdot), u^2(\cdot), \ldots, u^N(\cdot))$ with the following property: $u^i : [0, T\rangle \mapsto \mathbb{R}^{m^i}$ is a measurable function for all $i \in \{1, 2, \ldots, N\}$ and there exists a unique absolutely continuous state trajectory $x(\cdot)$ such that conditions (6.2) and (6.3) are satisfied.[1] An information structure for the game is a mapping

$$I : \{(u(\cdot), t) \mid u(\cdot) \in \mathcal{U}, t \in [0, T\rangle\} \mapsto Y^1 \times Y^2 \times \ldots \times Y^N,$$

where Y^i is the observation space of player i. The interpretation of the information structure I is as follows: if the players choose the feasible paths $u(\cdot) \in \mathcal{U}$, then player i's information at time t is given by $I^i(u(\cdot), t) \in Y^i$ and

$$I(u(\cdot), t) = (I^1(u(\cdot), t), I^2(u(\cdot), t), \ldots, I^N(u(\cdot), t)).$$

[1]Measurability is a weak regularity condition for a function. Any piecewise continuous function is measurable.

The information structure is called nonanticipating if for all $t \in [0, T\rangle$ and all pairs $(u(\cdot), v(\cdot)) \in \mathcal{U} \times \mathcal{U}$ with $u(s) = v(s)$ for all $s \in [0, t)$ it holds that $I(u(\cdot), t) = I(v(\cdot), t)$. It is clear that, in the case of a nonanticipating information structure, the information at time t, $I(u(\cdot), t)$, depends only on the restriction of $u(\cdot)$ to the interval $[0, t)$. An information structure I is called regular if $I(u(\cdot), t) = I(v(\cdot), t)$ holds for all $t \in [0, T\rangle$ and all pairs $(u(\cdot), v(\cdot)) \in \mathcal{U} \times \mathcal{U}$ such that $u(\cdot)$ and $v(\cdot)$ differ from each other only on a set of Lebesgue measure 0.

The information structure tells us on which endogenous variables each player can condition his actions at any point t during the game. As an example, consider the case of the Markovian information structure. Here, $I^i(u(\cdot), t) = x(t)$ where $x(\cdot)$ is the state trajectory corresponding to the control paths $u(\cdot)$. The Markovian information structure is obviously nonanticipating and regular. Another nonanticipating and regular information structure is given by

$$I^i(u(\cdot), t) = \begin{cases} z & \text{if } t \in [0, \delta), \\ x(t-\delta) & \text{if } t \in [\delta, T\rangle, \end{cases}$$

where $\delta > 0$ and $z \in X$ are fixed constants. This information structure describes a situation where player i can observe the state of the system only with a certain time delay δ. As a final example of a nonanticipating and regular information structure consider $I^i(u(\cdot), t) = \{x(T_l) \mid T_l \leq t\}$, where $\{T_1, T_2, \ldots\} \subseteq [0, T\rangle$ is a fixed finite or countably infinite set. This information structure captures a situation where players can observe the state of the game only at the given observation dates T_l.

An example of a regular and anticipating information structure is $I^i(u(\cdot), t) = x(t + \delta)$ with $\delta > 0$. Another anticipating information structure is the one used in an open-loop Stackelberg game (see chapter 5). Here the follower (say, player 2) is supposed to be able to condition his action at time t on the entire control path of the leader (player 1), that is, $I^2(u(\cdot), t) = u^1(\cdot)$. The anticipating information structure is the reason for the time inconsistency of the open-loop Stackelberg equilibrium (see also theorem 6.1 below). Finally, an example of a nonanticipating and irregular information structure is $I^1(u(\cdot), t) = u^2(t-1)$.

The complete formulation of an N-player differential game includes the specification of the objective functionals (6.1), the system dynamics (6.2), the control constraints (6.3), and the information structure I. We shall use the notation $\Gamma_I(x_0, 0)$ for this game, whereby the subscript I indicates that we are dealing with a general information structure I. A strategy for player i in $\Gamma_I(x_0, 0)$ is a mapping $\phi^i : Y^i \times [0, T\rangle \mapsto \mathbb{R}^{m^i}$. At any time t, player i determines his action $u^i(t)$ by applying the strategy ϕ^i

to the current information $I^i(u(\cdot), t)$. Formally this means that the control paths $u^1(\cdot)$, $u^2(\cdot)$, ..., $u^N(\cdot)$ must satisfy the fixed point condition

$$u^i(t) = \phi^i(I^i(u(\cdot), t), t) \tag{6.4}$$

for all $i \in \{1, 2, \ldots, N\}$ and all $t \in [0, T\rangle$. If the information structure is regular, then it suffices that (6.4) holds for almost all $t \in [0, T\rangle$.

An N-tuple $\phi = (\phi^1, \phi^2, \ldots, \phi^N)$ of strategies is called a strategy profile. It is clear that in this general setting we cannot guarantee that a strategy profile determines a unique and feasible N-tuple of control paths. Equation (6.4) may have no solution or multiple solutions in $\mathcal{U}$. If there exists a unique element $u(\cdot) \in \mathcal{U}$ satisfying (6.4), the strategy profile ϕ is called feasible and $u(\cdot)$ is said to be the N-tuple of control paths corresponding to ϕ.

Owing to the countless possibilities that exist for the specification of the information structure it is impossible to state any general results on the existence of feasible strategy profiles. Therefore, let us proceed by assuming that they exist and let us denote the set of all feasible strategy profiles for $\Gamma_I(x_0, 0)$ by $\mathcal{S}_I$. The set of all $(N-1)$-tuples $\phi^{-i} = (\phi^1, \phi^2, \ldots, \phi^{i-1}, \phi^{i+1}, \ldots, \phi^N)$ for which there exists a strategy ϕ^i for player i such that $(\phi^1, \phi^2, \ldots, \phi^N) \in \mathcal{S}_I$ is denoted by $\mathcal{S}_I^{-i}$. It is the set of all strategies of the opponents of player i for which there exists a feasible response by player i. Finally, the set of feasible responses by player i to a given $(N-1)$-tuple $\phi^{-i} \in \mathcal{S}_I^{-i}$ is denoted by $\mathcal{S}_I^i(\phi^{-i})$.

Throughout the following discussion we consider only feasible strategy profiles $\phi \in \mathcal{S}_I$. This can be justified by making the assumption that in the case where a strategy profile is infeasible (for example, because it does not generate a feasible and uniquely defined state trajectory) the payoffs to the players are not determined by (6.1) but are equal to $-\infty$.

It will be convenient to denote the objective functional of player i defined in (6.1) by $J^i(\phi^i; \phi^{-i}, x_0, 0)$. This notation has to be interpreted as follows. Given a strategy profile $\phi \in \mathcal{S}_I$ consisting of the strategies of player i's opponents, ϕ^{-i}, and player i's reponse, ϕ^i, a unique N-tuple of control paths is determined by (6.4). Of course, this also determines a unique state trajectory $x(\cdot)$. Using these control paths and the state trajectory one can compute the payoff for player i in $\Gamma_I(x_0, 0)$ from (6.1). We are now ready to define the Nash equilibrium concept for a non-Markovian differential game.

Definition 6.1 A Nash equilibrium for the differential game $\Gamma_I(x_0, 0)$ is a strategy profile $\phi = (\phi^1, \phi^2, \ldots, \phi^N) \in \mathcal{S}_I$ such that for all $i \in \{1, 2, \ldots, N\}$ and for all strategies $\psi^i \in \mathcal{S}_I^i(\phi^{-i})$ it holds that $J^i(\phi^i; \phi^{-i}, x_0, 0) \geq J^i(\psi^i; \phi^{-i}, x_0, 0)$.

One reason (certainly not the least important one) for the emphasis on Markovian equilibria in applications of differential game theory is analytical tractability. Recall that the optimization problem of any player, given that his opponents use Markovian strategies, is a standard optimal control problem as described in chapter 3. The player maximizes his own objective functional over the set of feasible control paths. In a non-Markovian Nash equilibrium the situation is different for at least two important reasons.

First, if at least one opponent of player i uses a non-Markovian strategy, the optimization problem of player i is, in general, no longer a standard optimal control problem because the opponents' strategies depend on variables other than t and $x(t)$. Consider, for example, a differential game with two players and assume that player 1 has delayed state information $I^1(u(\cdot), t) = x(t - \delta)$. In this case player 2 faces the system dynamics

$$\dot{x}(t) = f(x(t), \phi^1(x(t - \delta), t), u^2(t), t).$$

In contrast to the Markovian case, the right-hand side of the system dynamics depends on $x(t - \delta)$ because this variable is an argument of player 1's strategy. The integrand of the objective functional of player i depends on the variable $x(t - \delta)$ for the same reason. It follows that the HJB equation and the maximum principle from chapter 3 are not applicable for solving such a problem. Although similar optimization techniques exist for some modifications of standard optimal control problems, they are substantially more difficult to apply, especially in the framework of a differential game, where several interconnected optimization problems have to be solved simultaneously.

The second reason for the increased mathematical difficulty of handling non-Markovian Nash equilibria (as compared to Markovian Nash equilibria) is even more fundamental. According to definition 6.1, player i does not maximize his objective functional over the set of control paths but over the set of feasible responses $\phi^i \in S^i_I(\phi^{-i})$. In other words, he maximizes over a set of strategies. From a mathematical point of view, this is quite different from the optimal control framework where maximization takes place over a set of control paths. It is for these reasons that we do not present any general equilibrium conditions for non-Markovian Nash equilibria.

We now define one particular nonanticipating and regular information structure H (the letter H stands for 'history') to which we restrict attention in the rest of the chapter. To begin with recall that, if H is nonanticipating, then $H(u(\cdot), t)$ depends only on the restriction of $u(\cdot)$ to the interval $[0, t)$. Let us call this the t-truncation of $u(\cdot)$ and denote it by

$u_t(\cdot)$. We say that two t-truncations $u_t(\cdot)$ and $v_t(\cdot)$ are equivalent if the set $\{s \in [0,t) \mid u_t(s) \neq v_t(s)\}$ has Lebesgue measure 0. This defines an equivalence relation on the set of t-truncations. The information structure H can now be defined by saying that, for all $i \in \{1,2,\ldots,N\}$ and all $t \in [0,T\rangle$, $H^i(u(\cdot),t)$ is the equivalence class to which $u_t(\cdot)$ belongs. It is common to denote this equivalence class also by $u_t(\cdot)$ so that we may write $H^i(u(\cdot),t) = u_t(\cdot)$. We shall refer to the equivalence class $u_t(\cdot)$ as the t-history of the game.[2] By definition, H is nonanticipating and regular. A differential game which uses this information structure is called a differential game with history-dependent strategies. Following our notation for the general case we indicate the information structure defined by t-histories by the subscript H. That is, the game where the information available to player i at time t is the t-history is denoted by $\Gamma_H(x_0,0)$, and the set of feasible strategies is denoted by $\mathcal{S}_H$. Furthermore, we denote the set of all possible t-histories of the differential game $\Gamma_H(x_0,0)$ by $\mathcal{U}_t = \{u_t(\cdot) \mid u(\cdot) \in \mathcal{U}\}$. Note that 0-histories and $\mathcal{U}_0$ cannot be defined in that way. Therefore, we formally set $\mathcal{U}_0 = \{x_0\}$, which makes the notation introduced below consistent with the one used in earlier chapters of the book.

We proceed by defining subgames of $\Gamma_H(x_0,0)$. Let $\tilde{u}(\cdot) \in \mathcal{U}$ be a given N-tuple of feasible paths. For every $t \in [0,T\rangle$, the subgame $\Gamma_H(\tilde{u}_t(\cdot),t)$ is defined by the objective functionals

$$\int_t^T e^{-r^i(s-t)} F^i(x(s),u^1(s),u^2(s),\ldots,u^N(s),s)\,ds + e^{-r^i(T-t)}S^i(x(T)),$$

$$i = 1,2,\ldots,N,$$

the system dynamics (6.2), and the constraints (6.3), whereby (6.2) and (6.3) are only required to hold for $s \in [t,T\rangle$. The initial state for the system dynamics of $\Gamma_H(\tilde{u}_t(\cdot),t)$ is given by $\tilde{x}(t)$, which is uniquely determined by equation (6.2) with $u^i(s)$ replaced by $\tilde{u}^i_t(s)$ for all $s \in [0,t)$ and all $i \in \{1,2,\ldots,N\}$. The information available to every player at time $s \in [t,T\rangle$ is the s-history $\hat{u}_s(\cdot)$, which is given by

$$\hat{u}_s(\tau) = \begin{cases} \tilde{u}_t(\tau) & \text{if } \tau \in [0,t), \\ u(\tau) & \text{if } \tau \in [t,s). \end{cases}$$

Moreover, the objective functional of player i in the subgame $\Gamma(\tilde{u}_t(\cdot),t)$ will be denoted by $J^i(\phi^i;\phi^{-i},\tilde{u}_t(\cdot),t)$.

[2]Readers who are becoming confused at this point may proceed by assuming that control paths are continuous functions, in which case the equivalence class to which $u_t(\cdot)$ belongs consists of a single element, namely the t-truncation $u_t(\cdot)$ itself. In this case the terms 't-truncation' and 't-history' are synonyms.

We can now generalize the concepts of time consistency and subgame perfection to differential games with history-dependent strategies.

Definition 6.2 Let $\Gamma_H(x_0, 0)$ be a differential game with history-dependent strategies and let $\phi = (\phi^1, \phi^2, \ldots, \phi^N)$ be a Nash equilibrium with corresponding N-tuple of control paths $u(\cdot)$.

(i) The Nash equilibrium ϕ is called time consistent if, for all $t \in [0, T\rangle$, ϕ is also a Nash equilibrium for the subgame $\Gamma_H(u_t(\cdot), t)$.
(ii) The Nash equilibrium ϕ is called subgame perfect if, for all $t \in [0, T\rangle$ and all $\tilde{u}_t(\cdot) \in \mathcal{U}_t$, ϕ is also a Nash equilibrium for the subgame $\Gamma_H(\tilde{u}_t(\cdot), t)$.

These definitions are the natural generalizations of the corresponding definitions in the case of the Markovian information structure. The difference between time consistency and subgame perfection is the same as before: whereas time consistency only requires that no player wants to revise his strategy choice during the game, subgame perfection also assumes that there are no incentives to deviate from the equilibrium strategy off the equilibrium path, that is, in subgames defined by t-histories not corresponding to the equilibrium profile. The following result generalizes theorem 4.3 to history-dependent Nash equilibria.

Theorem 6.1 *Every Nash equilibrium in a differential game with history-dependent strategies is time consistent.*

Proof The proof is essentially the same as the one for theorem 4.3. Let ϕ be a Nash equilibrium with corresponding N-tuple of control paths $u(\cdot) \in \mathcal{U}$ and assume that ϕ is not time consistent. Then there exists a time $t \in [0, T\rangle$ and a player $i \in \{1, 2, \ldots, N\}$ such that $J^i(\tilde{\phi}^i; \phi^{-i}, u_t(\cdot), t) > J^i(\phi^i; \phi^{-i}, u_t(\cdot), t)$ holds for some $\tilde{\phi}^i \in \mathcal{S}^i_H(\phi^{-i})$. Now consider the compound strategy

$$\pi^i(\cdot, s) = \begin{cases} \phi^i(\cdot, s) & \text{if } s \in [0, t), \\ \tilde{\phi}^i(\cdot, s) & \text{if } s \in [t, T\rangle. \end{cases}$$

Because the information structure is nonanticipating, switching to $\tilde{\phi}^i$ at time t does not affect the s-histories $u_s(\cdot)$ for any $s < t$. From this fact and the integral form of the objective functionals we conclude that $J^i(\pi^i; \phi^{-i}, x_0, 0) > J^i(\phi^i; \phi^{-i}, x_0, 0)$ must hold. Since this is a contradiction to the assumption that ϕ is a Nash equilibrium, the theorem is proved. ■

6.2 Acceptable profiles and effective threats

Trigger strategies are a particular class of non-Markovian strategies which have received much attention in the game theoretic literature.[3] More specifically, in a trigger strategy equilibrium the players agree to follow a certain target path and sustain this agreement by threatening to punish any defector. In order that this qualifies as a Nash equilibrium, every player has to accept the target path and the threats must be effective so as to prevent any defections. We now discuss these properties in the framework of the differential game $\Gamma_H(x_0, 0)$ from the previous section under the additional assumption that $T = \infty$. For simplicity, let us assume that $r^i > 0$ holds for all $i \in \{1, 2, \dots, N\}$ and that the utility functions F^i are bounded and continuous. This ensures that the integrals in (6.1) are finite and that all optimality criteria introduced in section 3.6 are equivalent.

Let $\tilde{u}_t(\cdot) \in \mathcal{U}$ be an arbitrary but fixed N-tuple of control paths. The upper value of the subgame $\Gamma_H(\tilde{u}_t(\cdot), t)$ for player i is defined as

$$D^i(\tilde{u}_t(\cdot), t) = \inf\left\{\sup\left\{J^i(\phi^i; \phi^{-i}, \tilde{u}_t(\cdot), t) \,\middle|\, \phi^i \in \mathcal{S}^i_H(\phi^{-i})\right\} \,\middle|\, \phi^{-i} \in \mathcal{S}^{-i}_H\right\}.$$

$D^i(\tilde{u}_t(\cdot), t)$ is the lowest payoff player i can ensure for himself in the subgame $\Gamma_H(\tilde{u}_t(\cdot), t)$ if his opponents announce their strategies before the start of the game. One could also say that $D^i(\tilde{u}_t(\cdot), t)$ is the highest payoff player i can expect in $\Gamma_H(\tilde{u}_t(\cdot), t)$ if his opponents try to minimize his objective functional J^i.

Definition 6.3 Consider a feasible strategy profile $\tilde{\phi} = (\tilde{\phi}^1, \tilde{\phi}^2, \dots, \tilde{\phi}^N) \in \mathcal{S}$ for the game $\Gamma_H(x_0, 0)$ and denote by $\tilde{u}(\cdot)$ the N-tuple of control paths corresponding to this profile. We shall refer to $\tilde{\phi}$ as the target profile and to $\tilde{u}(\cdot)$ as the target path. The target profile (or path) is said to be ϵ-acceptable by player i if

$$J^i(\tilde{\phi}^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) \geq D^i(\tilde{u}_t(\cdot), t) + \epsilon \tag{6.5}$$

holds for all $t \in [0, \infty)$. Here, ϵ is a nonnegative constant. If the target profile (or path) is ϵ-acceptable by all players $i \in \{1, 2, \dots, N\}$ then we say that it is ϵ-acceptable. The terms '0-acceptable' and 'acceptable' will be used synonymously.

To see that the concepts introduced in definition 6.3 make sense, suppose that the target path $\tilde{u}(\cdot)$ is not acceptable. In this case there exists a time t and a player i such that (6.5) does not hold with $\epsilon = 0$. Thus, it must hold

[3]For a motivation we refer to section 2.4.

that $\gamma := D^i(\tilde{u}_t(\cdot), t) - J^i(\tilde{\phi}^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) > 0$. Player i can ensure that he does better than along the target path by switching at time t to a strategy $\phi^i \in \mathcal{S}_H^i(\tilde{\phi}^{-i})$ which yields a payoff $J^i(\phi^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) > D^i(\tilde{u}_t(\cdot), t) - \gamma$ for the subgame $\Gamma_H(\tilde{u}_t(\cdot), t)$. This is possible by the definition of the upper value $D^i(\tilde{u}_t(\cdot), t)$ and it does indeed improve player i's objective functional because of the definition of γ. Therefore, if $\tilde{u}(\cdot)$ is not acceptable to player i, one cannot expect player i to agree to the target path $\tilde{u}(\cdot)$.

Theorem 6.2 *Let $\Gamma_H(x_0, 0)$ be a differential game with history-dependent strategies. Then every Nash equilibrium is acceptable.*

Proof Assume $\tilde{\phi}$ is a Nash equilibrium and denote by $\tilde{u}(\cdot)$ the corresponding N-tuple of control paths. By theorem 6.1 we know that $\tilde{\phi}$ is time consistent. This implies that

$$J^i(\tilde{\phi}^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) = \sup\left\{J^i(\phi^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) \,\middle|\, \phi^i \in \mathcal{S}_H^i(\tilde{\phi}^{-i})\right\} \quad (6.6)$$

for all $t \in [0, \infty)$ and all $i \in \{1, 2, \dots, N\}$. Moreover, we know from the definition of the upper values that $D^i(\tilde{u}_t(\cdot), t) \leq \sup\{J^i(\phi^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) \mid \phi^i \in \mathcal{S}_H^i(\tilde{\phi}^{-i})\}$ holds for all $t \in [0, \infty)$ and all $i \in \{1, 2, \dots, N\}$. Together with (6.6), this proves the theorem.■

We now turn to the discussion of threats that can be used to enforce a given target profile $\tilde{\phi} \in \mathcal{S}_H$. At this point it is unimportant whether $\tilde{\phi}$ is acceptable or not. At any time instant $s' \in [0, \infty)$ player $i \in \{1, 2, \dots, N\}$ can decide whether to cooperate and continue to play his target strategy $\tilde{\phi}^i$ or to defect by deviating from $\tilde{\phi}^i$.[4]

In this section we assume that the players react to a defection by any opponent with a fixed positive time delay $\delta > 0$.[5] That is to say, if a player defects at time s', his opponents will start to punish him at time $s = s' + \delta$. It is furthermore assumed that the punishment lasts for ever. Under these assumptions a trigger strategy for player j with target profile $\tilde{\phi}$ can be defined as follows:

[4]By defecting at time s' we mean that player i chooses a strategy ϕ^i such that s' is the supremum over all nonnegative numbers τ for which the set $\{t \in [0, \tau) | \tilde{\phi}^i(\tilde{u}_t(\cdot), t) \neq \phi^i(u_t(\cdot), t)\}$ has Lebesgue measure 0. Here, $\tilde{u}(\cdot)$ and $u(\cdot)$ are the N-tuples of control paths corresponding to $(\tilde{\phi}^1, \tilde{\phi}^2, \dots, \tilde{\phi}^N)$ and $(\tilde{\phi}^1, \tilde{\phi}^2, \dots, \tilde{\phi}^{i-1}, \phi^i, \tilde{\phi}^{i+1}, \dots, \tilde{\phi}^N)$, respectively. Deviating from $\tilde{\phi}^i$ on a set of measure 0 does not count as a defection since it affects neither the information, nor the evolution of the state variable, nor the objective functional of any player.

[5]See section 6.4 for a discussion of this assumption and for an alternative framework in which defectors are punished immediately after their defection.

$$\psi^j(\cdot, t) = \begin{cases} \tilde{\phi}^j(\cdot, t) & \text{if no player has defected at or before time } t - \delta, \\ \theta^j(\cdot, t) & \text{if a defection has occurred at or before time } t - \delta. \end{cases} \tag{6.7}$$

Here $\theta = (\theta^1, \theta^2, \ldots, \theta^N) \in \mathcal{S}_H$ is a strategy profile which we call the threats or the punishment strategies. The interpretation of ψ^j is as follows. Player j starts by cooperating and playing the target strategy $\tilde{\phi}^j$ as long as no one defects. Exactly δ time units after the first defection has occurred, player j switches to the punishment strategy θ^j.

In the remainder of this section we state conditions under which the trigger strategies defined in (6.7) constitute a Nash equilibrium. The basic idea is to ensure that the threats are effective so that no player will ever defect. It is worth emphasizing that in this case the punishment is never executed along the equilibrium path.

Denote the target path corresponding to the target profile $\tilde{\phi}$ by $\tilde{u}(\cdot)$. Now consider the decision problem of player i at time t under the assumption that before time t no player has defected. He can either continue to cooperate, in which case his discounted utility over the remaining time horizon is $J^i(\tilde{\phi}^i(\cdot); \tilde{\phi}^{-i}(\cdot), \tilde{u}_t(\cdot), t)$, or he can defect at time t. In the latter case his defection will not lead to a reaction by his opponents until time $t + \delta$. The period $[t, t + \delta)$ is called the cheating period. Thereafter, at time $t + \delta$, the punishment period begins during which his opponents employ their threat strategies θ^{-i}. Punishment is assumed to last for ever. If we denote by ϕ^i player i's defection strategy then we can write his discounted utility over the time interval $[t, \infty)$ in the case of defection as

$$\begin{aligned} J^i_{\mathrm{DEF}}(\phi^i, \tilde{u}_t(\cdot), t) = \int_t^{t+\delta} & e^{-r^i(s-t)} F^i(\hat{x}(s), \hat{u}(s), s)\, ds \\ & + e^{-r^i\delta} J^i(\phi^i; \theta^{-i}, \hat{u}_{t+\delta}(\cdot), t + \delta), \end{aligned} \tag{6.8}$$

where $\hat{u}(\cdot)$ and $\hat{x}(\cdot)$ are the N-tuple of control paths and the state trajectory, respectively, corresponding to the strategy profile

$$(\tilde{\phi}^1, \tilde{\phi}^2, \ldots, \tilde{\phi}^{i-1}, \phi^i, \tilde{\phi}^{i+1}, \ldots, \tilde{\phi}^N).$$

The first term on the right-hand side of (6.8) is the utility derived during the cheating period and the second term is the utility derived during the punishment period. For simplicity we did not make the dependence of J^i_{DEF} on the target profile $\tilde{\phi}$ and the threats θ^{-i} explicit through the notation. We emphasize, however, that J^i_{DEF} depends on the history at the time of defection, $\tilde{u}_t(\cdot)$.

We can now state a first important result on trigger strategies.

Theorem 6.3 *Let $\tilde{\phi} \in \mathcal{S}_H$ be a given target profile for the game $\Gamma_H(x_0, 0)$ and let $\tilde{u}(\cdot)$ be the corresponding target path.*

(i) The strategy profile $\psi = (\psi^1, \psi^2, \ldots, \psi^N)$ defined in (6.7) constitutes a Nash equilibrium for the game $\Gamma_H(x_0, 0)$ if and only if

$$J^i_{\mathrm{DEF}}(\phi^i, \tilde{u}_t(\cdot), t) \leq J^i(\tilde{\phi}^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) \tag{6.9}$$

holds for all $i \in \{1, 2, \ldots, N\}$, all $t \in [0, \infty)$, and all feasible defection paths $\phi^i \in \mathcal{S}^i_H(\tilde{\phi}^{-i})$.

(ii) Assume that for all $i \in \{1, 2, \ldots, N\}$ and all $u(\cdot) \in \mathcal{U}$ the mapping $t \mapsto D^i(u_t(\cdot), t)$ is continuous from the right and that there exists $\epsilon > 0$ such that $\tilde{\phi}$ is ϵ-acceptable. If the delay δ is sufficiently small then it is possible to find threats $\theta \in \mathcal{S}_H$ such that the trigger strategy profile ψ defined in (6.7) constitutes a Nash equilibrium for the game $\Gamma_H(x_0, 0)$.

Proof (i) This assertion is quite obvious. If player i does not defect at time t he derives the utility $J^i(\tilde{\phi}^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t)$ over the interval $[t, \infty)$. If he defects at time t by switching to ϕ^i he will get $J^i_{\mathrm{DEF}}(\phi^i, \tilde{u}_t(\cdot), t)$. Obviously, player i will defect at time t (and thus the trigger strategy profile does not constitute a Nash equilibrium) if and only if (6.9) is violated.

(ii) Denote the first term on the right-hand side of (6.8) by $z^i(\delta)$ and note that boundedness of F^i implies that $\lim_{\delta \to 0} z^i(\delta) = 0$. Now consider the second term on the right-hand side of (6.8). Because of the definition of $D^i(\hat{u}_{t+\delta}(\cdot), t + \delta)$ it is possible to find feasible threat strategies $\theta^{-i} \in \mathcal{S}^{-i}_H$ such that

$$J^i(\phi^i; \theta^{-i}, \hat{u}_{t+\delta}(\cdot), t + \delta) < D^i(\hat{u}_{t+\delta}(\cdot), t + \delta) + \epsilon/2$$

holds for all feasible defection strategies $\phi^i \in \mathcal{S}^i_H(\theta^{-i})$. Furthermore, since $\tilde{\phi}$ is ϵ-acceptable, we know that the inequality $J^i(\tilde{\phi}^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) \geq D^i(\tilde{u}_t(\cdot), t) + \epsilon$ holds for all $i \in \{1, 2, \ldots, N\}$ and all $t \in [0, \infty)$. Taking all together, it follows that a sufficient condition for the inequality in (6.9) to hold is that

$$z^i(\delta) + e^{-r^i\delta}[D^i(\hat{u}_{t+\delta}(\cdot), t + \delta) + \epsilon/2] < D^i(\tilde{u}_t(\cdot), t) + \epsilon.$$

The continuity assumption stated in the theorem implies that in the limit, as δ approaches 0, this inequality becomes $D^i(\tilde{u}_t(\cdot), t) + \epsilon/2 \leq D^i(\tilde{u}_t(\cdot), t) + \epsilon$, which is obviously true. Therefore, condition (6.9) is satisfied when δ is sufficiently close to 0 and we conclude from part (i) of the theorem that the threats θ are strong enough to support the target path $\tilde{\phi}$ as a Nash equilibrium. ■

The first statement in the theorem is a general necessary and sufficient condition for the equilibrium property of trigger strategies. The second statement proves that ϵ-acceptable target paths can always be supported as Nash equilibria if the delay δ is sufficiently short. The additional continuity assumption on the upper values D^i is of a purely technical nature and can usually be directly verified (see, e.g., example 6.1). It holds under quite mild regularity conditions on the problem data, such as smoothness of the functions F^i, f, and U^i. The proof of theorem 6.3(ii) does not work for 0-acceptable target paths but requires the stronger property of ϵ-acceptability with $\epsilon > 0$. This is due to the existence of the fixed positive delay $\delta > 0$. We shall see in section 6.4 that, when δ is infinitesimally small, one can also support 0-acceptable target paths as trigger strategy equilibria.

The following example illustrates the results of this section.

Example 6.1 Consider a one-sector economy in which there are two types of agents: workers and capitalists. The capital–labour ratio at time t is denoted by $x(t)$ and serves as the state variable of the model. Per-capita output at time t is $f(x(t))$ where f is a standard neoclassical production function satisfying $f(0) = 0, f'(x) > 0$ and $f''(x) < 0$ for all $x \in (0, \infty)$, as well as the Inada conditions $\lim_{x\to 0} f'(x) = \infty$ and $\lim_{x\to\infty} f'(x) = 0$.

The workers run the government, which can redistribute income between the two classes by means of lump sum taxes and transfers denoted by $u^1(t)$. If $u^1(t)$ is positive it represents a transfer from the capitalists to the working class and if $u^1(t)$ is negative it is a transfer from the workers to the capitalists. The net income of workers consists of their wage income plus the transfer payments, $f(x(t)) - x(t)f'(x(t)) + u^1(t)$. It is assumed that workers don't save, thus their income equals consumption. Under the assumption that the utility function is linear with respect to consumption the objective functional of the workers (or, equivalently, of the government) can be written as

$$J^1(u^1(\cdot); u^2(\cdot), x_0, 0) = \int_0^\infty e^{-rt}[f(x(t)) - x(t)f'(x(t)) + u^1(t)]\, dt,$$

where $r > 0$ is the time preference rate and $u^2(\cdot)$ is the capitalists' control path described below.

Capitalists act collectively. Their profits after tax are given by the marginal product of capital minus $u^1(t)$. Denoting the capitalists' savings rate at time t by $u^2(t)$ we can therefore write their consumption at time t as $[1 - u^2(t)][x(t)f'(x(t)) - u^1(t)]$. If we assume that capitalists, too, have a linear utility function and make the additional assumption that workers and capitalists have the same time preference rate r then we can write the capitalists' objective functional as

$$J^2(u^2(\cdot); u^1(\cdot), x_0, 0) = \int_0^\infty e^{-rt}[1 - u^2(t)][x(t)f'(x(t)) - u^1(t)]\, dt.$$

The system dynamics are given by

$$\dot{x}(t) = u^2(t)[x(t)f'(x(t)) - u^1(t)] - \gamma x(t), \quad x(0) = x_0, \tag{6.10}$$

where $\gamma > 0$ denotes the population growth rate and $x_0 > 0$ is the initial state at time 0. Finally, the control constraints are given by

$$\begin{aligned} &u^1(t) \in [-f(x(t)) + x(t)f'(x(t)), x(t)f'(x(t))], \\ &u^2(t) \in [0, 1]. \end{aligned} \tag{6.11}$$

The constraint concerning the capitalists' savings rate $u^2(t)$ is obvious. The constraint concerning the lump sum tax $u^1(t)$ ensures that both workers and capitalists have nonnegative consumption. In the absence of a credit market this is a necessary feasibility condition for any redistributive tax policy.

Before we consider the issue of trigger strategy equilibria for this game let us briefly discuss Pareto optimal solutions which maximize the sum of the objective functionals $\tilde{J} = J^1(u^1(\cdot); u^2(\cdot), x_0, 0) + J^2(u^2(\cdot); u^1(\cdot), x_0, 0)$ subject to (6.10) and (6.11). Because of the simple (linear) structure of the model one can find these solutions without the maximum principle or the HJB equation. Indeed, by using the state equation (6.10) we obtain $\tilde{J} = \int_0^\infty e^{-rt}[f(x(t)) - \gamma x(t) - \dot{x}(t)]\, dt$. Integrating the term containing $\dot{x}(t)$ and using the initial condition $x(0) = x_0$ yields[6]

$$\tilde{J} = x_0 + \int_0^\infty e^{-rt}[f(x(t)) - (\gamma + r)x(t)]\, dt.$$

The integrand in this representation of the objective functional $\tilde{J}$ is a strictly concave function of the state variable $x(t)$, which attains its maximum at the unique value $x(t) = \bar{x}$ defined by the equation $f'(\bar{x}) = \gamma + r$. From these properties it follows immediately that, to maximize $\tilde{J}$, the state trajectory must reach the so-called golden-rule level $\bar{x}$ as fast as possible and remain there for ever. Using (6.10) it is straightforward to verify that this is the case if and only if the controls are chosen in such a way that

[6]This step uses also the property that every feasible state trajectory must remain bounded, which follows from the Inada conditions and the positive population growth rate $\gamma > 0$.

$$u^2(t)[x(t)f'(x(t)) - u^1(t)] = \begin{cases} f(x(t)) & \text{if } x(t) < \bar{x}, \\ \gamma\bar{x} & \text{if } x(t) = \bar{x}, \\ 0 & \text{if } x(t) > \bar{x}. \end{cases} \tag{6.12}$$

The definition of $\bar{x}$, concavity of f, and $f(0) = 0$ imply that $f(\bar{x}) > \gamma\bar{x}$ such that the interval $(-f(\bar{x}) + (\gamma + r)\bar{x}, r\bar{x}]$ is nonempty. For every β in this interval consider the stationary Markovian strategies $\tilde{\phi}^1_\beta$ and $\tilde{\phi}^2_\beta$ where

$$\tilde{\phi}^1_\beta(x) = \begin{cases} -f(x) + xf'(x) & \text{if } x < \bar{x}, \\ \beta & \text{if } x = \bar{x}, \\ xf'(x) & \text{if } x > \bar{x}, \end{cases}$$

$$\tilde{\phi}^2_\beta(x) = \begin{cases} 1 & \text{if } x < \bar{x}, \\ \gamma\bar{x}/[(\gamma + r)\bar{x} - \beta] & \text{if } x = \bar{x}, \\ 0 & \text{if } x > \bar{x}. \end{cases}$$

It is easy to show that, for $\beta \in (-f(\bar{x}) + (\gamma + r)\bar{x}, r\bar{x}]$, these strategies are feasible, that is, $(\tilde{\phi}^1_\beta, \tilde{\phi}^2_\beta) \in \mathcal{S}_H$. Moreover, if $u^1(\cdot)$ and $u^2(\cdot)$ are the control paths corresponding to $\tilde{\phi}^1_\beta$ and $\tilde{\phi}^2_\beta$, respectively, then (6.12) is satisfied so that the pair $(\tilde{\phi}^1_\beta, \tilde{\phi}^2_\beta)$ is Pareto efficient.[7]

Note that in the Pareto optimal solution $(\tilde{\phi}^1_\beta, \tilde{\phi}^2_\beta)$ neither the workers nor the capitalists consume anything if the capital–labour ratio is smaller than $\bar{x}$. They postpone their consumption until the capital–labour ratio reaches its steady state value $\bar{x}$. Also note that for $x \neq \bar{x}$ the strategies are independent of the parameter β, which implies that the state trajectory $\tilde{x}(\cdot)$ generated by $(\tilde{\phi}^1_\beta, \tilde{\phi}^2_\beta)$ is independent of β.

From now on we assume that the initial state x_0 is positive but smaller than the golden-rule per-capita capital stock $\bar{x}$. This implies that for the Pareto optimal solutions described above it will always be the case that the state satisfies $\tilde{x}(t) \leq \bar{x}$ for all $t \in [0, \infty)$. More specifically, we have $\dot{\tilde{x}}(t) = f(\tilde{x}(t)) - \gamma\tilde{x}(t)$ and $\tilde{x}(t) < \bar{x}$ for $t < \tau(x_0)$, and $\tilde{x}(t) = \bar{x}$ for $t \geq \tau(x_0)$. Here

$$\tau(x_0) = \int_{x_0}^{\bar{x}} \frac{dz}{f(z) - \gamma z}$$

is the finite time at which the golden-rule capital stock is reached from the initial state x_0. Using this property it is easy to calculate the payoffs over

[7] $\tilde{\phi}^1_\beta$ and $\tilde{\phi}^2_\beta$ are not the only strategies with this property. As a matter of fact, all one needs for (6.12) to be satisfied in the case $x > \bar{x}$ is that either $\tilde{\phi}^1_\beta(x) = xf'(x)$ or $\tilde{\phi}^2_\beta(x) = 0$. We do not consider all strategies with this property because, in the sequel, we shall restrict ourselves to initial states $x_0 \leq \bar{x}$ for which the case $x > \bar{x}$ never occurs in equilibrium.

the period $[t, \infty)$ for both workers and capitalists under the assumption that the Pareto optimal strategy profile $(\tilde{\phi}^1_\beta, \tilde{\phi}^2_\beta)$ is used. More specifically, we have

$$\begin{aligned} J^1(\tilde{\phi}^1_\beta; \tilde{\phi}^2_\beta, \tilde{u}_t(\cdot), t) &= (1/r)e^{-r\tau(\tilde{x}(t))}[f(\bar{x}) - (\gamma + r)\bar{x} + \beta], \\ J^2(\tilde{\phi}^2_\beta; \tilde{\phi}^1_\beta, \tilde{u}_t(\cdot), t) &= (1/r)e^{-r\tau(\tilde{x}(t))}[r\bar{x} - \beta]. \end{aligned} \tag{6.13}$$

Our next goal is to compute the upper values $D^i(\tilde{u}_t(\cdot), t)$ for both players $i \in \{1, 2\}$. To this end first note that, if the workers choose the strategy $u^1(s) = x(s)f'(x(s))$, they obtain the utility

$$J^1(u^1(\cdot); u^2(\cdot), \tilde{u}_t(\cdot), t) = \int_t^\infty e^{-r(s-t)} f(\tilde{x}(t)e^{-\gamma(s-t)})\, ds = \int_0^\infty e^{-rs} f(\tilde{x}(t)e^{-\gamma s})\, ds$$

independently of the capitalists' strategy. This implies that

$$\sup\left\{J^1(\phi^1; \phi^2, \tilde{u}_t(\cdot), t) \,\middle|\, \phi^1 \in \mathcal{S}^1_H(\phi^2)\right\} \geq \int_0^\infty e^{-rs} f(\tilde{x}(t)e^{-\gamma s})\, ds$$

holds for all $\phi^2 \in \mathcal{S}^{-1}_H$. On the other hand, by choosing $u^2(s) = 0$ for all $s \in [t, \infty)$ the capitalist can ensure that the workers do not get higher utility than that and we therefore conclude that $D^1(\tilde{u}_t(\cdot), t) = \int_0^\infty e^{-rs} f(\tilde{x}(t)e^{-\gamma s})\, ds$. The calculation of $D^2(\tilde{u}_t(\cdot), t)$ is even simpler. As a matter of fact, $J^2(u^2(\cdot); u^1(\cdot), \tilde{u}_t(\cdot), t)$ is always nonnegative because of the control constraints (6.11). Since $J^2(u^2(\cdot); u^1(\cdot), \tilde{u}_t(\cdot), t)$ is equal to 0 in the case where $u^1(s) = x(s)f'(x(s))$ holds for all $s \in [t, \infty)$, it follows that $D^2(\tilde{u}_t(\cdot), t) = 0$. Using these results and (6.13) we conclude that the Pareto efficient strategy profile $(\tilde{\phi}^1_\beta, \tilde{\phi}^2_\beta)$ is ϵ-acceptable if and only if the two inequalities

$$\begin{aligned} &(1/r)e^{-r\tau(\tilde{x}(t))}[f(\bar{x}) - (\gamma + r)\bar{x} + \beta] \geq \int_0^\infty e^{-rs} f(\tilde{x}(t)e^{-\gamma s})\, ds + \epsilon, \\ &(1/r)e^{-r\tau(\tilde{x}(t))}[r\bar{x} - \beta] \geq \epsilon \end{aligned} \tag{6.14}$$

hold for all $t \in [0, \infty)$.[8] Without specifying the functional form of the production function f, condition (6.14) is quite difficult to analyse. However, it can be used to show that there exist Pareto optimal solutions which are ϵ-acceptable for some $\epsilon > 0$. For example, consider the case where the initial state is equal to the golden-rule, i.e., $x_0 = \bar{x}$. Obviously, this implies $\tau(x_0) = 0$ and $\tilde{x}(t) = \bar{x}$ for all $t \in [0, \infty)$. It follows that (6.14) holds for some positive ϵ if and only if

[8] Recall that $\tilde{x}(t)$ is the value of the state variable at time t in the Pareto optimal solution. In other words, $\tilde{x}(t)$ is equal to $\bar{x}$ for $t \geq \tau(x_0)$ and it is defined by the initial value problem $\dot{x}(t) = f(x(t)) - \gamma x(t)$, $x(0) = x_0$ for $t < \tau(x_0)$.

$$(r+\gamma)\bar{x} - f(\bar{x}) + r\int_0^\infty e^{-rs} f(\bar{x}e^{-\gamma s})\,ds < \beta < r\bar{x}.$$

A parameter β satisfying this inequality exists if and only if

$$f(\bar{x}) - \gamma\bar{x} > r\int_0^\infty e^{-rs} f(\bar{x}e^{-\gamma s})\,ds. \tag{6.15}$$

Condition (6.15) is therefore sufficient for the existence of an ϵ-acceptable and Pareto optimal target path. The continuity assumptions for $D^i(\tilde{u}_t(\cdot), t)$ used in theorem 6.3 are trivially satisfied in this example and we can therefore conclude that under (6.15) there exist Pareto optimal solutions which can be supported as trigger strategy Nash equilibria (see exercise 1 at the end of this chapter for an additional discussion of this example).

6.3 Credible threats

The threats in a trigger strategy equilibrium are intended to prevent defections from the target path. In equilibrium, defections do not occur and the punishment is never executed. For this reason there is only one condition to be satisfied by threats in order for a trigger strategy profile to constitute a Nash equilibrium: the threats must be effective as described by condition (6.9). However, in many situations the most effective threats may not be credible. For example, firms in an oligopolistic market may threaten a potential defector by announcing that they would give away their products for free, thereby leaving no demand for the defector. Although this is certainly a very effective threat it is not credible because it would create losses for the punishing firms. Any rational firm considering the possibility of defection would anticipate that the punishment by the opponents will not be carried out (at least not for very long time).

Consider a trigger strategy equilibrium ψ with target profile $\tilde{\phi}$ and threats θ as in (6.7). In order for a threat to be credible it has to be in the punisher's own best interest to carry out the punishment once a defection has occurred. Now assume that some player i defects at time t. By construction of the trigger strategies ψ, all players switch to their punishment strategies $\theta^j, j \in \{1, 2, \ldots, N\}$, at time $t+\delta$.[9] Saying that this is in all players' own interest is tantamount to saying that the strategy profile θ itself constitutes a Nash equilibrium of the subgame that starts at time $t+\delta$. We can therefore conclude that a necessary condition for threats to be credible is that they constitute a Nash equilibrium of any

[9]Note that the defector also switches.

subgame that can arise δ time units after a defection. This is obviously the case if θ is a subgame perfect Nash equilibrium as defined in definition 6.2. In the present section we discuss trigger strategy equilibria with this property.

Before we continue, it is worth pointing out that subgame perfectness of θ does not make the trigger strategy equilibrium ψ subgame perfect. To see this, note that there are three types of subgames to be considered: subgames before any defection, subgames starting at least δ time periods after a defection, and subgames which are characterized by the fact that a defection has occurred less than δ time periods before the start of the subgame. Let us denote by t the time of the first defection and by $\tilde{u}(\cdot)$ the target path and consider the three types of subgames in turn. If ψ is a trigger strategy equilibrium, then we know from theorem 6.2 that it is time consistent, which implies that it is an equilibrium for any subgame $\Gamma_H(\tilde{u}_s(\cdot), s)$ with $s \in [0, \infty)$. If $s < t$, that is, if no defection has occurred before or at time s, then it follows that $\tilde{u}_s(\cdot)$ is the only possible s-history of the game, which proves that ψ is an equilibrium for all subgames of the first type. In a subgame of the second type, the defection has occurred at least δ time periods ago so that everyone has already switched to his punishment strategy θ^j. Since θ is by assumption subgame perfect, it follows that ψ is a Nash equilibrium of any subgame in the second category, too. Finally, consider a subgame of the third type, one that starts at time $s \in (t, t+\delta)$. In such a subgame all players know that a defection has occurred because $u_s(\cdot)$ is different from $\tilde{u}_s(\cdot)$. Despite this knowledge, the players wait by assumption until time $t+\delta$ before they start to punish the defector. It is quite obvious that this is, in general, not an optimal behaviour, because it would be better to start the punishment right away. We can therefore conclude that ψ is, in general, not a Nash equilibrium for subgames of the third type.

The reason why the trigger strategies defined in (6.7) are usually not subgame perfect is apparently the existence of the positive delay δ. Intuitively one would think that subgame perfection can be achieved if it is assumed that punishment starts as soon as a defection occurs. The problem with this assumption is that, in a continuous time model, there does not exist a first time after the defection. A possible way to circumvent this problem will be presented in section 6.4.

Let us now proceed with the discussion of trigger strategies with subgame perfect threats. More specifically, we consider trigger strategies ψ^j, $j \in \{1, 2, \ldots, N\}$, as defined in (6.7) where $\tilde{\phi}$ denotes the target profile and θ the punishment strategy profile. The target path will be denoted by $\tilde{u}(\cdot)$. We assume that θ is a Markov perfect Nash equilibrium of the game $\Gamma(x_0, 0)$.

As in the previous section, we can compute the discounted utility of a defecting agent by decomposing it into the utility derived during the cheating period and that derived during the punishment period. This leads to

$$J^i_{\mathrm{DEF}}(\phi^i, \tilde{u}_t(\cdot), t) = \int_t^{t+\delta} e^{-r^i(s-t)} F^i(\hat{x}(s), \hat{u}(s), s)\, ds + e^{-r^i\delta} J^i(\phi^i; \theta^{-i}, \hat{u}_{t+\delta}(\cdot), t+\delta),$$

where $\hat{u}(\cdot)$ and $\hat{x}(\cdot)$ are the N-tuple of control paths and the state trajectory, respectively, corresponding to the strategy profile

$$(\tilde{\phi}^1, \tilde{\phi}^2, \ldots, \tilde{\phi}^{i-1}, \phi^i, \tilde{\phi}^{i+1}, \ldots, \tilde{\phi}^N),$$

and ϕ^i is player i's defection strategy. Since θ is a Markov perfect Nash equilibrium we know that

$$J^i(\phi^i; \theta^{-i}, \hat{u}_{t+\delta}(\cdot), t+\delta) \leq J^i(\theta^i; \theta^{-i}, \hat{u}_{t+\delta}(\cdot), t+\delta).$$

The term on the right-hand side of this inequality is the optimal value function of player i's optimization problem given the strategies θ^{-i} of his opponents and evaluated at the state $\hat{x}(t+\delta)$ and time $t+\delta$. Following the notation from the previous chapters we denote this function by $V^i(\hat{x}(t+\delta), t+\delta; \theta)$. Using the above inequality and the definition of J^i_{DEF} we obtain

$$J^i_{\mathrm{DEF}}(\phi^i, \tilde{u}_t(\cdot), t) \leq \int_t^{t+\delta} e^{-r^i(s-t)} F^i(\hat{x}(s), \hat{u}(s), s)\, ds + e^{-r^i\delta} V^i(\hat{x}(t+\delta), t+\delta; \theta). \tag{6.16}$$

This yields the following result.

Theorem 6.4 *Let $\tilde{\phi} \in S_H$ be a given target profile for the game $\Gamma_H(x_0, 0)$ and let $\theta \in S_H$ be a Markov perfect Nash equilibrium. Denote the N-tuple of control paths and the state trajectory corresponding to $\tilde{\phi}$ by $\tilde{u}(\cdot)$ and $\tilde{x}(\cdot)$, respectively, and let $\hat{u}(\cdot)$ and $\hat{x}(\cdot)$ be as defined above.*

(i) The trigger strategies defined in (6.7) constitute a Nash equilibrium for the game $\Gamma_H(x_0, 0)$ if

$$\int_t^{t+\delta} e^{-r^i(s-t)} F^i(\hat{x}(s), \hat{u}(s), s)\, ds + e^{-r^i\delta} V^i(\hat{x}(t+\delta), t+\delta; \theta) \leq J^i(\tilde{\phi}^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t)$$

holds for all $i \in \{1, 2, \ldots, N\}$, all $t \in [0, \infty)$, and all defection strategies $\phi^i \in \mathcal{S}^i_H(\tilde{\phi}^{-i})$.

(ii) Assume that the optimal value functions $V^i(x, t; \theta)$ are continuous in x and t and that there exists $\epsilon > 0$ such that

$$J^i(\tilde{\phi}^i; \tilde{\phi}^{-i}, \tilde{u}_t(\cdot), t) \geq V^i(\tilde{x}(t), t; \theta) + \epsilon \tag{6.17}$$

holds for all $i \in \{1, 2, \ldots, N\}$ and all $t \in [0, \infty)$. If the detection delay δ is sufficiently small, then the trigger strategies defined in (6.7) constitute a Nash equilibrium for the game $\Gamma_H(x_0, 0)$.

Proof The proof is analogous to the proof of theorem 6.3. The reader is asked to provide the details in exercise 2.■

The assumption that (6.17) must hold for some strictly positive ϵ is analogous to the assumption of ϵ-acceptability in theorem 6.3. It has to hold as a consequence of the positive delay δ between a defection and the start of punishment. If ϵ in (6.17) could be chosen to be equal to 0, then (6.17) would simply say that the target profile $\tilde{\phi}$ dominates (in the Pareto sense) the Nash equilibrium θ in all subgames along the target path. Theorem 6.4(ii) can thus be reformulated in the following way: any strategy profile $\tilde{\phi}$ for which there exists a Markov perfect Nash equilibrium θ which is strictly and uniformly dominated by $\tilde{\phi}$ in any subgame along the target path can be supported as a trigger strategy equilibrium with credible threats if the delay δ is sufficiently small. The threats are simply the Markov perfect Nash equilibrium θ.

Example 6.1 (continued) Let us see if one can support a Pareto optimal solution in the game between workers and capitalist by credible threats. In order to apply theorem 6.4 we have to find a Markov perfect Nash equilibrium (θ^1, θ^2). For the present model this is quite easy and can be done without using the conditions of theorem 4.4. Indeed, assume that the capitalists use the strategy not to save at all, i.e., $\theta^2(x, t) = 0$ for all $(x, t) \in X \times [0, \infty)$. This implies that the workers' control variable, the tax policy, has no influence on the evolution of the state. Consequently, the best response by the workers to the strategy θ^2 is obviously to transfer as much as possible to themselves, i.e., $\theta^1(x, t) = xf'(x)$ for all $(x, t) \in X \times [0, \infty)$. On the other hand, if the workers use θ^1, then the capitalists get zero utility independently of their own actions. Since this makes them indifferent between all their possible strategies, they might as well choose to save nothing. It follows that θ^1 is a best response to θ^2 and vice versa so that (θ^1, θ^2) qualifies as a Markov perfect Nash equilibrium.

The optimal value functions are also easily computed. Because in equilibrium nothing is saved, one has $\dot{x}(s) = -\gamma x(s)$ which yields $x(s) = x(t)e^{-\gamma(s-t)}$ for all $s \in [t, \infty)$. Therefore, the workers' utility in the Markov perfect Nash equilibrium is given by

$$V^1(x, t; \theta) = \int_t^\infty e^{-r(s-t)} f(xe^{-\gamma(s-t)})\, ds = \int_0^\infty e^{-rs} f(xe^{-\gamma s})\, ds.$$

Furthermore, since capitalists do not consume anything in the Markov perfect Nash equilibrium we obtain

$$V^2(x, t; \theta) = 0.$$

Comparing this with the results derived in our previous discussion of this model in section 6.2 we see that the identity $V^i(\tilde{x}(t), t; \theta) = D^i(\tilde{u}_t(\cdot), t)$ holds for $i = 1, 2$ and for all $t \in [0, \infty)$, which means that in this Markov perfect Nash equilibrium each player harms his opponent as much as possible. In other words, the lack of cooperation leads in this model to the most severe consequences one can think of. Moreover, because $V^i(\tilde{x}(t), t; \theta) = D^i(\tilde{u}_t(\cdot), t)$ it follows that condition (6.17) is identical to the ϵ-acceptability of the target path. Thus, condition (6.14) ensures that there exists a trigger strategy equilibrium with credible threats which supports the Pareto optimal strategy profile $(\tilde{\phi}^1_\beta, \tilde{\phi}^2_\beta)$.

6.4 δ-strategies

In our discussion of trigger strategy equilibria we have been using two important structural assumptions. The first is that the game has an infinite time horizon $T = \infty$, and the second is that there exists a positive delay δ between the defection of a player and the start of punishment by his opponents. Let us start this section by briefly discussing the consequences of dropping these assumptions.

First assume that T is finite but that there is still a positive delay δ between defection and punishment. If player i deviates from the target path during the period $(T - \delta, T]$ his opponents are not able to execute their punishment before the end of the game. Therefore, unless the target path is individually rational for player i (which a Pareto optimal target path, in general, is not), he will defect during the period $(T - \delta, T]$. The same argument applies to all players so that cooperation cannot be sustained at the end of the planning horizon. Now consider the interval $(T - 2\delta, T - \delta]$. If a player deviates from the target strategy during this interval, his opponents would have the chance to punish him but we know already from the previous argument that they are going to play some individually rational strategy during the possible punishment period. Therefore, the threat to punish during $(T-\delta, T]$ is void and all players will behave noncooperatively during $(T-2\delta, T-\delta]$ as well. This argument can be repeated over and over again so that we finally conclude that noncooperative behaviour will prevail throughout the entire interval $[0, T]$.[10] Pareto optimal solutions can usually not be supported as trigger strategy equilibria in this setting. The reader should note that both the finite time horizon and the positive delay δ are crucial for the validity of the backward-induction argument used to demonstrate this result.

Now consider the second main assumption of our previous analysis, namely that there exists a positive delay between defection and punishment. We have already pointed out that this prevents trigger strategy equilibria from being subgame perfect. The reason is related to the so-called Archimedean property of the real numbers which says that there is no smallest time s which is strictly greater than t. If a defection occurs at time t and we assume that the opponents of the defector start punishing him as soon as possible after the defection, then we either have to assume that they start punishing him at the same time as he defects or that there is a positive delay between defection or punishment. Note that starting the punishment at the time of defection requires that the defection can be detected immediately. This is not possible with a nonanticipating and regular information structure. In particular, it is not possible with history-dependent strategies. In order to circumvent this obstacle a new concept of strategies has been invented which we now discuss.

We maintain the basic framework introduced in section 6.1 but assume, for simplicity, that the control constraints (6.3) are of the simple form $u^i(s) \in U^i(s)$ where, for each $s \in [0, T\rangle$, $U^i(s)$ is a given subset of $\mathbb{R}^{m^i}$. Let δ be a positive real number such that T/δ is an integer if T is finite. The reader should note that from now on δ no longer denotes the delay between defection and punishment. Moreover, let $L = \infty$ if $T = \infty$ and $L = T/\delta + 1$ if T is finite. For each nonnegative integer l strictly smaller than L define $T_{l,\delta} = \delta l$. Let $\mathcal{V}^i_{l,\delta}$ be the set of admissible control functions $u^i : [T_{l,\delta}, T_{l+1,\delta}) \mapsto \mathbb{R}^{m^i}$ satisfying $u^i(s) \in U^i(s)$. Furthermore, define $\mathcal{V}_{l,\delta} = \mathcal{V}^1_{l,\delta} \times \mathcal{V}^2_{l,\delta} \times \ldots \times \mathcal{V}^N_{l,\delta}$, $\mathcal{U}_{0,\delta} = \{x_0\}$, and $\mathcal{U}_{l,\delta} = \mathcal{V}_{0,\delta} \times \mathcal{V}_{1,\delta} \times \ldots \times \mathcal{V}_{l,\delta}$. A δ-strategy for player i is a mapping $\phi^i_\delta : \{(u(\cdot), l) \mid u(\cdot) \in \mathcal{U}_{l-1,\delta}, 0 \le l < L\} \mapsto \mathcal{V}^i_{l,\delta}$.

So far, there is no essential difference between the present framework and that of section 6.1. As a matter of fact, δ-strategies defined above can also be interpreted as strategies in the sense of section 6.1, when player i's

[10]This argument is well known in repeated games, see also section 2.4.

information at time $t \in [T_{l,\delta}, T_{l+1,\delta})$ is given by $\mathcal{U}_{l,\delta}$. The crucial step by which we leave the setting of the earlier sections of this chapter is when we consider sequences of δ-strategies as δ approaches 0. More specifically, we define a strategy for player i as a set of δ-strategies $\{\phi^i_{\delta_k} \mid k = 1, 2, \ldots\}$ where $(\delta_k)_{k=1,2,\ldots}$ is a sequence of positive real numbers satisfying $\lim_{k\to\infty} \delta_k = 0$. If T is finite, we require in addition that T/δ_k is an integer for all k. We denote the set of all strategies for player i by Δ^i.

By saying that player i uses a strategy $\phi^i \in \Delta^i$ we mean that he can use any δ-strategy $\phi^i_{\delta_k} \in \phi^i$. This implies that he can use δ-strategies with arbitrary small but positive δ. Although the actual application of such a strategy is hard to imagine, Δ^i is a well defined mathematical object and differential games using these strategy spaces have been rigorously studied. We do not attempt to present the details of this rather cumbersome theory but refer the interested reader to the literature mentioned at the end of the chapter.

The basic steps of the theory can be explained as follows. First one can show that, under rather mild technical assumptions on the problem data, every N-tuple of δ-strategies $(\phi^1_{\delta^1}, \phi^2_{\delta^2}, \ldots, \phi^N_{\delta^N})$, where different players may use different values for δ, defines a unique N-tuple of control paths, a unique state trajectory, and unique values for the objective functionals. Let us call this the outcome of the game corresponding to the δ-strategy profile $(\phi^1_{\delta^1}, \phi^2_{\delta^2}, \ldots, \phi^N_{\delta^N})$. Then one shows that for every N-tuple of sequences of positive numbers δ^i_k, $i \in \{1, 2, \ldots, N\}$, with the property that $\lim_{k\to\infty} \delta^i_k = 0$, the set of accumulation points of the outcomes corresponding to $(\phi^1_{\delta^1_k}, \phi^2_{\delta^2_k}, \ldots, \phi^N_{\delta^N_k})$ is nonempty. Finally, one has to define the outcome set corresponding to an N-tuple of strategies $\phi = (\phi^1, \phi^2, \ldots, \phi^N)$. To this end, one selects for each player i a sequence δ^i_k and a corresponding sequence of δ^i_k-strategies in ϕ^i. For every choice of these sequences, the previous step yields a nonempty set of accumulation points. The set of outcomes corresponding to ϕ is simply the union of these sets of accumulation points for all possible choices of sequences. It is worth emphasizing that this construction yields, in general, an outcome set corresponding to ϕ which is not a singleton. This problem is taken care of in the definition of a Nash equilibrium in the following way. One defines a Nash equilibrium $\phi = (\phi^1, \phi^2, \ldots, \phi^N)$ by the usual Nash equilibrium property (ϕ^i is the best response to ϕ^{-i}) and, in addition, by the property that the outcome set corresponding to ϕ is a singleton.

The advantage of this approach is that it allows consideration of non-Markovian strategies in which a player can react to a certain event immediately after it happens. In particular, it allows us to consider trigger strategies without a fixed positive delay between defection and punishment. Using these trigger strategies it becomes possible to generalize theorem 6.3(ii) by showing that any acceptable target path can be supported as a trigger strategy equilibrium. Note that this result, together with theorem 6.2, implies that a strategy profile consisting of history-dependent strategies is a Nash equilibrium if and only if it is acceptable. Analogously, one can generalize theorem 6.4(ii) by showing that any target profile, which dominates (in the Pareto sense) a fixed Markov perfect Nash equilibrium in any subgame along the target path, can be supported as a subgame perfect trigger strategy equilibrium. A rigorous statement of these results is beyond the scope of this book.

6.5 Further reading

Trigger strategies have been mainly discussed in the framework of infinitely repeated games in discrete time (so-called supergames); see e.g. Friedman [99, 100]. In differential games, these equilibria were analysed by Benhabib and Radner [7], Hämäläinen et al. [113], Kaitala and Pohjola [141, 142], Mehlmann [177], and Tolwinski et al. [231].

Papers that study similar equilibria in continuous time game models are Ehtamo and Hämäläinen [69], Haurie et al. [123], and Kaitala [139].

The concept of δ-strategies was first introduced by Friedman [98]. Our presentation is based on Tolwinski et al. [231].

Example 6.1 is taken from Kaitala and Pohjola [142], who built on an earlier model by Hamada [112]. Differential game models of the conflict between workers and capitalists can be found in Hoel [127], Lancaster [156], and Sorger [221].

6.6 Exercises

1. Consider example 6.1 under the assumption that the production function is given by $f(x) = x^\alpha$ where $\alpha \in (0, 1)$. Compute the golden-rule capital–labour ratio $\bar{x}$. Assume $x_0 \in (0, \bar{x})$ and verify that the Pareto optimal state trajectory $\tilde{x}(\cdot)$ is given by

$$\tilde{x}(t) = \begin{cases} x_0 \left[\dfrac{x_0^{\alpha-1} + (\gamma - x_0^{\alpha-1}) e^{\gamma(\alpha-1)t}}{\gamma} \right]^{1/(1-\alpha)} & \text{if } t < \tau(x_0), \\ \bar{x} & \text{if } t \geq \tau(x_0). \end{cases}$$

Use this result to compute $\tau(x_0)$. Show that condition (6.15) is satisfied.

2. Prove theorem 6.4.

3. Consider the game $\Gamma_I(x_0, 0)$ defined in section 6.1 under the additional assumption that there are only two players ($N = 2$) and that T is finite. For each of the following information structures determine whether it is nonanticipating and/or regular.

   (i) $I^1(u(\cdot), t) = x(t)$ and $I^2(u(\cdot), t) = y(t)$ where
   $$y(t) = \begin{cases} 1 \text{ if } u^1(s) \leq 0 \text{ for all } s \in [0, t), \\ 0 \text{ otherwise.} \end{cases}$$

   (ii) $I^1(u(\cdot), t) = x(t)$ and $I^2(u(\cdot), t) = y(t)$ where
   $$y(t) = \begin{cases} 1 \text{ if } u^1(s) \leq 0 \text{ for almost all } s \in [0, t), \\ 0 \text{ otherwise.} \end{cases}$$

   (iii) $I^1(u(\cdot), t) = x(t)$ and $I^2(u(\cdot), t) = y(t)$, where $y(\cdot)$ is the solution of the boundary value problem
   $$\dot{y}(s) = g(x(s), u^1(s), u^2(s), y(s)), \quad y(T) = 0$$

   and g is a smooth function which is not constant with respect to any of its arguments.

4. Consider an infinite horizon differential game with two players. The utility functions are $F^1(x, u^1, u^2, t) = x - (u^1)^2/2$ and $F^2(x, u^1, u^2, t) = u^1 + u^2 - x^2/2$, and the system dynamics are $f(x, u^1, u^2, t) = u^2$. Assume that both players have the same discount rate $r > 0$ and that the control constraints are given by $u^1(t) \in [-1, 1]$ and $u^2(t) \in [-1, 0]$.

   (i) Find the Pareto optimal solution $(\tilde{\phi}^1, \tilde{\phi}^2)$ which maximizes the sum of the two objective functionals, $\tilde{J} = J^1(\phi^1; \phi^2, x_0, 0) + J^2(\phi^2; \phi^1, x_0, 0)$.

   (ii) Find a Markov perfect Nash equilibrium (θ^1, θ^2).

   (iii) Under the assumption that the initial state is $x_0 = 1 + r$, show that condition (6.17) in theorem 6.4 is satisfied for some $\epsilon \geq 0$ if and only if $e^{-r} < 1 - r/2$.

   (iv) Under the assumption that the initial state satisfies $x_0 \leq r$ show that condition (6.17) in theorem 6.4 cannot be satisfied for any $\epsilon \geq 0$.

7 Differential games with special structures

In this chapter we are concerned with the analysis of tractable game structures. More specifically, we investigate various classes of differential games for which one can derive analytical characterizations of open-loop and Markov perfect Nash equilibria. We stress analytical tractability since we believe that analytical solutions have the advantage of shedding light on the qualitative properties of equilibria in a general way.

We identify three different classes of games for which we discuss the derivation of both open-loop and Markov perfect Nash equilibria. The first class is that of linear quadratic games. Such a game is characterized by a linear system of state equations and quadratic objective functions. The class of linear quadratic differential games has not only gained popularity among dynamic game theorists but also among macroeconomists interested in issues such as policy coordination, optimal stabilization policies, and the like. For this reason we present some macroeconomic interpretations of linear quadratic games.

The second class of solvable differential games we consider consists of games in which the state variables enter both the state equations and the objective functions linearly. We refer to this class as linear state games. We show that these games have the property that an open-loop Nash equilibrium is Markov perfect and that the optimal value functions are linear with respect to the state variables.

The third class of games discussed in this chapter is exponential games. In this class the state variables enter the objective functions via an exponential term while the state equation is independent of the state variables. We discuss a general solution to this class of games when the planning horizon is finite and point out that, after appropriate transformations, exponential games become linear state games.

We do not discuss the three classes of games in their most general form but, instead, illustrate their properties by means of simple examples. We believe that this form of presentation highlights the most important points concerning the different solution techniques without distracting the reader by (often cumbersome) analytical details. Our goal then is to emphasize the general principles of the analysis of these games and show the reader how to study similar and perhaps more general games which share the same basic structure.[1]

7.1 Linear quadratic games

We consider a two-person differential game in which player 1 minimizes the quadratic cost function

$$\frac{1}{2}\int_0^T e^{-rt}\left\{g_1 x(t)^2 + g_2[u^1(t)]^2\right\} dt$$

and player 2 minimizes

$$\frac{1}{2}\int_0^T e^{-rt}\left\{m_1 x(t)^2 + m_2[u^2(t)]^2\right\} dt.$$

The state equation is given by

$$\dot{x}(t) = ax(t) + bu^1(t) + cu^2(t), \qquad x(0) = x_0. \tag{7.1}$$

Here g_i, m_i, a, b, c, and r are constants, $x(t)$ is the single state variable, $u^i(t)$ is the scalar control variable of player i, and T is either a finite or infinite time horizon. Because the characteristic features of this game are the linear state equation and the quadratic objective functions, we refer to it as a linear quadratic game.

In the present section we discuss several approaches to analysing this game. The general principles of these approaches can also be used to study more general linear quadratic games. For example, one could include quadratic salvage value functions in the finite horizon game, consider games with more than two players, more than one state variable, or more than one control variable for each player, and allow the cost functions to be more general quadratic forms in the variables $x(t)$, $u^1(t)$, and $u^2(t)$. It is also possible to apply the same techniques to nonstationary games, that is, games in which a, b, c, g_i, and m_i are time functions instead of constants. The results are also easily generalized to the situation where the players use different discount factors. Presenting the solution techniques for more general linear quadratic games would not lead to essentially new insights but would make the analytical derivations more cumbersome.

[1] In an appendix to this chapter we present the general case of linear quadratic games and derive open-loop and Markov perfect Nash equilibria.

We have introduced the linear quadratic differential game in the formulation which is most often encountered in economic applications. This means that the objective of the players is to minimize the present value of future costs. Since in this book we deal with maximization problems, we rewrite the game by multiplying both objective functions by -1 and turning the minimization problems into maximization problems. Thus, the objective functionals of player 1 and 2 are

$$J_1 = -\frac{1}{2}\int_0^T e^{-rt}\{g_1 x(t)^2 + g_2[u^1(t)]^2\}\,dt, \tag{7.2}$$

$$J_2 = -\frac{1}{2}\int_0^T e^{-rt}\{m_1 x(t)^2 + m_2[u^2(t)]^2\}\,dt. \tag{7.3}$$

The players seek to maximize these functionals subject to (7.1).

Several (macro)economic interpretations can be given for linear quadratic differential games. They differ especially with respect to the identities of the players. The common feature of all interpretations is that the state variables (and perhaps also the control variables) describe deviations of certain economic indicators (or instrument variables) from their target values or their natural long-run levels. These deviations are penalized by quadratic cost functions, that is, the goal of the decision makers is to minimize the discounted quadratic deviations from targets. If the underlying economic system relating the state and control variables is a linear system (as is often assumed to be the case) then the resulting game is of the linear quadratic form. In the literature the following situations, for instance, have been modelled by linear quadratic differential games:

- *Stabilization policies in a closed economy with multiple policy authorities* Player 1 is the government, player 2 is the central bank. Both have quadratic cost functions with different weights attached to economic target variables such as the rates of unemployment or inflation. The instruments of the government may be public expenditures, the budget deficit, or another measure of fiscal policy, whereas the money supply is under the central bank's control. Some applications consider the growth rates of these variables as controls instead of the levels. Target and instrument variables are linked by a linear macroeconomic model containing, for example, output and money market equilibrium conditions and an expectations-augmented Phillips curve. A dynamic relation of the type of equation (7.1) may arise by assuming that expectations about the inflation are formed in an adaptive way.
- *International policy coordination* The two players represent the governments of two countries. They attach possibly different weights to deviations of a common state variable (such as the balance of pay-

ments or the exchange rate) from a certain target level. This state variable, in turn, is influenced by both governments' instrument variables which may include those listed in the previous paragraph plus, for example, import tariffs.

The interpretations of the linear quadratic game given here are only a few within the field of macroeconomics. In addition there exist many applications in the fields of microeconomics, finance, and management science. In the second part of the book we will discuss some of these applications in more detail.

7.1.1 Open-loop Nash equilibrium

In deriving equilibria for linear quadratic games we first consider the open-loop version of the above game. That is, we assume that both players commit themselves to certain courses of action for the entire planning period, from which they cannot deviate even if that would be in their interest. The only information on which the equilibrium strategies are based is calendar time together with the parameters of the model, which are common knowledge. For the purpose of interpretation, we can imagine $x(t)$ to be the deviation of a target variable from its ideal level (which is assumed to be politically desired and sustainable in the long run). The control variables $u^1(t)$ and $u^2(t)$ denote actions of player 1 and 2, respectively, which exert influence on $x(t)$ via (7.1), but which are costly to use for the decision makers. In most economic applications it makes sense to assume that g_1, m_1, g_2, and m_2 are strictly positive and that r is nonnegative. The parameters b and c may be positive or negative, whereas a is in most cases negative so that the uncontrolled system is stable (although this is not required for the most part of the following analysis).

Let $(u^1(\cdot), u^2(\cdot))$ be an open-loop Nash equilibrium. The current-value Hamiltonians of players 1 and 2 are

$$H_1(x, u^1, p^1, t) = -\frac{1}{2}[g_1 x^2 + g_2(u^1)^2] + p^1[ax + bu^1 + cu^2(t)],$$
$$H_2(x, u^2, p^2, t) = -\frac{1}{2}[m_1 x^2 + m_2(u^2)^2] + p^2[ax + bu^1(t) + cu^2],$$

where p^i, $i = 1, 2$, denotes the costate variable of player i.

If there are neither control nor state constraints, maximizing the Hamiltonians with respect to the control variables results in

$$u^1(t) = \left(\frac{b}{g_2}\right)p^1(t) \tag{7.4}$$

and

$$u^2(t) = \left(\frac{c}{m_2}\right)p^2(t) \tag{7.5}$$

for all t. The adjoint equations are

$$\dot{p}^1(t) = g_1 x(t) + (r - a)p^1(t),$$
$$\dot{p}^2(t) = m_1 x(t) + (r - a)p^2(t).$$

Substituting (7.4) and (7.5) into the state equation (7.1) and combining it with the adjoint equations results in the so-called canonical system

$$\begin{aligned} \dot{x}(t) &= ax(t) + (b^2/g_2)p^1(t) + (c^2/m_2)p^2(t), \\ \dot{p}^1(t) &= g_1 x(t) + (r - a)p^1(t), \\ \dot{p}^2(t) &= m_1 x(t) + (r - a)p^2(t). \end{aligned} \tag{7.6}$$

If the time horizon is finite, system (7.6) has the initial condition $x(0) = x_0$ and two transversality conditions $p^i(T) = 0$ for $i = 1, 2$. Under the stated parameter restrictions, the maximized Hamiltonian functions are concave with respect to the state variable such that (7.4)–(7.6) together with the boundary conditions are sufficient for the Nash equilibrium property of $(u^1(\cdot), u^2(\cdot))$. In the case of an infinite time horizon one has to replace the terminal conditions $p^i(T) = 0$ by appropriate limiting transversality conditions (see theorem 4.2 for details). We will demonstrate below how to construct a solution to (7.4)–(7.6) such that the limiting transversality conditions hold.

In vector notation the canonical system (7.6) can be written as

$$\dot{y}(t) = A \cdot y(t), \tag{7.7}$$

where $y(t) = \big(x(t), p^1(t), p^2(t)\big)'$ (here $'$ indicates transposition), and

$$A = \begin{pmatrix} a & b^2/g_2 & c^2/m_2 \\ g_1 & r-a & 0 \\ m_1 & 0 & r-a \end{pmatrix}.$$

Since an open-loop Nash equilibrium is fully characterized by (7.4), (7.5), and system (7.7), our next objective is to derive a solution to (7.7). This is a system of linear first order differential equations that can be solved analytically. We have to distinguish between the case of a finite horizon and the case of an infinite time horizon.

Let us first consider the finite horizon problem in which $T < \infty$. We have already mentioned that in this case we have the initial condition $x(0) = x_0$ at $t = 0$ and the terminal conditions $p^1(T) = p^2(T) = 0$ at $t = T$

so that system (7.7) is a two-point boundary value problem. In order to find a solution to that problem we proceed as follows. First we derive the three eigenvalues of the system matrix A as well as the corresponding eigenvectors. Using these results we derive the fundamental solution of (7.7), and finally we calculate a particular solution that satisfies the boundary conditions. Let us carry out these steps in turn.

The determinant of A is

$$\det A = (r - a)[a(r - a) - M],$$

where

$$M = c^2(m_1/m_2) + b^2(g_1/g_2).$$

The eigenvalues s_1, s_2, and s_3 of A can be derived from the characteristic equation $\det(sI - A) = 0$, where I is the identity matrix.[2] More specifically, this equation is

$$\det(sI - A) = (s - r + a)^2(s - a) - (s - r + a)M = 0.$$

The three solutions of this equation are the eigenvalues

$$s_1 = r/2 - \sqrt{r^2/4 - a(r - a) + M},$$
$$s_2 = r/2 + \sqrt{r^2/4 - a(r - a) + M},$$
$$s_3 = r - a.$$

We can distinguish five different cases according to the signs of the eigenvalues:

(1) One eigenvalue is negative, two are positive: $s_1 < 0$, $s_2 > 0$, $s_3 > 0$. This case occurs whenever the determinant of the system matrix A is negative.
(2) One eigenvalue is positive, two are zero: $s_1 = s_2 = 0$, $s_3 > 0$. This case occurs if $0 < a < r$ and $a(r - a) = M$ holds so that $\det A = 0$.
(3) All three eigenvalues are positive: $s_1 > 0$, $s_2 > 0$, $s_3 > 0$. This case occurs if $0 < a < r$ and $a(r - a) > M$ holds so that $\det A > 0$.
(4) One eigenvalue is positive, one is negative, and one is zero: $s_1 < 0$, $s_2 > 0$, $s_3 = 0$. This case occurs if $a = r$ so that $\det A = 0$.
(5) One eigenvalue is positive, two are negative: $s_1 < 0$, $s_2 > 0$, $s_3 < 0$. This case occurs if $r < a$ so that $\det A > 0$.

Because of $r^2/4 - a(r - a) = (r/2 - a)^2 > 0$ and $M > 0$ by our assumptions, no complex solutions can occur in any case.

[2] If one applies this solution technique to a more general differential game with several state variables, one can solve the characteristic equation only by numerical methods.

In the following we restrict ourselves to the assumption of a stable uncontrolled system ($a \leq 0$), i.e. to case 1.[3] Therefore two eigenvalues of A are positive and one is negative, implying that the canonical system has the saddle point property. The general solution of the canonical system (7.7) can be written as

$$y(t) = e^{At}y(0) = We^{\Lambda t}\alpha, \tag{7.8}$$

where, for $i = 1, 2, 3$, we denote an eigenvector of A corresponding to s_i by $w_i = (w_{1i}, w_{2i}, w_{3i})' \in \mathbb{R}^3$. The matrices W and Λ are given by

$$W = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{pmatrix}.$$

Furthermore, we have

$$e^{\Lambda t} = \begin{pmatrix} e^{s_1 t} & 0 & 0 \\ 0 & e^{s_2 t} & 0 \\ 0 & 0 & e^{s_3 t} \end{pmatrix}$$

and $\alpha = W^{-1}y(0)$. Note that under our assumptions the three eigenvalues are different and, thus, the eigenvectors are linearly independent.

Using (7.8) we can express the three boundary conditions in the case of a finite horizon game as

$$\begin{aligned} x_0 &= \alpha_1 w_{11} + \alpha_2 w_{12} + \alpha_3 w_{13}, \\ 0 &= \alpha_1 w_{21} e^{s_1 T} + \alpha_2 w_{22} e^{s_2 T} + \alpha_3 w_{23} e^{s_3 T}, \\ 0 &= \alpha_1 w_{31} e^{s_1 T} + \alpha_2 w_{32} e^{s_2 T} + \alpha_3 w_{33} e^{s_3 T}. \end{aligned}$$

This system of linear equations can easily be solved for the unknown parameters α_1, α_2, and α_3. Using this solution, (7.4), and (7.5) the open-loop Nash equilibrium strategies are completely determined.

In the case of an infinite horizon we have to proceed in a slightly different way. We have a given initial state x_0 but no terminal values for the costate variables $p^i(t)$. In order to satisfy the limiting transversality conditions (cf. theorem 3.4), we choose a solution to (7.7) which converges to a steady state. In the case under consideration, this can be done by setting $\alpha_2 = \alpha_3 = 0$, because s_1 and w_1 are the stable eigenvalue and eigenvector, respectively. The unique particular solution satisfying $\alpha_2 = \alpha_3 = 0$ and $x(0) = x_0$ is

[3] A similar analysis also holds for cases 4 and 5, whereas in cases 2 and 3 the canonical system is totally unstable.

$$x(t) = x_0 e^{s_1 t},$$
$$p^1(t) = x_0(w_{21}/w_{11})e^{s_1 t},$$
$$p^2(t) = x_0(w_{31}/w_{11})e^{s_1 t}.$$

As in the finite horizon case we can use (7.4) and (7.5) to find the open-loop Nash equilibrium strategies. In the present case they are given by

$$u^1(t) = x_0\left(\frac{b}{g_2}\right)\left(\frac{w_{21}}{w_{11}}\right)e^{s_1 t},$$
$$u^2(t) = x_0\left(\frac{c}{m_2}\right)\left(\frac{w_{31}}{w_{11}}\right)e^{s_1 t}.$$

A nondegenerate Markovian representation of these control paths is $u^1(t) = \phi^1(x(t))$, $u^2(t) = \phi^2(x(t))$, where

$$\phi^1(x) = \left(\frac{b}{g_2}\right)\left(\frac{w_{21}}{w_{11}}\right)x,$$
$$\phi^2(x) = \left(\frac{c}{m_2}\right)\left(\frac{w_{31}}{w_{11}}\right)x.$$

Given the explicit form of the state and control variables it is also easy to calculate the equilibrium cost of each player over the entire planning horizon. These costs are given by

$$-J_1 = \frac{x_0^2}{2(r-2s_1)}\left[g_1 + \left(\frac{b^2}{g_2}\right)\left(\frac{w_{21}}{w_{11}}\right)^2\right],$$
$$-J_2 = \frac{x_0^2}{2(r-2s_1)}\left[m_1 + \left(\frac{c^2}{m_2}\right)\left(\frac{w_{31}}{w_{11}}\right)^2\right].$$

Since the open-loop equilibrium strategies determined above depend on the initial state x_0 of the game, it is obvious that the open-loop Nash equilibrium is not Markov perfect. This is the case for general linear quadratic games, too. It is therefore of interest to derive Markov perfect Nash equilibria by considering equilibrium strategies that depend explicitly on the state of the game. This is done in section 7.1.3 below. Before that, however, we demonstrate another approach to determining open-loop Nash equilibria of a linear quadratic differential game.

7.1.2 Open-loop Nash equilibrium: an alternative approach

In the previous subsection we characterized the unique open-loop Nash equilibrium for a linear quadratic differential game by explicitly solving the linear canonical system in the state and costate variables. In this section we make use of an alternative approach that rests on the derivation of a system of Riccati equations. In order to demonstrate this approach we restrict ourselves to a finite horizon game but, in contrast to the previous subsection, we allow for nonstationarity and include quadratic salvage value functions. Thus, the linear quadratic differential game has the objective functionals

$$J_1 = -\frac{1}{2}\int_0^T e^{-rt}\{g_1(t)x(t)^2 + g_2(t)[u^1(t)]^2\}\,dt - \frac{1}{2}e^{-rt}q^1x(T)^2, \qquad (7.9)$$

$$J_2 = -\frac{1}{2}\int_0^T e^{-rt}\{m_1(t)x(t)^2 + m_2(t)[u^2(t)]^2\}\,dt - \frac{1}{2}e^{-rt}q^2x(T)^2. \qquad (7.10)$$

Each player seeks to maximize his payoff functional subject to the state equation

$$\dot{x}(t) = a(t)x(t) + b(t)u^1(t) + c(t)u^2(t), \qquad x(0) = x_0. \qquad (7.11)$$

The functions $a(\cdot)$, $b(\cdot)$, $c(\cdot)$, $g_1(\cdot)$, $g_2(\cdot)$, $m_1(\cdot)$, and $m_2(\cdot)$ are continuous and exogenously given and r, T, x_0, q^1, and q^2 are given constants. Because of the explicit time dependence of the system dynamics and the cost functions, the game is nonautonomous.

In exactly the same way as in the previous subsection we obtain the optimality conditions

$$u^1(t) = \frac{b(t)}{g_2(t)}p^1(t), \qquad (7.12)$$

$$u^2(t) = \frac{c(t)}{m_2(t)}p^2(t), \qquad (7.13)$$

and the canonical system

$$\begin{aligned}
\dot{x}(t) &= a(t)x(t) + [b(t)^2/g_2(t)]p^1(t) + [c(t)^2/m_2(t)]p^2(t),\\
\dot{p}^1(t) &= [r - a(t)]p^1(t) + g_1(t)x(t), \qquad (7.14)\\
\dot{p}^2(t) &= [r - a(t)]p^2(t) + m_1(t)x(t),
\end{aligned}$$

where, as before, $p^i(t)$, $i = 1, 2$, denotes the costate variable of player i. Due to the presence of salvage value functions the transversality conditions are now

$$p^i(T) = -q^i x(T), \qquad i = 1, 2.$$

So far, everything has been completely analogous to the approach in section 7.1.1. Now we depart from that approach by assuming that the costate variables can be written in the form

$$p^i(t) = N^i(t)x(t) + n^i(t), \tag{7.15}$$

where $N^i(\cdot)$ and $n^i(\cdot)$ are differentiable functions which will be determined below.[4] To satisfy the transversality conditions we require

$$N^i(T) = -q^i \qquad \text{and} \qquad n^i(T) = 0. \tag{7.16}$$

Differentiation of (7.15) with respect to time yields

$$\dot{p}^i(t) = \dot{N}^i(t)x(t) + N^i(t)\dot{x}(t) + \dot{n}^i(t).$$

Now we replace the time derivatives $\dot{p}^i(t)$ and $\dot{x}(t)$ in this equation by the expressions given in the state and costate equations (7.14) and replace the costate variables by the right-hand side of (7.15). In the resulting equations we collect terms involving the same power of $x(t)$. The equations are affine with respect to $x(t)$ so that every equation contains a constant term and a term involving $x(t)$. One way to satisfy the equations is to require that the constant term and the coefficient of $x(t)$ in the linear term is identically equal to 0. This gives the following system of Riccati differential equations

$$\dot{N}^1(t) = g_1(t) + [r - 2a(t)]N^1(t) - \frac{b(t)^2}{g_2(t)}[N^1(t)]^2 - \frac{c(t)^2}{m_2(t)}N^1(t)N^2(t), \tag{7.17}$$

$$\dot{N}^2(t) = g_2(t) + [r - 2a(t)]N^2(t) - \frac{c(t)^2}{m_2(t)}[N^2(t)]^2 - \frac{b(t)^2}{g_2(t)}N^1(t)N^2(t), \tag{7.18}$$

and

$$\dot{n}^1(t) = [r - a(t)]n^1(t) - \frac{b(t)^2}{g_2(t)}N^1(t)n^1(t) - \frac{c(t)^2}{m_2(t)}N^1(t)n^2(t), \tag{7.19}$$

$$\dot{n}^2(t) = [r - a(t)]n^2(t) - \frac{c(t)^2}{m_2(t)}N^2(t)n^2(t) - \frac{b(t)^2}{g_2(t)}N^2(t)n^1(t). \tag{7.20}$$

System (7.17)–(7.18) is a system of two coupled Riccati differential equations for the functions $N^1(\cdot)$ and $N^2(\cdot)$. The boundary conditions are stated in (7.16). Solutions to such systems can be found using standard numerical packages. Once the solutions are found and substituted in (7.19)–(7.20) the latter system is a homogeneous system of linear differential equations. Since the boundary conditions stated in (7.16) are $n^1(T) = n^2(T) = 0$, it follows from the theory of linear differential equations that $n^1(t) = n^2(t) = 0$ for all $t \in [0, T]$. Since we have determined the

[4]This assumption involves no loss of generality because it is trivially satisfied if $N^i(t) = 0$ and $n^i(t) = p^i(t)$ for all t. In the following, however, we shall determine $N^i(\cdot)$ and $n^i(\cdot)$ in a different way so that (7.15) is still satisfied.

functions $N^i(\cdot)$ and $n^i(\cdot)$, equations (7.12), (7.13), and (7.15) yield a nondegenerate Markovian representation of the open-loop Nash equilibrium control paths. The corresponding state trajectory can be found by substituting these strategies into the state equation, which yields

$$\dot{x}(t) = \left[a(t) + \frac{b(t)^2}{g_2(t)}N^1(t) + \frac{c(t)^2}{m_2(t)}N^2(t)\right]x(t), \qquad x(0) = x_0.$$

7.1.3 Markov perfect Nash equilibrium

We now turn to the derivation of a Markov perfect Nash equilibrium of the game considered in section 7.1. Denoting the equilibrium strategies by $\phi^1(x, t)$ and $\phi^2(x, t)$ the HJB equations for this game are

$$rV^1(x,t) - \frac{\partial V^1(x,t)}{\partial t}$$
$$= \max\left\{-\frac{1}{2}[g_1x^2 + g_2(u^1)^2] + \frac{\partial V^1(x,t)}{\partial x}[ax + bu^1 + c\phi^2(x,t)] \,\Big|\, u^1 \in \mathbb{R}\right\},$$
$$rV^2(x,t) - \frac{\partial V^2(x,t)}{\partial t}$$
$$= \max\left\{-\frac{1}{2}[m_1x^2 + m_2(u^2)^2] + \frac{\partial V^2(x,t)}{\partial x}[ax + b\phi^1(x,t) + cu^2] \,\Big|\, u^2 \in \mathbb{R}\right\}.$$

The boundary conditions for the finite horizon game are $V(x, T) = 0$ for all x. To find a solution to these equations we proceed as follows. Given the linear quadratic structure of the game, we guess that the optimal value functions are quadratic and that the equilibrium strategies are linear with respect to the state variable. We then demonstrate that this conjecture does indeed lead to optimal value functions that satisfy the HJB equations and the boundary conditions. More specifically, we postulate a quadratic value function of the form[5]

$$V^i(x,t) = (1/2)v^i(t)x^2,$$

where $v^i(\cdot)$, $i = 1, 2$, is a function that has to be determined. The transversality condition for the finite horizon problem is satisfied if $v^i(T) = 0$.

[5]In this simple example of a linear quadratic differential game the initial guess for the strategies is irrelevant. In more complicated linear quadratic games, one has to use a parametric guess of the form $\phi(x, t) = A(t)x + B(t)$. In that case one also needs a more general guess for the optimal value fraction, such as $V^i(x, t) = (1/2)v^i(t)x^2 + w^i(t)x + z^i(t)$.

Substituting the quadratic value functions into the HJB equations and carrying out the maximization on the right-hand side yields

$$u^1 = \phi^1(x,t) = \left(\frac{b}{g_2}\right)v^1(t)x, \tag{7.21}$$

$$u^2 = \phi^2(x,t) = \left(\frac{c}{m_2}\right)v^2(t)x. \tag{7.22}$$

Substituting this back into the HJB equations, collecting terms with equal powers of x, and equating the coefficients of these terms to zero, one obtains the following system of coupled Riccati differential equations:

$$\dot{v}^1(t) = g_1 + (r-2a)v^1(t) - \frac{b^2}{g_2}[v^1(t)]^2 - 2\frac{c^2}{m_2}v^1(t)v^2(t), \tag{7.23}$$

$$\dot{v}^2(t) = m_1 + (r-2a)v^2(t) - \frac{c^2}{m_2}[v^2(t)]^2 - 2\frac{b^2}{g_2}v^1(t)v^2(t). \tag{7.24}$$

The reader should note the similarity of this system to the Riccati differential equations (7.17)–(7.18) which characterize the unique open-loop Nash equilibrium of the game. The two systems differ only in the coefficients of the cross products $N^1(t)N^2(t)$ (respectively $v^1(t)v^2(t)$). This difference reflects the different informational assumptions of the two equilibria. In the Markov perfect Nash equilibrium discussed in the present section the players take into account that their opponents react to changes of the state variable. In the open-loop equilibrium discussed in the previous subsections the strategies do not depend on the state which, in turn, leads to the different coefficients.

As has already been mentioned in the previous subsection, a system such as (7.23)–(7.24) can readily be solved by numerical methods. The boundary conditions are given by $v^1(T) = v^2(T) = 0$. Once the functions $v^1(\cdot)$ and $v^2(\cdot)$ are determined, equations (7.21) and (7.22) determine the equilibrium strategies. Since the conditions of theorem 4.4 are satisfied, the equilibrium is Markov perfect.

In case of an infinite horizon we proceed basically in the same way as before, but we use stationary (i.e. time independent) value functions $V^i(x)$ as well as stationary strategies $\phi^i(x)$. For the problem under consideration, the guess for the optimal value functions is therefore

$$V^i(x) = (1/2)v^i x^2, \qquad i = 1, 2,$$

where v^i, $i = 1, 2$, are constant parameters of the unknown functions $V^i(x)$ which are to be determined. Since $v^i(t) = v^i$ does not depend on t we have $\dot{v}^i(t) = 0$ and equations (7.23) and (7.24) become a system of algebraic Riccati equations, namely

$$g_1 + (r - 2a)v^1 - (b^2/g_2)(v^1)^2 - 2(c^2/m_2)v^1v^2 = 0, \tag{7.25}$$

$$m_1 + (r - 2a)v^2 - (c^2/m_2)(v^2)^2 - 2(b^2/g_2)v^1v^2 = 0. \tag{7.26}$$

This system has in general multiple solutions. Under the parameter restrictions stated in section 7.1.1 (including the condition that $a \leq 0$ so that the uncontrolled system is stable) it can be shown that there exists a unique negative solution, i.e., $v^1 < 0$ and $v^2 < 0$.[6] This solution is the one that produces value functions satisfying the equilibrium conditions of theorem 4.4. To see this, first note that the parameter restrictions and $v^i < 0$ imply that

$$\begin{aligned}\dot{x}(t) &= ax(t) + b\phi^1(x(t)) + c\phi^2(x(t)) \\ &= ax(t) + \left(\frac{b^2}{g_2}\right)v^1x(t) + \left(\frac{c^2}{m_2}\right)v^2x(t)\end{aligned} \tag{7.27}$$

has the opposite sign of $x(t)$ (recall that we have assumed $a \leq 0$).

Now assume for a moment that the state space is the bounded interval $X = [-\alpha, \beta]$, where α and β are positive constants so that the steady state is contained in this interval. Because of the properties mentioned in the previous paragraph $u^1 = \phi^1(x)$ and $u^2 = \phi(x)$ are feasible even if x is a boundary point of the state space. Thus, the derivation of the above formulas (in particular, equations (7.25) and (7.26)) remains valid even under the assumption of a bounded state space. Since the quadratic value functions $V^i(x)$ are bounded on any bounded interval, the conditions of theorem 4.4 are satisfied and the proposed linear strategies $\phi^1(\cdot)$ and $\phi^2(\cdot)$ constitute a Markov perfect Nash equilibrium. Note that this argument is not necessarily valid if we had chosen a solution of (7.25)–(7.26) for which $v^1 > 0$ or $v^2 > 0$.

The above argument holds for any bounded state space $[\alpha, \beta]$ which contains 0. One can choose α and β arbitrarily large (but finite). In the case where the state space is unbounded, $X = \mathbb{R}$, the argument is technically more complicated and requires an approach based on the conditions of theorem 3.4 (see also example 3.5).

The equilibrium dynamics for the Markov perfect Nash equilibrium is given by (7.27). Hence, the equilibrium state trajectory is $x(t) = x_0e^{st}$ with

$$s = a + (b^2/g_2)v^1 + (c^2/m_2)v^2 < 0.$$

[6]An analogous result holds also for the case of more general linear quadratic games with more than one state variable whereby negativity has to be replaced by negative definiteness. In that case the solutions to (7.25)–(7.26) can only be obtained numerically.

From this result it follows immediately that the open-loop representations of the control paths generated by the Markov perfect Nash equilibrium strategies are

$$u^1(t) = \phi^1(x(t)) = x_0(b/g_2)v^1 e^{st} \quad \text{and} \quad u^2(t) = \phi^2(x(t)) = x_0(c/m_2)v^2 e^{st}.$$

The resulting costs for the two players over the infinite planning horizon are

$$-J_1 = \frac{x_0^2}{2(r-2s)}\left[g_1 + \left(\frac{b^2}{g_2}\right)(v^1)^2\right],$$
$$-J_2 = \frac{x_0^2}{2(r-2s)}\left[m_1 + \left(\frac{c^2}{m_2}\right)(v^2)^2\right].$$

7.1.4 A numerical example

In order to get an idea about the quantitative differences of open-loop and Markov perfect equilibria of linear quadratic differential games let us present a numerical example. For that matter we formulate a macroeconomic model in which the governments of two countries aim at stabilizing domestic policies.

Consider two symmetric countries with private sector agents that have perfect foresight, but do not act strategically with respect to government stabilization policies. Short-run equilibria in the goods markets are given by

$$q(t) = \rho q^*(t) - \gamma r(t) + \delta[e(t) + p^*(t) - p(t)]$$
$$q^*(t) = \rho q(t) - \gamma r^*(t) - \delta[e(t) + p^*(t) - p(t)]$$

where $q(t)$ denotes the deviation of the real output of country 1 from its natural level, $r(t)$ is the real interest rate, $e(t)$ is the exchange rate measured in terms of units of the currency of country 1 per unit of foreign currency and $p(t)$ is the price level.[7] Variables with an asterisk refer to country 2. The equilibrium equations state that output of each country must be equal to domestic demand, which depends on income (output) and the interest rate, plus foreign demand, which depends on foreign income and the real exchange rate. Under perfect foresight real interest rates are determined by

[7] All variables in this model (except interest rates) are measured on a logarithmic scale. This allows us to formulate the model equations in linear form.

$$r(t) = i(t) - \dot{p}(t),$$
$$r^*(t) = i^*(t) - \dot{p}^*(t),$$

where $i(t)$ denotes the nominal rate of interest. Asset market equilibrium requires that the real money supply equals real money demand, where real money demand depends on the current level of income and the nominal interest rate, i.e.

$$\tilde{m}(t) - p(t) = \kappa q(t) - \lambda i(t),$$
$$\tilde{m}^*(t) - p^*(t) = \kappa q^*(t) - \lambda i^*(t),$$

with $\tilde{m}(t)$ denoting nominal money supply. The equilibrium in the money market assumes that residents of each country only hold portfolios in domestic currency. Under the assumption of perfect capital markets and perfect foresight, the interest rate parity condition

$$i(t) = i^*(t) + \dot{e}(t)$$

holds in each country and the price levels adjust to domestic excess demand according to

$$\dot{p}(t) = \xi q(t), \tag{7.28}$$
$$\dot{p}^*(t) = \xi q^*(t). \tag{7.29}$$

The variables γ, δ, κ, λ, and ξ are assumed to be positive constants and $0 < \rho < 1$.

We assume that each government uses the domestic real money supply $m(t) = \tilde{m}(t) - p(t)$ (respectively, $m^*(t) = \tilde{m}^*(t) - p^*(t)$) as its control variable. In this simple model we can identify the real exchange rate $s(t) = e(t) + p^*(t) - p(t)$ as a measure of international competitiveness. With these definitions it is now possible, after appropriate substitutions, to derive a reduced form model of this two-country economy that is fully described by a dynamic equation for the real exchange rate,

$$\dot{s}(t) = \phi_1 m^*(t) - \phi_1 m(t) + \phi_2 s(t), \tag{7.30}$$

and two output equations

$$q(t) = am(t) + \frac{\rho}{k} am^*(t) + bs(t),$$
$$q^*(t) = \frac{\rho}{k} am(t) + am^*(t) - bs(t).$$

Here we have

$$k = 1 + \frac{\gamma\kappa}{\lambda} - \gamma\xi, \; a = \frac{\gamma k}{\lambda(k^2 - \rho^2)}, \; b = \frac{\delta}{k + \rho},$$
$$\phi_1 = \frac{1 + \rho}{\lambda(k + \rho)}, \; \phi_2 = \frac{2\delta(k - 1)}{\gamma(k - \rho)}.$$

The policy makers in each country choose their optimal (monetary) policy so as to minimize the costs of the output gap and inflation and the loss of international competitiveness due to a devaluation of the currency. If the cost functions are additive and quadratic and policy makers plan for an infinite horizon we get the objective functional

$$\frac{1}{2}\int_0^\infty e^{-\tilde{r}t}\left[\alpha q(t)^2 + \beta\dot{p}(t)^2 + \eta s(t)^2\right] dt$$

for country 1 and

$$\frac{1}{2}\int_0^\infty e^{-\tilde{r}t}\left[\alpha q^*(t)^2 + \beta\dot{p}^*(t)^2 + \eta s(t)^2\right] dt$$

for country 2, where α, β, and η are positive constants. The constant $\tilde{r}$ denotes the common time preference rate of the two governments. Substitution of the price adjustment equations (7.28) and (7.29) shows that the objective functionals can be rewritten in the form

$$\frac{1}{2}\int_0^\infty e^{-\tilde{r}t}\left[q(t)^2 + cs(t)^2\right] dt \tag{7.31}$$

and

$$\frac{1}{2}\int_0^\infty e^{-\tilde{r}t}\left[(q^*)(t)^2 + cs(t)^2\right] dt, \tag{7.32}$$

respectively, where $c = \eta/(\alpha + \beta\xi^2)$. From these transformations it is clear that the game (7.30)–(7.32) is of the linear quadratic type and fits exactly to the structure (7.1)–(7.3).

To get some flavour of the quantitative magnitudes involved in the comparison of different equilibria for this linear quadratic game, we illustrate our analytical derivations by numerical calculations. For that purpose we relax the assumption that the two countries are symmetric and introduce a reduced form state equation of the type

$$\dot{s}(t) = \phi_1^1 m^*(t) - \phi_1^2 m(t) + \phi_2 s(t).$$

This asymmetry can arise if the two countries face different speeds of adjustments of the price levels to domestic excess demand. We specify the following values for the parameters of the model: $\phi_2 = -0.375$, $\phi_1^1 = 1$, $\phi_1^2 = 0.6$, $c = 1$, $\tilde{r} = 0$, $s_0 = 25$. The following tables summarize the resulting values of the state and the control variables for the first five periods (the first one being subdivided into ten subperiods), the stable eigenvalues of the controlled system, and the total costs (over the infinite planning

Table 7.1. *Open-loop Nash equilibrium*

$s(t)$	$m(t)$	$m^*(t)$	t
25.000	15.627	−9.368	0.0
22.118	13.826	−8.288	0.1
19.569	12.232	−7.333	0.2
17.313	10.822	−6.488	0.3
15.317	9.575	−5.740	0.4
13.552	8.471	−5.078	0.5
11.989	7.495	−4.493	0.6
10.607	6.631	−3.975	0.7
9.385	5.866	−3.517	0.8
8.303	5.190	−3.111	0.9
7.344	4.589	−2.753	1.0
2.157	1.348	−0.809	2.0
0.634	0.396	−0.238	3.0
0.186	0.116	−0.070	4.0
0.055	0.034	−0.021	5.0

Notes: Stable eigenvalue −1.255; costs to player 1: 177.429; costs to player 2: 145.493.

Table 7.2. *Markov perfect Nash equilibrium*

$s(t)$	$m(t)$	$m^*(t)$	t
25.000	14.840	−7.111	0.0
22.308	12.242	−6.345	0.1
19.907	11.817	−5.662	0.2
17.763	10.544	−5.052	0.3
15.851	9.409	−4.508	0.4
14.144	8.396	−4.023	0.5
12.622	7.492	−3.590	0.6
11.263	6.686	−3.203	0.7
10.050	5.966	−2.858	0.8
8.968	5.324	−2.551	0.9
8.083	4.732	−2.321	1.0
2.531	1.943	−0.703	2.0
0.912	0.772	−0.159	3.0
0.202	0.154	−0.053	4.0
0.071	0.042	−0.015	5.0

Notes: Stable eigenvalue: −1.139; costs to player 1: 185.502; costs to player 2: 148.265.

Table 7.3. *Minimization of the sum of both countries' costs*

$s(t)$	$m(t)$	$m^*(t)$	t
25.000	24.164	−30.000	0.0
21.111	20.405	−25.333	0.1
17.826	17.230	−21.392	0.2
15.053	14.550	−18.064	0.3
12.711	12.286	−15.252	0.4
10.734	10.375	−12.880	0.5
9.064	8.761	−10.876	0.6
7.654	7.398	−9.184	0.7
6.463	6.247	−7.755	0.8
5.457	5.275	−6.549	0.9
4.608	4.454	−5.530	1.0
0.849	0.821	−1.019	2.0
0.157	0.151	−0.188	3.0
0.029	0.028	−0.035	4.0
0.005	0.005	−0.006	5.0

Notes: Stable eigenvalue: −1.691; costs to player 1: 178.725; costs to player 2: 123.478.

horizon) to both players, for the open-loop Nash equilibrium, the Markov perfect Nash equilibrium, and the joint maximization solution.

The numerical results confirm the impressions obtained from the previous analytical discussion. The speed of convergence does not differ too much between the equilibria, with the Markov perfect equilibrium having the largest eigenvalue and, hence, the slowest convergence.

7.2 Linear state games

Linear quadratic differential games as discussed in the preceding section are characterized by the property that the system dynamics is a first order polynomial and the utility functions are second order polynomials with respect to the state and the control variables. The present section considers another analytically tractable class of differential games. It consists of those games for which the system dynamics and the utility functions are polynomials of degree 1 with respect to the state variables and which satisfy a certain property (described below) concerning the interaction between control variables and state variables. We call this class of games linear state games. It will be shown that these games have the property that their open-loop Nash equilibria are Markov perfect.

Let us start with a two-person differential game with state equations given by

$$\dot{x}(t) = f(x(t), u^1(t), u^2(t), t), \tag{7.33}$$

where $u^1(t) \in \mathbb{R}^{m^1}$ and $u^2(t) \in \mathbb{R}^{m^2}$ are the control variables of player 1 and player 2, respectively, and $x(t) \in \mathbb{R}^n$ is an n-dimensional vector of state variables. The objective functional of player i is given by

$$J^i = \int_0^T e^{-rt} L^i(x(t), u^1(t), u^2(t), t)\, dt + e^{-rT} S^i(x(T)). \tag{7.34}$$

We define the function $\tilde{H}^i : \mathbb{R}^{2n+m^1+m^2} \times [0, T] \mapsto \mathbb{R}$ by

$$\tilde{H}^i(x, u^1, u^2, p^i, t) = L^i(x, u^1, u^2, t) + p^i f(x, u^1, u^2, t),$$

where $p^i \in \mathbb{R}^n$ is a vector of costate variables. The reader should note the close relationship of this function to player i's Hamiltonian function. As a matter of fact, the latter is given by $H^1(x, u^1, p^1, t) = \tilde{H}^1(x, u^1, \phi^2(x, t), p^1, t)$ for player 1 and by $H^2(x, u^2, p^2, t) = \tilde{H}^2(x, \phi^1(x, t), u^2, p^2, t)$ for player 2 (here we denote as usual the strategy of player i by ϕ^i). The class of linear state games is defined by the structural properties of the functions $\tilde{H}^i$ and S^i. More specifically, a differential game is referred to as a linear state game if the conditions

$$\tilde{H}^i_{xx}(x, u^1, u^2, p^i, t) = S^i_{xx}(x) = 0 \tag{7.35}$$

and

$$\tilde{H}^i_{u^i}(x, u^1, u^2, p^i, t) = 0 \;\Rightarrow\; \tilde{H}^i_{u^i x}(x, u^1, u^2, p^i, t) = 0 \tag{7.36}$$

hold for $i = 1, 2$ and all $(x, u^1, u^2, p^i, t) \in \mathbb{R}^{2n+m^1+m^2} \times [0, T]$. Note that (7.36) is automatically satisfied if

$$\tilde{H}^1_{u^i x}(x, u^1, u^2, p^1, t) = \tilde{H}^2_{u^i x}(x, u^1, u^2, p^2, t) = 0 \tag{7.37}$$

holds for $i = 1, 2$ and all $(x, u^1, u^2, p^1, p^2, t) \in \mathbb{R}^{3n+m^1+m^2} \times [0, T]$. Conditions (7.36) and (7.37) are ordered in increasing strength. Let us discuss and illustrate them in reverse order.

Condition (7.37) implies that there is no multiplicative interaction at all between the state and the control variables in the game. In terms of the state equations, the objective functions, and the salvage value term this implies

$$\begin{aligned} f(x, u^1, u^2, t) &= A(t)x + g(u^1, u^2, t), \\ L^i(x, u^1, u^2, t) &= C^i(t)x + k^i(u^1, u^2, t), \\ S^i(x) &= W^i x, \end{aligned}$$

where $A : [0, T] \mapsto \mathbb{R}^{n \times n}$, $g : \mathbb{R}^{m^1+m^2} \times [0, T] \mapsto \mathbb{R}^n$, $C^i : [0, T] \mapsto \mathbb{R}^n$, $k^i : \mathbb{R}^{m^1+m^2} \times [0, T] \mapsto \mathbb{R}$, and $W^i \in \mathbb{R}^n$.

Example 7.1 As an example of a game that has the special structure just described, consider the situation in which there are two individuals who invest in a public stock of knowledge. Let $x(t)$ be the stock of knowledge at time t and $u^i(t)$ the investment of agent i in public knowledge at time t. The stock of knowledge evolves according to the accumulation equation

$$\dot{x}(t) = u^1(t) + u^2(t) - \alpha x(t),$$

where α is a constant rate of depreciation. Investment is costly and the cost function of agent i is given by $K^i(u^i(t))$. If each agent derives linear utility from the consumption of the stock of knowledge, the objective functions are given by

$$J^i = \int_0^T e^{-rt}\left[x(t) - K^i(u^i(t))\right] dt + e^{-rT} W^i x(T),$$

where r is the common discount rate and W^i is the terminal value of knowledge for player i. (See also section 9.5.)

Next consider condition (7.36). To understand this condition, first note that there are no control constraints. Thus, if the pair $(u^1(\cdot), u^2(\cdot))$ of open-loop strategies satisfies the sufficient equilibrium conditions stated in theorem 4.2 it holds that

$$(\partial/\partial u^i)\tilde{H}^i(x(t), u^1(t), u^2(t), p^i(t), t) = 0.$$

Condition (7.36) says that this condition does not involve the state variables. Consequently, the control paths $u^1(\cdot)$ and $u^2(\cdot)$ are uniquely determined by the costate trajectories. The adjoint equations and the transversality condition, too, do not involve the state variables of the problem because of condition (7.35). Thus, one can also compute the costate variables (and, consequently, the control paths) independently from the state variables. In particular, the open-loop strategies are independent of the initial state x_0. Using arguments similar to those in section 4.3, we can show that this property implies that the open-loop equilibrium is Markov perfect.

Conditions (7.35) and (7.36) are satisfied if the state equations and the objective functions have the structure

$$f(x, u^1, u^2, t) = A(t)x + g(u^1, u^2, t)x,$$
$$L^i(x, u^1, u^2, t) = C^i(t)x + k^i(u^1, u^2, t)x,$$
$$S^i(x) = W^i x,$$

where A, g, C^i, k^i, and W^i are as before.

Example 7.2 Consider two firms, each of which introduces a new durable good. Let M be the total market potential for both firms and $X^i(t)$ accumulated sales of firm i up to time t. The sales dynamics are influenced by saturation effects so that the sales rates at time t are proportional to the remaining market potential, $s^i(t) = \alpha^i(t)[M - X^1(t) - X^2(t)]$. One can interpret $\alpha^i(t)$ as the rate at which current nonadopters of the product are transformed into adopters of firm i's product at time t. If we assume that $\alpha^i(t)$ depends on the current prices $u^1(t)$ and $u^2(t)$ charged by the firms we get the equations

$$\dot{X}^i(t) = s^i(t) = \alpha^i(u^1(t), u^2(t))\big[M - X^1(t) - X^2(t)\big].$$

If firm i faces constant average production costs c^i and if the current level of production is identical to current sales we get the objective functionals

$$J^i = \int_0^T e^{-rt}\big[u^i(t) - c^i\big]\alpha^i(u^1(t), u^2(t))\big[M - X^1(t) - X^2(t)\big]\,dt.$$

Defining the state of the game by $x(t) = M - X^1(t) - X^2(t)$ we get the state dynamics

$$\dot{x}(t) = -\big[\alpha^1(u^1(t), u^2(t)) + \alpha^2(u^1(t), u^2(t))\big]x(t),$$

and the objective functions can be written as

$$J^i = \int_0^T e^{-rt}\big[u^i(t) - c^i\big]\alpha^i(u^1(t), u^2(t))x(t)\,dt,$$

which fits exactly the structure outlined above.

The main structural properties of linear state games are the linearity in the state variable and the lack of multiplicative interactions between the state and the control variables. There exist games that possess the same qualitative characteristics as linear state games, but which do admit multiplicative interactions between the state and the control variables. To see this let us look at an example in which there are two state variables (x^1, x^2) and two control variables (u^1, u^2). In particular we assume

$$
\begin{aligned}
f^1(x^1, u^1, u^2, t) &= A^1(t)x^1 + g^1(u^1, u^2, t) + x^1 h^1(u^2, t),\\
f^2(x^2, u^1, u^2, t) &= A^2(t)x^2 + g^2(u^1, u^2, t) + x^2 h^2(u^1, t),\\
L^1(x^1, u^1, u^2, t) &= C^1(t)x^1 + k^1(u^1, u^2, t),\\
L^2(x^2, u^1, u^2, t) &= C^2(t)x^2 + k^2(u^1, u^2, t),\\
S^i(x^i) &= W^i x^i.
\end{aligned}
$$

Games with this structure are characterized by two facets. Firstly, we can associate a state variable with each player. Secondly, the state equation and the objective function of each player are not influenced by the state variable of the opponent. This implies that, in an open-loop Nash equilibrium, the corresponding costate variables are redundant in the sense that they do not affect the equilibrium play of the agents. It is, however, possible that there is a mixed multiplicative interaction between the state variable of player i and the control variables of player j and we still get a Markov perfect open-loop Nash equilibrium.

Example 7.3 As an example of the structure just discussed, consider the case in which there are two firms each investing in the individual stock of goodwill through advertising. Let $x^i(t)$ be the goodwill stock of firm i at time t and $u^i(t)$ its current level of advertising activities. The state equations are given by

$$\dot{x}^i(t) = u^i(t) - \alpha^i(u^j(t))x^i(t),$$

where the rate of depreciation of the stock of goodwill depends on the current level of advertising activities of the rival firm. While advertising is costly, the firm generates revenues that are proportional to its current stock of goodwill. If the cost functions are given by $K^i(u^i(t))$, the objective functions become

$$J^i = \int_0^T e^{-rt}\left[\pi^i x^i(t) - K^i(u^i(t))\right] dt + e^{-rT} W^i x^i(T),$$

where π^i is a positive constant that measures the profitability of the existing stock of goodwill.

Let us examine some other special properties of linear state games and see how they imply the Markov perfectness of open-loop Nash equilibria. We start by deriving the adjoint equations. Since (7.35) and (7.36) hold, the costate equations are independent of the state variable and given by

$$\dot{p}^i(t) = B^i(u^1(t), u^2(t), t)$$

for appropriate functions $B^i : \mathbb{R}^{m^1+m^2} \times [0, T] \mapsto \mathbb{R}^n$. The transversality conditions are $p^i(T) = W^i$. Both the adjoint equations and the transversality conditions do not involve the state variables. From condition (7.36) it follows that the Hamiltonian maximizing conditions allow us to write the controls as functions of calendar time and the adjoint variables only, i.e.,

$$u^i(t) = \mu^i(p^1(t), p^2(t), t). \tag{7.38}$$

Substituting this into the adjoint equations results in a terminal value problem for the costate variables given by

$$\dot{p}^i(t) = B^i(\mu^1(p^1(t), p^2(t), t), \mu^2(p^1(t), p^2(t), t), t)\,, \qquad p^i(T) = W^i. \tag{7.39}$$

This system can be solved for the costate trajectories. Substituting the result into (7.38) gives the open-loop strategies. Note that the maximized Hamiltonian functions are always linear with respect to the state variables so that the sufficient equilibrium conditions of theorem 4.2 hold. Since the solution of (7.39) is independent of the initial states, the same is true for the equilibrium strategies. This property implies that the open-loop Nash equilibria for linear state games are Markov perfect.

As an alternative to the approach outlined above (which is based on the conditions of the maximum principle) one can also solve these games by the dynamic programming approach. To illustrate this we now derive Markov perfect open-loop Nash equilibria for each of the three examples introduced above using the HJB equations.

Example 7.1 (continued) We want to derive a Markov perfect open-loop Nash equilibrium $(u^1(\cdot), u^2(\cdot))$ of the knowledge accumulation game. Denoting the value functions by $V^i(x, t)$, the HJB equations can be written as

$$rV^i(x, t) - V^i_t(x, t) = \max\{x - K^i(u^i) + V^i_x(x, t)[u^i + u^j(t) - \alpha x] \mid u^i \in \mathbb{R}\}.$$

We make the informed guess that the value functions have the linear form

$$V^i(x, t) = a_i(t)x + b_i(t).$$

Using this conjecture, the maximization on the right-hand side of the HJB equation implies that

$$\frac{dK^i(u^i(t))}{du^i} = a_i(t)$$

for all t. This equation can be solved to give $u^i(t)$ as a function of $a_i(t)$, say $u^i(t) = \mu^i(a_i(t))$. Substituting this back into the HJB equation, collecting terms with equal powers of x, and equating the coefficients of these powers to 0 yields

$$\dot{a}_i(t) = (r+\alpha)a_i(t) - 1,$$

$$\dot{b}_i(t) = rb_i(t) + K^i(\mu^i(a_i(t))) - a_i(t)[\mu^1(a_1(t)) + \mu^2(a_2(t))].$$

To satisfy the boundary conditions for the optimal value functions we require $a_i(T) = W^i$ and $b^i(T) = 0$. From these conditions we get

$$a^i(t) = \frac{1}{r+\alpha} + \left(W^i - \frac{1}{r+\alpha}\right)e^{-(r+\alpha)(T-t)}.$$

The Markov perfect Nash equilibrium is therefore completely characterized.

Example 7.2 (continued) Now we derive a Markov perfect Nash equilibrium for the pricing problem discussed above. Let us denote the equilibrium strategies by $u^1(\cdot)$ and $u^2(\cdot)$. The HJB equations for this example are

$$\begin{aligned} &rV^i(x,t) - V^i_t(x,t) \\ &= \max\{(u^i - c^i)\alpha^i(u^i, u^j(t))x \\ &-V^i_x(x,t)[\alpha^i(u^i, u^j(t)) + \alpha^j(u^j(t), u^i)]x \mid u^i \in \mathbb{R}\}. \end{aligned}$$

Now we guess value functions of the form

$$V^i(x,t) = a_i(t)x.$$

Substituting this into the HJB equations, the first order conditions for the maximization on the right-hand side are

$$\begin{aligned} &\alpha^i(u^i(t), u^j(t)) + [u^i(t) - c^i]\alpha^i_{u^i}(u^i(t), u^j(t)) \\ &= a_i(t)\left[\alpha^i_{u^i}(u^i(t), u^j(t)) + \alpha^j_{u^i}(u^j(t), u^i(t))\right]. \end{aligned}$$

These two conditions can be solved to get $u^i(t) = \mu^i(a_i(t), a_j(t))$. The same procedure that we have already used in the previous example shows that the functions $a^i(\cdot)$ must satisfy

$$\dot{a}_i(t) = [r + \alpha^i(u^i(t), u^j(t)) + \alpha^j(u^j(t), u^i(t))]a_i(t) - [u^i(t) - c^i]\alpha^i(u^i(t), u^j(t)),$$

where $u^i(t) = \mu^i(a_i(t), a_j(t))$. The terminal conditions for these differential equations are $a_i(T) = 0$.

Example 7.3 (continued) Finally we derive a Markov perfect Nash equilibrium consisting of open-loop strategies $u^1(\cdot)$ and $u^2(\cdot)$ for the advertising game. The HJB equations are

$$\begin{aligned} &rV^i(x^1, x^2, t) - V^i_t(x^1, x^2, t) \\ &= \max\{\pi^i x^i - K^i(u^i) + V^i_{x^i}(x^1, x^2, t)[u^i - \alpha^i(u^j(t))x^i] \\ &+ V^i_{x^j}(x^1, x^2, t)[u^j(t) - \alpha^j(u^i)x^j] \mid u^i \in \mathbb{R}\}. \end{aligned}$$

For this example we conjecture linear value functions of the form[8]

$$V^i(x^i, x^j, t) = a_i(t)x^i + b_i(t).$$

Substituting this into the HJB equations and performing the maximization yields the same first order condition as in example 7.1, namely

$$\frac{dK^i(u^i(t))}{du^i} = a_i(t).$$

Again we can write the solution to this condition as $u^i(t) = \mu^i(a_i(t))$. By the same procedure that we used in example 7.1 we obtain the differential equations

$$\begin{aligned} \dot{a}_i(t) &= [r + \alpha^i(\mu^j(a_j(t)))]a_i(t) - \pi^i, \\ \dot{b}_i(t) &= rb_i(t) + K^i(\mu^i(a_i(t))) - a_i(t)\mu^i(a_i(t)), \end{aligned}$$

and the terminal conditions $a_i(T) = W^i$ and $b_i(T) = 0$. As in example 7.1 we see that the Nash equilibrium strategies are completely characterized once we have found the functions $a_i(\cdot)$ that satisfy the above system.

Summing up, we have demonstrated that the class of linear state games has a very useful property. The linearity in the state variables together with the decoupled structure between the state variables and the control variables implies that the open-loop equilibrium is Markov perfect and that the value functions are linear in the state variables. In the next section we introduce so-called exponential games and show that they can be transformed into linear state games by a simple state variable transformation.

7.3 Exponential games

The third class of analytically tractable differential games that we describe are termed exponential games. These games have the following general structure. The state equations are given by

[8]Note that we conjecture that the optimal value function for player i does not depend on x^j. Alternatively, one could start with the weaker conjecture $V^i(x^i, x^j, t) = a_i(t) = a_i(t)x^i + b_i(t) + c_i(t)x^j$. Following the same steps that we outline in the text, one will find that $c_i(t) = 0$ for all t.

$$\dot{x}(t) = f(u^1(t), u^2(t), t), \tag{7.40}$$

where $u^1(t) \in \mathbb{R}^{m^1}$ and $u^2(t) \in \mathbb{R}^{m2}$ are the control variables and $x(t) \in \mathbb{R}^n$ is the state vector. The objective function of player i is

$$J^i = \int_0^T e^{-rt} L^i(u^1(t), u^2(t), t) e^{-\lambda_i x(t)}\, dt, \tag{7.41}$$

where $\lambda_i \in \mathbb{R}^n$ is a fixed vector. The special structure of exponential games is that the state variables do not enter the right-hand side of the system dynamics and that they enter the objective functions in an exponential way. At first sight these games seem to be very different from the linear state games, but we now show that these two classes are equivalent. To this end we define new state variables by

$$y^i(t) = e^{-\lambda_i x(t)} \qquad i = 1, 2.$$

Differentiating both sides of this equation with respect to time t we get

$$\dot{y}^i(t) = -y^i(t)[\lambda_i \dot{x}(t)] = -y^i(t)\big[\lambda_i f(u^1(t), u^2(t), t)\big].$$

The objective functions can be written in terms of the new state variables as

$$J^i = \int_0^T e^{-rt} L^i(u^1(t), u^2(t), t) y^i(t)\, dt.$$

It is obvious that the transformed game is a linear state game. Therefore we can conclude that open-loop Nash equilibria of exponential games are Markov perfect.

7.4 Further reading

A more detailed discussion about linear quadratic differential games can be found in Başar and Olsder [4] and Mehlmann [176]. Our discussion here follows the paper by Dockner and Neck [60], in which several macroeconomic applications are discussed. Other macroeconomic applications of linear quadratic differential games can be found in Buiter [13], Calvo [14], Cohen and Michel [33], Dockner and Neck [59], Fisher [95], Hughes Hallet [129], and Miller and Salmon [179]. The macroeconomic model used for the numerical example follows Neck and Dockner [186] (see also Turnovsky et al. [234] for a discrete-time version).

The existence of solutions to algebraic Riccati equations is discussed in Papavassilopoulos et al. [189]. Discussions of the Riccati differential equations can be found in textbooks on differential equations.

Engwerda [75] deals with existence and uniqueness of open-loop Nash equilibria for linear quadratic games. In Engwerda [76] the existence of solutions to the Riccati equation associated with a Markov perfect Nash equilibrium of a game with scalar state equation and infinite horizon is discussed. Lockwood [162] studies the uniqueness of a Markov perfect Nash equilibrium in linear quadratic games with infinite horizon.

The class of linear state games was introduced by Dockner et al. [49]. In that paper the authors do not refer to these games as linear state games but they call them qualitatively solvable games. Linear state games are also discussed in Mehlmann [176] and Fershtman [88]. The class of linear state games has been applied to many different fields of economics and management science. Applications include the limit pricing problem in Dockner [45], a game between a thief and the police in Feichtinger [82], duopolistic pricing problems as described in Feichtinger and Dockner [84], and advertising games as in Jorgensen [131] or Feichtinger and Dockner [83].

Example 7.1 is a variation of the model used by Fershtman and Nitzan [94] and Wirl [244]. Example 7.3 is taken from Leitmann and Schmitendorf [158]. Example 7.2 is due to Dockner [43].

Exponential differential games were first introduced by Reinganum [199] and applied to models of R&D by Reinganum [197, 198]. This class of games is also studied in Mehlmann [176], where state dependent and state independent Markovian strategies are derived. (See also sections 10.2 and 10.4 below.)

7.5 Exercises

1. Consider the knowledge accumulation game of example 7.1. Assume that the cost function is quadratic, i.e., $K^i(u^i) = (1/2)(u^i)^2$. Derive an open-loop Nash equilibrium for the finite horizon game. Show that the equilibrium strategies of the finite horizon game converge to equilibrium strategies of the infinite horizon game as the planning horizon, T, approaches infinity.
2. Consider example 7.2 with an infinite time horizon and assume that $\alpha^i(u^i, u^j) = a - bu^i + d(u^j - u^i)$, where a, b, and d are positive constants. Show that there exists a Markov perfect equilibrium consisting of constant strategies. Prove that the equilibrium prices which correspond to this equilibrium are higher than the prices of a static Bertrand model with the demand curve given by $\alpha^i(u^i, u^j)$ and marginal costs given by c^i.
3. In deriving the Markov perfect equilibrium of example 7.2 we made use of the dynamic programming approach (the HJB equation).

Derive the Markov perfect equilibrium by applying the maximum principle, but now using an infinite time horizon.

4. Show by means of an appropriate state variable transformations that the game with objective functionals

$$J^i = \int_0^T e^{-rt} L^i(u^1(t), u^2(t), t)[y(t)]^\alpha \, dt$$

and the state equations

$$\dot{y}(t) = y(t) f(u^1(t), u^2(t), t)$$

fits into the class of linear state games.

5. Consider the following simple advertising differential game in which $x(t)$ is the market potential of firm 1 and $1 - x(t)$ that of firm 2. The state dynamics is given by

$$\dot{x}(t) = \ln u^1(t) - \ln u^2(t),$$

where $u^i(t)$ are the advertising expenditures of firm i at time t. The discounted stream of profits of firm 1 is given by

$$J^1 = \int_0^T e^{-rt} \{\pi^1 x(t) - [u^1(t)]^2\} \, dt$$

and that of firm 2 is

$$J^2 = \int_0^T e^{-rt} \{\pi^2 [1 - x(t)] - [u^2(t)]^2\} \, dt.$$

Show that the open-loop equilibrium advertising strategies are characterized by a system of differential equations in the control variables. Discuss the qualitative properties of this equilibrium by means of a phase diagram analysis.

Appendix

N-player affine quadratic differential games

In section 7.1 we introduced a simple example of a linear quadratic game and derived open-loop and Markov perfect Nash equilibria. In this appendix we apply the techniques outlined in the main text to a rather general class of affine quadratic differential games with N players. In such a game the state equations are given by

$$\dot{x}(t)=f(x(t),u^1(t),u^2(t),\ldots,u^N(t),t):=A(t)x(t)+\sum_{i=1}^{N}B^i(t)u^i(t)+c(t),$$
(7.42)

where $u^i(t)\in\mathbb{R}^{m^i}$ are the control variables and $x(t)\in\mathbb{R}^n$ is an n-dimensional vector of state variables. $A(t)$ and $B(t)$ are matrices of appropriate dimensions and $c(t)$ is a vector of dimension n. Restricting ourselves to affine quadratic games with finite time horizon T, the objective functional of player i is

$$J^i=\int_0^T L^i(x(t),u^1(t),u^2(t),\ldots,u^N(t),t)\,dt+S^i(x(T)), \tag{7.43}$$

where

$$L^i(x,u^1,u^2,\ldots,u^N,t)=\frac{1}{2}\left[x'Q^i(t)x+\sum_{j=1}^{N}(u^j)'R^{ij}(t)u^j\right].$$

Here, $Q^i(t)\in\mathbb{R}^{n\times n}$ and $R^{ij}(t)\in\mathbb{R}^{m^j\times m^j}$ are symmetric matrices for all $i,j\in\{1,2,\ldots,N\}$ and $Q^i(t)$ and $R^{ii}(t)$ are negative definite for all $i\in\{1,2,\ldots,N\}$. Furthermore,

$$S^i(x)=x'W^ix,$$

where $W^i\in\mathbb{R}^{n\times n}$ is symmetric and negative definite.

Theorem 7.1 *Consider the N-player affine quadratic differential game (7.42)–(7.43) and assume that there exists a unique solution $(M^1(\cdot),M^2(\cdot),\ldots,M^N(\cdot))$ of the system of Riccati equations*

$$\dot{M}^i(t)=-M^i(t)A(t)-A'(t)M^i(t)-Q^i(t)+M^i(t)\sum_{j=1}^{N}B^j(t)R^{jj}(t)^{-1}B^j(t)'M^j(t)$$

with the terminal condition $M^i(T)=W^i$, $i\in\{1,2,\ldots,N\}$. Let $(m^1(\cdot),m^2(\cdot),\ldots,m^N(\cdot))$ be the unique solution of the system of linear differential equations

$$\dot{m}^i(t)=-A(t).'m^i(t)-M^i(t)c(t)+M^i(t)\sum_{j=1}^{N}B^j(t)R^{jj}(t)^{-1}B^j(t)'m^j(t)$$

with boundary condition $m^i(T)=0$, $i\in\{1,2,\ldots,N\}$, and let $x^*(\cdot)$ be the unique solution of the linear differential equation

$$\dot{x}^*(t) = \left[A(t) - \sum_{i=1}^{N} B^i(t)R^{ii}(t)^{-1}B^i(t)'M^i(t)\right]x^*(t) + c(t)$$
$$- \sum_{i=1}^{N} B^i(t)R^{ii}(t)^{-1}B^i(t)'m^i(t)$$

satisfying the initial condition $x^*(0) = x_0$. *Then the N-tuple* $(\phi^1, \phi^2, \ldots, \phi^N)$ *defined by*

$$\phi^i(t) = -R^{ii}(t)^{-1}B^i(t)'\left[M^i(t)x^*(t) + m^i(t)\right]$$

is the unique open-loop Nash equilibrium of the game.

Proof The proof of this theorem follows by applying the maximum principle in exactly the same way as in the main text.■

It is not only possible completely to characterize an open-loop Nash equilibrium but also to derive a Markov perfect Nash equilibrium for the game (7.42)–(7.43).

Theorem 7.2 *Consider the N-player affine quadratic differential game (7.42)–(7.43) and assume that there exists a unique solution* $(W^1(\cdot), W^2(\cdot), \ldots, W^N(\cdot))$ *of the system of Riccati equations*

$$\dot{W}^i(t) = -W^i(t)F(t) - F'(t)W^i(t) - Q^i(t)$$
$$- \sum_{j=1}^{N} W^j(t)B^j(t)R^{jj}(t)^{-1}R^{ij}(t)R^{jj}(t)^{-1}B^j(t)'W^j(t)$$

with the terminal condition $W^i(T) = W^i$, $i \in \{1, 2, \ldots, N\}$, *where*

$$F(t) = A(t) - \sum_{k=1}^{N} B^k(t)R^{kk}(t)^{-1}B^k(t)'W^k(t).$$

Let $(w^1(\cdot), w^2(\cdot), \ldots, w^N(\cdot))$ *be the solution of the system of coupled linear differential equations*

$$\dot{w}^i(t) = -F(t)'w^i(t) - W^i(t)c(t) + W^i(t)\left[\sum_{j=1}^{N} B^j(t)R^{jj}(t)^{-1}B^j(t)'w^j(t)\right]$$
$$- \sum_{j=1}^{N} W^j(t)B^j(t)R^{jj}(t)^{-1}R^{ij}(t)R^{jj}(t)^{-1}B^j(t)'w^j(t)$$

with boundary condition $w^i(T) = 0$, $i \in \{1, 2, \ldots, N\}$, *and let* $(z^1(\cdot), z^2(\cdot), \ldots, z^N(\cdot))$ *be the solution of*

$$\dot{z}^i(t) = \left[-c(t) + \sum_{j=1}^{N} B^j(t)R^{jj}(t)^{-1}B^j(t)'w^j(t)\right]' w^i(t)$$
$$- 1/2 \sum_{j=1}^{N} w^j(t)B^j(t)R^{jj}(t)^{-1}R^{ij}(t)R^{jj}(t)^{-1}B^j(t)'w^j(t)$$

with boundary condition $z^i(T) = 0$, $i \in \{1, 2, \ldots, N\}$. *Then the N-tuple* $(\psi^1, \psi^2, \ldots, \psi^N)$ *defined by*

$$\psi^i(x, t) = -R^{ii}(t)^{-1}B^i(t)'\left[W^i(t)x + w^i(t)\right]$$

is a Markov perfect Nash equilibrium of the game. The optimal value function for this equilibrium is given by

$$V^i(x, t) = \frac{1}{2}x'W^i(t)x + x'w^i(t) + z^i(t). \tag{7.44}$$

Proof The proof follows from applying the HJB equation and assuming that the value function is of the form given in (7.44). ■

8 Stochastic differential games

So far we have considered only differential games in which the fundamentals (utility functions, system dynamics, initial state, etc.) do not contain any uncertainty. For some applications this is quite satisfactory, but for others it presents a severe limitation. We now take a look at stochastic optimal control problems and differential games in which some of the fundamentals involve random variables or stochastic processes. Uncertainty can be incorporated into our basic framework in many different ways, so we shall not give a complete treatment of this matter but concentrate on two forms of stochastic models which seem to be most useful in applications. The first is based on so-called piecewise deterministic processes whereas the second makes use of the Wiener process. The notion of piecewise deterministic processes captures the idea that uncertain changes in the system occur at discrete (but random) time instants. Between these so-called jump times, the system evolves in a deterministic way. The Wiener process, on the other hand, is used to model situations characterized by continuous stochastic noise.

8.1 Piecewise deterministic games

8.1.1 A piecewise deterministic control model

A piecewise deterministic process is a system which evolves in a deterministic way, except at certain jump times $T_1, T_2, \dots$ at which the deterministic law of motion switches from one mode to another. Both the jump times T_l and the system modes which govern the motion between jump times are randomly selected. We restrict ourselves to models where there are finitely many different modes and where the evolution of the process between the jump times is described by a deterministic differential equation which may depend on the current mode. Furthermore, we concentrate on the case of an autonomous problem defined over the unbounded time interval $[0, \infty)$ because this is the case most often encountered in applications. We assume that the discount rate is strictly positive and that the utility function is bounded so that the present value of utility is finite for all feasible control paths. The changes that have to be made in order to deal with a non-autonomous problem are straightforward and will be briefly outlined at the end of the section.

Let $X \subseteq \mathbb{R}^n$ denote the state space of the model as in chapter 3 and let M be a finite set of modes. One may think of the mode of the system as an additional state variable with discrete values. We denote by $x(t)$ the state at time t and by $u(t)$ the control value chosen at time t. The set of controls, which are feasible at time t when the current mode is $h \in M$ and the state is equal to $x(t) \in X$, is given by $U(h, x(t)) \subseteq \mathbb{R}^m$. The evolution of the state during an interval in which the mode of the system equals h is described by the differential equation $\dot{x}(t) = f(h, x(t), u(t))$, where $f(h, \cdot, \cdot)$ maps the set $\Omega(h) = \{(x, u) \mid x \in X, u \in U(h, x)\}$ into $\mathbb{R}^n$. The instantaneous payoff rate at time t, when the current mode is equal to h, is given by $F(h, x(t), u(t))$, where $F(h, \cdot, \cdot)$ is a real-valued function defined on $\Omega(h)$. In addition, the decision maker receives the lump sum payoff $S_{hk}(x(t))$ if a jump from mode h to mode $k \neq h$ occurs at time t. The function $S_{hk}(\cdot)$ is defined on X and has values in $\mathbb{R}$. All payoffs are discounted at the constant rate $r > 0$ and the initial state and mode are deterministic constants, $x_0 \in X$ and $h_0 \in M$, respectively.

If there were only a single mode $h \in M$ we would be in the situation described in chapter 3. The present model differs from that of chapter 3 because the dynamic system can switch between modes in a nondeterministic way.

The mathematical tool to describe the evolution of the system mode as a function of time is a continuous-time stochastic process $h : [0, \infty) \times \Xi \mapsto M$. Here, Ξ is a set of points ξ representing possible realizations of some random phenomenon (like a coin toss). Subsets of Ξ are called events and each event has a certain probability. If the event A is characterized as the set of all those $\xi \in \Xi$ which satisfy a certain condition a, then we denote the probability of A simply by $\text{Prob}\{a\}$. For example, the event that the mode of the system at time t is equal to h is $\{\xi \in \Xi \mid h(t, \xi) = h\}$ and its probability is denoted by $\text{Prob}\{h(t, \xi) = h\}$. Quite often the argument ξ is suppressed for notational simplicity so that the probability is written as $\text{Prob}\{h(t) = h\}$. Conditional probabilities play a major role in probability theory and are also important in the description of the stochastic processes in the present chapter. If A and B are two events characterized by conditions a and b, respectively, and if $\text{Prob}\{b\} > 0$, then the probability of A conditional on the occurrence of B is given by $\text{Prob}\{a \mid b\} = \text{Prob}\{a \text{ and } b\}/\text{Prob}\{b\}$. Probability theory tells us how to calculate conditional probabilities even if $\text{Prob}\{b\} = 0$; we assume that the reader is familiar with these rules.

Returning to our stochastic control model let us see how we can describe the probability law which governs the switches between modes. Assume that for every pair of modes $(h, k) \in M \times M$ with $h \neq k$ there exists a function $q_{hk} : \Omega(h) \mapsto \mathbb{R}_+$ such that the following relation holds

$$\lim_{\Delta \to 0} \frac{1}{\Delta} \text{Prob}\{h(t + \Delta) = k \mid h(t) = h\} = q_{hk}(x(t), u(t)) \qquad k \neq h. \tag{8.1}$$

This means that the probability that the system switches from mode h to another mode k during the short time interval $(t, t + \Delta]$, given that the mode was h at time t, is, to a first approximation, proportional to the length of the interval Δ. The factor of proportionality is equal to $q_{hk}(x(t), u(t))$. In addition to (8.1) we assume that the probability that two or more switches occur during the interval $(t, t + \Delta]$, divided by Δ, converges to 0 as Δ approaches 0.

Throughout this section we assume that the functions $f(h, \cdot)$, $F(h, \cdot)$, $S_{hk}(\cdot)$, and $q_{hk}(\cdot)$ are sufficiently smooth and that $F(h, \cdot)$, $S_{hk}(\cdot)$, and $q_{hk}(\cdot)$ are bounded. Given the initial state $x_0 \in X$, the initial mode $h_0 \in M$, and the control path $u(\cdot)$, the system dynamics

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

and equation (8.1) determine a unique stochastic process $h : [0, \infty) \times \Xi \mapsto M$.[1] This process is called a piecewise deterministic process and it is known to be continuous from the right. This means that $\lim_{k \to \infty} h(s_k, \xi) = h(t, \xi)$ holds for almost all $\xi \in \Xi$, for all $t \in [0, \infty)$, and for every sequence of real numbers $s_k > t$ with $\lim_{k \to \infty} s_k = t$. For simplicity, we do not make the dependence of the process on the initial conditions explicit in the notation. However, we emphasize that the probability law governing the process depends also on the chosen control path $u(\cdot)$ and we henceforth write $\text{Prob}_{u(\cdot)}$ or $E_{u(\cdot)}$ to denote probabilities or expectations computed with respect to that law. We also emphasize that not only is $h(\cdot)$ a stochastic process but so is $x(\cdot)$. This follows from the fact that the system dynamics f depend on the realization of $h(\cdot)$. A correct notation for the state at time t would therefore be $x(t, \xi)$ but, as before, we suppress the argument ξ and simply write $x(t)$.

[1]We do not explicitly state the smoothness assumptions for f and q_{hk} which are necessary for the stochastic process to be well defined. The interested reader may consult the literature cited in section 8.3 for more details. We maintain this somewhat informal level throughout the chapter.

Different realizations of the stochastic process (i.e., different realizations $\xi \in \Xi$) result in different payoff streams for the agent even if the control path is fixed. Therefore, we assume that the decision maker maximizes the expectation of the discounted payoff stream, conditional on the given initial state and mode. In other words, he maximizes the objective functional

$$J(u(\cdot)) = E_{u(\cdot)}\left\{ \int_0^\infty e^{-rt} F(h(t), x(t), u(t))\, dt + \sum_{l \in \mathbb{N}} e^{-rT_l} S_{h(T_l-)h(T_l)}(x(T_l)) \,\Big|\, x(0) = x_0, h(0) = h_0 \right\}$$

over the set of feasible control paths $u(\cdot)$. Here T_l denotes the lth jump time, $h(T_l-)$ is the mode of the system immediately before the switch, and $h(T_l)$ is the mode after the switch. It can be shown that the above definition of $J(u(\cdot))$ is equivalent to

$$J(u(\cdot)) = E_{u(\cdot)}\left\{ \int_0^\infty e^{-rt}\Big[F(h(t), x(t), u(t)) + \sum_{k \neq h(t)} q_{h(t)k}(x(t), u(t)) S_{h(t)k}(x(t))\Big]\, dt \,\Big|\, x(0) = x_0, h(0) = h_0 \right\}. \tag{8.2}$$

At this point we have to comment on what is meant by feasible control paths. In section 3.5 we pointed out that certain representations of control paths (i.e., certain strategies) do not make sense as they involve inaccessible information or future values of endogenous variables. In a stochastic setting this issue is even more relevant because uncertainty is resolved only gradually as time evolves. To explain this in more detail let us consider an example where $h(t)$ describes the set of energy resources that can be exploited at time t.[2] This can be considered as a stochastic process since it is not known which new technologies are going to be invented in the future or at which time technological breakthroughs will occur. Firms engaged in energy research and development can directly influence the rate at which new inventions are made by allocating more or fewer resources to their research and development departments. It is clear that the profit rate of such a firm depends on the current mode of the system. But also firms with no research activities depend crucially on the available energy resources and must take into account the possibility of future technological developments in the energy sector. Therefore, such a firm faces a stochastic optimization problem of the

[2] In this interpretation the system mode as of, for example, 2000 would certainly include oil, solar energy, or nuclear fission but not nuclear fusion.

sort described in this section. There are two important points to note. The first one is that the firm should use the most recent information in order to determine its actions. If a new technology already has been developed, a firm engaged in research and development may decide to abandon its own research activities in that field and start a different research programme. This implies that control paths are random processes like the state and mode processes. On the other hand, it is clear that even if the decisions of such a firm may depend on the possibility of future events, they cannot depend on those events themselves. In particular, a decision rule for the control $u(t, \xi)$ can never have the form $u(t, \xi) = \phi(h(t + \tau, \xi), \ldots)$ where τ is strictly positive. If $h(\cdot)$ is a random process then its realization at time $t + \tau$ is not known at time t, so that decisions at time t cannot depend on the future mode of the system (although they may very well depend on the probability distribution of the future modes at time $t + \tau$, conditional on the information available at time t). In short, a feasible strategy for an optimal action at time t can only depend on information that has been revealed by that time. We call such a strategy (and the corresponding control path) nonanticipating.[3]

In mathematical terms, one would require that the control at time t be a measurable function with respect to the sigma-algebra generated by the realizations of the stochastic process $(h(\cdot), x(\cdot))$ up to time t. This sigma-algebra consists of all those events which can be distinguished by observations of $h(\cdot)$ and $x(\cdot)$ up to time t. Thus, a nonanticipating strategy $u(t, \xi) = \phi(\ldots)$ must not contain arguments $h(t + \tau, \xi)$ or $x(t + \tau, \xi)$ on the right-hand side for any $\tau > 0$. In the remainder of this chapter, we do not make the dependence of $u(\cdot)$ on $\xi \in \Xi$ explicit; we simply write $u(t)$ unless special emphasis has to be put on this dependence.

Definition 8.1 A control path $u : [0, \infty) \times \Xi \mapsto \mathbb{R}^m$ is feasible for the stochastic optimal control problem stated above if it is nonanticipating, if the piecewise deterministic process $(h(\cdot), x(\cdot))$ is well defined, if the constraints $x(t) \in X$ and $u(t) \in U(h(t), x(t))$ are satisfied with probability 1 for all t, and if the integral in (8.2) is well defined. The control path $u(\cdot)$ is optimal if it is feasible and if $J(u(\cdot)) \geq J(\tilde{u}(\cdot))$ holds for all feasible control paths $\tilde{u}(\cdot)$.

Note that we do not have to distinguish between the various optimality criteria discussed in section 3.6 because we have assumed that $r > 0$ and that F, q_{hk}, and S_{hk} are bounded functions so that the integral in (8.2) is finite. Before we can prove the main result of this section we need the

[3]See section 6.1 for a discussion of nonanticipating strategies in a deterministic framework.

following, technical lemma. The proof of the lemma is beyond the scope of the book and can be found in the literature cited in section 8.3.

Lemma 8.1 *Let* $\tilde{u}(\cdot)$ *be a feasible control path and denote by* $\tilde{h}(\cdot)$ *and* $\tilde{x}(\cdot)$ *the corresponding mode and state processes. Furthermore, let* $V : M \times X \mapsto \mathbb{R}$ *be any bounded function such that* $V(h, x)$ *is continuously differentiable in* x *for all* $h \in M$. *Then it holds that*

$$\begin{aligned}
&- V(h, x) \\
&= E_{\tilde{u}(\cdot)}\Bigg\{\int_0^\infty e^{-rt}\Big[V_x(\tilde{h}(t), \tilde{x}(t))f(\tilde{h}(t), \tilde{x}(t), \tilde{u}(t)) - rV(\tilde{h}(t), \tilde{x}(t)) \qquad (8.3)\\
&+ \sum_{k \neq \tilde{h}(t)} q_{\tilde{h}(t)k}(\tilde{x}(t), \tilde{u}(t))[V(k, \tilde{x}(t)) \\
&- V(\tilde{h}(t), \tilde{x}(t))]\Big]\, dt \,\Big|\, \tilde{x}(0) = x, \tilde{h}(0) = h\Bigg\}.
\end{aligned}$$

Applying a heuristic argument similar to the one used in section 3.2 to motivate equation (3.7), one can derive the following HJB equation for the stochastic optimal control problem at hand

$$\begin{aligned}
rV(h, x) = \max \Bigg\{ &F(h, x, u) + V_x(h, x)f(h, x, u) \\
&+ \sum_{k \neq h} q_{hk}(x, u)[S_{hk}(x) + V(k, x) - V(h, x)] \,\Big|\, u \in U(h, x)\Bigg\}.
\end{aligned}$$

Note that, in contrast to section 3.2, we are now dealing with an autonomous problem. Therefore, the optimal value function V does not depend explicitly on time t and no term involving the partial derivative V_t shows up in the equation.[4] On the other hand, the possibility of random switches from one mode to another makes it necessary to include the summation term on the right-hand side of the HJB equation. We now state the main result of this section.

Theorem 8.1 *Let* $V : M \times X \mapsto \mathbb{R}$ *be a bounded function such that* $V(h, x)$ *is continuously differentiable in* x *for all* $h \in M$ *and such that the HJB equation*

$$\begin{aligned}
rV(h, x) = \max \Bigg\{ &F(h, x, u) + V_x(h, x)f(h, x, u) \qquad (8.4)\\
&+ \sum_{k \neq h} q_{hk}(x, u)[S_{hk}(x) + V(k, x) - V(h, x)] \,\Big|\, u \in U(h, x)\Bigg\}
\end{aligned}$$

[4]A version of the HJB equation for a nonautonomous problem will be presented after theorem 8.1.

is satisfied for all $(h, x) \in M \times X$. *Let* $\Phi(h, x)$ *denote the set of controls* $u \in U(h, x)$ *maximizing the right-hand side of (8.4). If* $u(\cdot)$ *is a feasible control path, and* $(h(\cdot), x(\cdot))$ *is the corresponding stochastic process describing the mode and the state of the system, and if* $u(t) \in \Phi(h(t), x(t))$ *holds with probability 1 for almost all* $t \in [0, \infty)$ *then* $u(\cdot)$ *is an optimal control path.*

Proof Let $\tilde{u}(\cdot)$ be any feasible control path and denote by $(\tilde{h}(\cdot), \tilde{x}(\cdot))$ the corresponding mode and state processes. Because V is assumed to satisfy the assumptions of lemma 8.1 we know that (8.3) must hold. Using the HJB equation and the fact that $\tilde{u}(t) \in U(\tilde{h}(t), \tilde{x}(t))$ we obtain

$$\begin{aligned} &V_x(\tilde{h}(t), \tilde{x}(t))f(\tilde{h}(t), \tilde{x}(t), \tilde{u}(t)) - rV(\tilde{h}(t), \tilde{x}(t)) \\ &+ \sum_{k \neq \tilde{h}(t)} q_{\tilde{h}(t)k}(\tilde{x}(t), \tilde{u}(t))\Big[V(k, \tilde{x}(t)) - V(\tilde{h}(t), \tilde{x}(t))\Big] \\ &\leq -F(\tilde{h}(t), \tilde{x}(t), \tilde{u}(t)) - \sum_{k \neq \tilde{h}(t)} q_{\tilde{h}(t)k}(\tilde{x}(t), \tilde{u}(t))S_{\tilde{h}(t)k}(\tilde{x}(t)). \end{aligned}$$

Together with (8.3), this inequality implies that

$$\begin{aligned} V(h_0, x_0) \geq E_{\tilde{u}(\cdot)}\Bigg\{ \int_0^\infty e^{-rt}\Big[&F(\tilde{h}(t), \tilde{x}(t), \tilde{u}(t)) \\ &+ \sum_{k \neq \tilde{h}(t)} q_{\tilde{h}(t)k}(\tilde{x}(t), \tilde{u}(t))S_{\tilde{h}(t)k}(\tilde{x}(t))\Big]\, dt \,\Big|\, \tilde{x}(0) = x_0, \tilde{h}(0) = h_0\Bigg\}. \end{aligned}$$

Because of (8.2) we therefore obtain $J(\tilde{u}(\cdot)) \leq V(h_0, x_0)$. If we repeat exactly the same argument with $u(\cdot)$ instead of $\tilde{u}(\cdot)$ then we obtain $J(u(\cdot)) = V(h_0, x_0)$ because $u(t) \in \Phi(h(t), x(t))$ holds for almost all t. Consequently, $J(u(\cdot)) \geq J(\tilde{u}(\cdot))$ which completes the proof. ■

Let us conclude this section with a few remarks on possible generalizations and an important special case of theorem 8.1. First, we have already mentioned that nonautonomous problems, in which one or all of the functions U, f, F, S_{hk}, and q_{hk} depend explicitly on time t, can easily be handled. In such a case, the optimal value function V also depends on t and is a mapping $V : M \times X \times [0, \infty) \mapsto \mathbb{R}$. The HJB equation now becomes

$$\begin{aligned} rV(h, x, t) - V_t(h, x, t) = \max \Bigg\{&F(h, x, u, t) + V_x(h, x, t)f(h, x, u, t) \\ &+ \sum_{k \neq h} q_{hk}(x, u, t)[S_{hk}(x, t) + V(k, x, t) - V(h, x, t)] \,\Big|\, u \in U(h, x, t)\Bigg\}. \end{aligned}$$

The only important difference between this equation and (8.4) is the term $-V_t(h, x, t)$ on the left-hand side.

The second generalization concerns the possible nondifferentiability of the value function V. As in section 3.7 one can prove a version of theorem 8.1, in which $V(h, \cdot)$ is only required to be locally Lipschitz continuous instead of continuously differentiable, by replacing the gradient $V_x(h, x)$ (or $(V_x(h, x, t), V_t(h, x, t))$ for the nonautonomous problem) by a generalized gradient in the sense of definition 3.4. The appropriate HJB equation for this case would be

$$rV(h, x) = \max \Bigg\{ F(h, x, u) + \alpha f(h, x, u)$$

$$+ \sum_{k \neq h} q_{hk}(x, u)[S_{hk}(x) + V(k, x) - V(h, x)] \,\Bigg|\, u \in U(h, x), \alpha \in \partial V(h, x) \Bigg\},$$

where $\partial V(h, x)$ denotes the generalized gradient of the function $V(h, \cdot)$ evaluated at x.

An important special case of the stochastic optimal control problem described in this section is a deterministic optimal control problem with an uncertain time horizon T. In this case, T is the only random variable of the model. We can cast this situation in the above framework simply by specifying two modes, $M = \{1, 2\}$, with the interpretation that $h(t) = 1$ if and only if $T > t$ and $h(t) = 2$ if and only if $T \leq t$. In other words, the mode $h = 1$ is activated as long as control can be exerted and the mode switches to $h = 2$ at the end of the decision period. By setting $f(h, \cdot) = F(h, \cdot) = 0$ for $h = 2$ and $S_{21}(\cdot) = q_{21}(\cdot) = 0$ it follows that the objective functional is given by

$$E_{u(\cdot)} \left\{ \int_0^T e^{-rt} F(1, x(t), u(t))\, dt + e^{-rT} S_{12}(x(T)) \right\},$$

where the probability distribution of the random variable T may depend on the chosen control path. The function $z_{u(\cdot)}(t) = \text{Prob}_{u(\cdot)}\{T > t\}$ is called the survivor function of T and is related to the cumulative distribution function $\pi_{u(\cdot)}(t) = \text{Prob}_{u(\cdot)}\{T \leq t\}$ by the equation $z_{u(\cdot)}(t) + \pi_{u(\cdot)}(t) = 1$. Since the initial mode is $h(0) = 1$ with probability 1, we know that $z_{u(\cdot)}(0) = 1$. Applying these definitions and the rules for calculating conditional probabilities we obtain

$$\begin{aligned}
&\text{Prob}_{u(\cdot)}\{T \in (t, t+\Delta] \mid T > t\} \\
&= [\text{Prob}_{u(\cdot)}\{T \leq t+\Delta\} - \text{Prob}_{u(\cdot)}\{T \leq t\}]/\text{Prob}_{u(\cdot)}\{T > t\} \\
&= [\pi_{u(\cdot)}(t+\Delta) - \pi_{u(\cdot)}(t)]/z_{u(\cdot)}(t) \\
&= \; -[z_{u(\cdot)}(t+\Delta) - z_{u(\cdot)}(t)]/z_{u(\cdot)}(t).
\end{aligned}$$

In the limit, as Δ approaches 0, we obtain from this equation and (8.1) that

$$q_{12}(x(t), u(t), t) = -\dot{z}_{u(\cdot)}(t)/z_{u(\cdot)}(t).$$

Together with the initial condition $z_{u(\cdot)}(0) = 1$ this implies

$$z_{u(\cdot)}(t) = \exp\left\{-\int_0^t q_{12}(x(s), u(s), s)\, ds\right\}. \tag{8.5}$$

This equation relates the switching rate (in this context also called the hazard rate) q_{12} to the survivor function $z_{u(\cdot)}$. Whenever we know the survivor function of the terminal time T (or, equivalently, its cumulative distribution function) we can easily compute the corresponding switching rate by the above formulas, provided the cumulative distribution function is differentiable. To sum up, optimal control problems with uncertain time horizon can be regarded as simple special cases of the piecewise deterministic control problem discussed above.

8.1.2 Stationary Markovian Nash equilibria

Let us consider a piecewise deterministic differential game with N players, restricting ourselves to an autonomous model defined over an infinite time horizon.

Denote by $u^i(t)$ the control value chosen by player $i \in \{1, 2, \ldots, N\}$ at time t and define (cf. section 4.1)

$$u^{-i}(t) = (u^1(t), u^2(t), \ldots, u^{i-1}(t), u^{i+1}(t), \ldots, u^N(t))$$

and

$$u(t) = (u^1(t), u^2(t), \ldots, u^N(t)).$$

In other words, $u(t)$ denotes the vector of all controls and $u^{-i}(t)$ denotes the vector of controls of the opponents of player i.

If the system is in mode $h(t) \in M$ and state $x(t) \in X$ then player i's set of feasible controls is given by $U^i(h(t), x(t), u^{-i}(t)) \subseteq \mathbb{R}^{m^i}$. Since we are restricting ourselves to the autonomous case, this set does not explicitly depend on the time variable t. The state equation for the game is

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

and the switching rates are given by

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\text{Prob}\{h(t+\Delta)=k\mid h(t)=h\}=q_{hk}(x(t),u(t))\qquad k\neq h.$$

The objective functional of player $i\in\{1,2,\ldots,N\}$ is

$$J^i(u^i(\cdot))=E_{u(\cdot)}\Bigg\{\int_0^\infty e^{-r^it}F^i(h(t),x(t),u(t))\,dt$$
$$+\sum_{l\in\mathbb{N}}e^{-r^iT_l}S^i_{h(T_l-)h(T_l)}(x(T_l))\,\Big|\,x(0)=x_0,h(0)=h_0\Bigg\},$$

where $F^i(h(t),x(t),u(t))$ denotes the payoff rate at time t and $S^i_{h(T_l-)h(T_l)}(x(T_l))$ the payoff received if a jump from $h(T_l-)$ to $h(T_l)$ occurs at time $t=T_l$. As in the previous section we assume that all functions are sufficiently smooth. We also require that the payoff functions $F^i(h,\cdot)$ and $S^i_{hk}(\cdot)$ as well as the switching rates $q_{hk}(\cdot)$ are bounded. The sequence T_1, T_2, ... is the sequence of jump times and $r^i>0$ is the discount rate of player i. Finally, we assume that an initial mode $h_0\in M$ and an initial state $x_0\in X$ are given. We denote the piecewise deterministic differential game defined in this way by $\Gamma(h_0,x_0)$.

In this section we are interested in Nash equilibria in which every player represents his strategy in the form $u^i(t)=\phi^i(h(t),x(t))$ where ϕ^i is a mapping from $M\times X$ into $\mathbb{R}^{m^i}$. In other words, the control value chosen by player i at time t depends in general on the current mode $h(t)$ as well as on the current state $x(t)$. We do not consider Nash equilibria consisting of functions ϕ^i which depend also explicitly on time t because we are considering an autonomous problem. As mentioned in section 4.2, this sort of time independence is called stationarity and we refer to a representation of the form $u^i(t)=\phi^i(h(t),x(t))$ as a stationary Markovian strategy. There may, of course, exist nonstationary (i.e., time dependent) equilibria also in autonomous games but they tend to be less interesting and are therefore not often analysed.

If all players $j\neq i$ use the strategies $u^j(t)=\phi^j(h(t),x(t))$ then player i's optimal control problem consists in maximizing

$$J^i_{\phi^{-i}}(u^i(\cdot))=E_{u(\cdot)}\Bigg\{\int_0^\infty e^{-r^it}F^i_{\phi^{-i}}(h(t),x(t),u^i(t))\,dt \tag{8.6}$$
$$+\sum_{l\in\mathbb{N}}e^{-r^iT_l}S^i_{h(T_l-)h(T_l)}(x(T_l))\,\Big|\,x(0)=x_0,h(0)=h_0\Bigg\}$$

subject to the constraints

$$\dot{x}(t) = f^i_{\phi^{-i}}(h(t), x(t), u^i(t))$$
$$x(0) = x_0 \tag{8.7}$$
$$u^i(t) \in U^i_{\phi^{-i}}(h(t), x(t))$$

where the piecewise deterministic process $h(\cdot)$ is determined by the initial condition $h(0) = h_0$ and the switching rates

$$q^i_{\phi^{-i},hk}(x(t), u^i(t)). \tag{8.8}$$

The functions $F^i_{\phi^{-i}}$, $f^i_{\phi^{-i}}$, $U^i_{\phi^{-i}}$, and $q^i_{\phi^{-i},hk}$ are defined by

$$F^i_{\phi^{-i}}(h, x, u^i) = F^i\big(h, x, \phi^1(h, x), \dots, \phi^{i-1}(h, x), u^i, \phi^{i+1}(h, x), \dots, \phi^N(h, x)\big),$$
$$f^i_{\phi^{-i}}(h, x, u^i) = f\big(h, x, \phi^1(h, x), \dots, \phi^{i-1}(h, x), u^i, \phi^{i+1}(h, x), \dots, \phi^N(h, x)\big),$$
$$U^i_{\phi^{-i}}(h, x) = U^i\big(h, x, \phi^1(h, x), \dots, \phi^{i-1}(h, x), \phi^{i+1}(h, x), \dots, \phi^N(h, x)\big),$$
$$q^i_{\phi^{-i},hk}(x, u^i) = q_{hk}\big(x, \phi^1(h, x), \dots, \phi^{i-1}(h, x), u^i, \phi^{i+1}(h, x), \dots, \phi^N(h, x)\big).$$

This optimal control problem is of the same form as the one discussed in section 8.1.1.

The concepts of Nash equilibrium and subgame perfectness are now defined in the following, obvious way.

Definition 8.2 An N-tuple $(\phi^1, \phi^2, \dots, \phi^N)$ of functions $\phi^i : M \times X \mapsto \mathbb{R}^{m^i}$, $i \in \{1, 2, \dots, N\}$, is a stationary Markovian Nash equilibrium of the game $\Gamma(h_0, x_0)$ if, for each player $i \in \{1, 2, \dots, N\}$, an optimal control path $u^i(\cdot)$ of the problem (8.6)–(8.8) exists and is given by the stationary Markovian strategy $u^i(t) = \phi^i(h(t), x(t))$. If $(\phi^1, \phi^2, \dots, \phi^N)$ is a stationary Markovian Nash equilibrium for all games $\Gamma(h, x)$ with $(h, x) \in M \times X$ then it is called subgame perfect.

To verify that a given N-tuple of stationary Markovian strategies $(\phi^1, \phi^2, \dots, \phi^N)$ qualifies as a stationary Markovian Nash equilibrium for a piecewise deterministic game it suffices to verify, for each $i \in \{1, 2, \dots, N\}$, that ϕ^i is an optimal strategy for the problem (8.6)–(8.8). The latter problem is of the form discussed in section 8.1.1 and we can apply theorem 8.1. The following theorem summarizes the resulting equilibrium conditions. Since these conditions do not involve the given initial state and mode they are sufficient conditions for a subgame perfect stationary Markovian Nash equilibrium.

Theorem 8.2 *Let* $(\phi^1, \phi^2, \dots, \phi^N)$ *be a given N-tuple of functions* $\phi^i : M \times X \mapsto \mathbb{R}^{m^i}$, $i \in \{1, 2, \dots, N\}$, *and assume that the piecewise deterministic process defined by the state dynamics*

$$\dot{x}(t) = f(h(t), x(t), \phi^1(h(t), x(t)), \phi^2(h(t), x(t)), \ldots, \phi^N(h(t), x(t)))$$

and the switching rates

$$q_{hk}(x(t), \phi^1(h(t), x(t)), \phi^2(h(t), x(t)), \ldots, \phi^N(h(t), x(t)))$$

is well defined for all initial conditions $(h(0), x(0)) = (h, x) \in M \times X$. *Let there exist* N *bounded functions* $V^i : M \times X \mapsto \mathbb{R}$, $i = 1, 2, \ldots, N$, *such that* $V^i(h, x)$ *is continuously differentiable in* x *and such that the HJB equations*

$$r^i V^i(h, x) = \max \Big\{ F^i_{\phi^{-i}}(h, x, u^i) + V^i_x(h, x) f^i_{\phi^{-i}}(h, x, u^i) \qquad (8.9)$$
$$+ \sum_{k \neq h} q^i_{\phi^{-i},hk}(x, u^i)\big[S^i_{hk}(x) + V^i(k, x) - V^i(h, x)\big] \,\Big|\, u^i \in U^i_{\phi^{-i}}(h, x) \Big\}$$

are satisfied for all $i \in \{1, 2, \ldots, N\}$ *and all* $(h, x) \in M \times X$. *Denote by* $\Phi^i(h, x)$ *the set of all* $u^i \in U^i_{\phi^{-i}}(h, x)$ *which maximize the right-hand side of (8.9). If* $\phi^i(h, x) \in \Phi^i(h, x)$ *holds for all* $i \in \{1, 2, \ldots, N\}$ *and all* $(h, x) \in M \times X$ *then* $(\phi^1, \phi^2, \ldots, \phi^N)$ *is a stationary Markovian Nash equilibrium. Moreover, the equilibrium is subgame perfect.*

Proof To verify the equilibrium property apply theorem 8.1 to (8.6)–(8.8). To see that the equilibrium is subgame perfect just note that the condition $\phi^i(h, x) \in \Phi^i(h, x)$ holds for all $(h, x) \in M \times X$ so that the N-tuple $(\phi^1, \phi^2, \ldots, \phi^N)$ actually constitutes an equilibrium for all initial states $(h, x) \in M \times X$. ■

We now illustrate the derivation of stationary Markovian Nash equilibria in a piecewise deterministic differential game. For another application of theorem 8.2 see section 10.3.

Example 8.1 Let there be two players $i \in \{1, 2\}$ and two modes $h \in \{1, 2\}$, and let the state space be the closed interval $X = [0, \bar{x}]$. Both players have the same sets of feasible controls given by

$$U^i(h, x, u^{-i}) = \begin{cases} [0, 1] & \text{if } x > 0, \\ \{0\} & \text{if } x = 0, \end{cases}$$

and the same discount rate $r^1 = r^2 = r$. Since $U^i(h, x, u^{-i})$ is independent of i, h, and u^{-i} we simply denote it by $U(x)$. The system dynamics are given by

$$f(h, x, u^1, u^2) = -u^1 - u^2$$

and the switching rates are

$$q_{hk}(x, u^1, u^2) = \begin{cases} u^2 + \alpha & \text{if } (h, k) = (1, 2), \\ u^1 + \alpha & \text{if } (h, k) = (2, 1), \end{cases}$$

where α is a positive constant. Finally, the profit rates are

$$F^i(h, x, u^1, u^2) = \begin{cases} x & \text{if } i = h, \\ 0 & \text{if } i \neq h, \end{cases}$$

and

$$S^i_{hk}(x) = \begin{cases} 0 & \text{if } i = h \neq k, \\ \beta & \text{if } i = k \neq h, \end{cases}$$

where β is a positive constant.

A verbal description of this example is as follows. Each player i has a favourable mode $h = i$ and an unfavourable mode $k \neq i$. In the unfavourable mode, the player can influence the switching rate to get into the favourable mode by choosing a high control value. If he succeeds in triggering a switch he gets an immediate reward of $\beta > 0$ at the time of the switch, as well as a continuous profit flow of $x(t)$ as long as the system remains in the favourable mode. Exerting control, however, causes the stock x to decrease so that trying to induce a fast switch into the favourable mode reduces the future payoff rate. The state variable x can be interpreted as a resource and must therefore have a positive value (shadow price). If the player finds that the system is in his favourable mode, he cannot influence the switching rate so as to get into the unfavourable mode, although he certainly can decrease the stock x by choosing a nonzero control.

First, note that the game has a symmetric structure: mode 1 is favourable for player 1 and mode 2 is favourable for player 2. This suggests that we look for stationary Markovian Nash equilibria, (ϕ^1, ϕ^2), which display the same kind of symmetry, i.e.,

$$\phi^1(1, x) = \phi^2(2, x) \text{ and } \phi^2(1, x) = \phi^1(2, x). \tag{8.10}$$

We restrict the analysis to equilibria with this structure, although nothing in the model formulation rules out the existence of asymmetric Nash equilibria. In a symmetric Nash equilibrium the optimal value functions of the two players must exhibit symmetry, i.e.,

$$V^1(1, x) = V^2(2, x) \text{ and } V^2(1, x) = V^1(2, x).$$

Assume that player 2 plays the stationary Markovian strategy $u^2(t) = \phi^2(h(t), x(t))$ and consider the resulting piecewise deterministic control problem of player 1. The HJB equation for this problem is

$$\begin{aligned} rV^1(1,x) &= \max\{x - [u^1 + \phi^2(1,x)]V_x^1(1,x) \\ &\quad + [\phi^2(1,x) + \alpha][V^1(2,x) - V^1(1,x)] \mid u^1 \in U(x)\}, \\ rV^1(2,x) &= \max\{-[u^1 + \phi^2(2,x)]V_x^1(2,x) \\ &\quad + (u^1 + \alpha)[\beta + V^1(1,x) - V^1(2,x)] \mid u^1 \in U(x)\}. \end{aligned}$$

Substituting from (8.10) into these equations, and writing u, $\phi(h,x)$, and $V(h,x)$ instead of u^1, $\phi^1(h,x)$, and $V^1(h,x)$, we obtain

$$\begin{aligned} rV(1,x) &= \max\{x - [u + \phi(2,x)]V_x(1,x) \\ &\quad + [\phi(2,x) + \alpha][V(2,x) - V(1,x)] \mid u \in U(x)\}, \end{aligned} \tag{8.11}$$

$$\begin{aligned} rV(2,x) &= \max\{-[u + \phi(1,x)]V_x(2,x) \\ &\quad + (u + \alpha)[\beta + V(1,x) - V(2,x)] \mid u \in U(x)\}. \end{aligned} \tag{8.12}$$

If we can find continuously differentiable functions $V(1,\cdot)$ and $V(2,\cdot)$ such that (8.11) and (8.12) hold and such that $\phi(1,x)$ maximizes the right-hand side of (8.11) and $\phi(2,x)$ maximizes the right-hand side of (8.12) then we have established that $u^1(t) = \phi^1(h(t),x(t)) = \phi(h(t),x(t))$ is an optimal stationary Markovian strategy for player 1's problem. By symmetry, $u^2(t) = \phi^2(h(t),x(t)) = \phi(3-h(t),x(t))$ is an optimal stationary Markovian strategy of player 2's problem so that (ϕ^1,ϕ^2) is a stationary Markovian Nash equilibrium.[5]

To simplify the exposition we choose the specific parameter values

$$\bar{x} = 1\,,\ r = 1\,,\ \alpha = 2\,,\ \beta = 1/3 \tag{8.13}$$

for which we now derive three different stationary Markovian Nash equilibria. All of them have the property that $\phi(1,x) = 0$ holds for all $x \in [0,1]$. To see this, note that a high value of x is beneficial to both players because x enters positively in their utility functions, and because it represents the common stock from which they can extract resources to influence the switching rates. The shadow price interpretation of V_x suggests that $V_x(h,x) > 0$. In particular, we expect

$$V_x(1,x) \geq 0. \tag{8.14}$$

If this is true for all $x \in [0,1]$ then the maximum on the right-hand side of (8.11) is attained at the lower boundary of the control set $U(x)$, that is, $\phi(1,x) = 0$ for all $x \in [0,1]$. Of course, we need to check that (8.14) holds for any equilibrium candidate.

Using our conjecture $\phi(1,x) = 0$ and the parameter specifications in (8.13) we can rewrite (8.11) and (8.12) as follows

[5] We recommend that the reader makes sure that these arguments and the symmetric structure of the game have been understood before proceeding to the derivation of ϕ and V.

$$V(1,x) = x - \phi(2,x)V_x(1,x) + [\phi(2,x)+2][V(2,x) - V(1,x)], \tag{8.15}$$

$$V(2,x) = \max\{-uV_x(2,x) + (u+2)[1/3 + V(1,x) - V(2,x)] \mid u \in U(x)\}. \tag{8.16}$$

The coefficient of u on the right-hand side of (8.16) is

$$\sigma(x) := 1/3 + V(1,x) - V(2,x) - V_x(2,x) \tag{8.17}$$

and the maximization with respect to u yields the following necessary conditions

$$u = 0 \Rightarrow \sigma(x) \leq 0,$$
$$u = 1 \Rightarrow \sigma(x) \geq 0,$$
$$u \in (0,1) \Rightarrow \sigma(x) = 0.$$

We now show that, for each of these three cases, one can construct a stationary Markovian Nash equilibrium.

Case 1: Suppose that $\sigma(x) \leq 0$ for all $x \in [0,1]$. In this case we choose $u = \phi(2,x) = 0$ for all $x \in [0,1]$ and equations (8.15) and (8.16) can be written as

$$V(1,x) = x + 2[V(2,x) - V(1,x)],$$
$$V(2,x) = 2[1/3 + V(1,x) - V(2,x)].$$

The unique solution to this system of two linear equations is given by $V(1,x) = (4/15) + (3/5)x$ and $V(2,x) = (2/5) + (2/5)x$. We therefore obtain from (8.17) that $\sigma(x) = (x-1)/5 \leq 0$, which verifies the characterizing assumption of this case. Finally, it is clear that (8.14) holds so that indeed all the conditions of theorem 8.2 are satisfied. We have proved that (ϕ^1, ϕ^2) with $\phi^1(h,x) = \phi^2(h,x) = 0$ for $h \in \{1,2\}$ and all $x \in [0,1]$ constitutes a subgame perfect, stationary Markovian Nash equilibrium.

Case 2: Suppose that $\sigma(x) \geq 0$ for all $x \in [0,1]$. In this case we choose $u = \phi(2,x) = 0$ for $x = 0$ and $u = \phi(2,x) = 1$ for all $x \in (0,1]$. Substituting this into (8.15) and (8.16) we obtain for $x = 0$

$$V(1,0) = 2[V(2,0) - V(1,0)],$$
$$V(2,0) = 2[1/3 + V(1,0) - V(2,0)],$$

and for $x > 0$

$$\left.\begin{aligned} V(1,x) &= x - V_x(1,x) + 3[V(2,x) - V(1,x)], \\ V(2,x) &= -V_x(2,x) + 3[1/3 + V(1,x) - V(2,x)]. \end{aligned}\right. \tag{8.18}$$

The first pair of equations yields $V(1,0) = 4/15$ and $V(2,0) = 2/5$. Using these values as initial conditions for the system of two linear, first-order differential equations in (8.18) we obtain the unique solution

$$V(1,x) = \frac{1}{3}e^{-x} + \frac{11}{735}e^{-7x} + \frac{4}{7}x - \frac{4}{49},$$
$$V(2,x) = \frac{1}{3}e^{-x} - \frac{11}{735}e^{-7x} + \frac{3}{7}x + \frac{4}{49}.$$

This leads to

$$\sigma(x) = \frac{1}{3}e^{-x} - \frac{11}{147}e^{-7x} + \frac{1}{7}x - \frac{38}{147}.$$

It is a straightforward exercise in calculus to show that $\sigma(0) = 0$ and $\sigma(x) > 0$ for all $x \in (0, 1]$ so that the qualification for this case is satisfied. Similarly, it can be verified that (8.14) holds, so that we have established all the conditions ensuring the Nash equilibrium property. We conclude that (ϕ^1, ϕ^2) defined by

$$\phi^1(h,x) = \begin{cases} 0 \text{ if } h=1 \text{ or } x=0, \\ 1 \text{ if } h=2 \text{ and } x \in (0,1], \end{cases} \qquad \phi^2(h,x) = \begin{cases} 0 \text{ if } h=2 \text{ or } x=0, \\ 1 \text{ if } h=1 \text{ and } x \in (0,1], \end{cases}$$

is a subgame perfect, stationary Markovian Nash equilibrium.

Case 3: Suppose that $\sigma(x) = 0$ for all $x \in [0, 1]$. In the former cases we found equilibria with boundary solutions for the optimal control u. Now we are going to construct an equilibrium admitting an interior solution. First of all, if $x = 0$ the only feasible value for u is $u = 0$ so we obtain $V(1,0) = 4/15$ and $V(2,0) = 2/5$ by the same argument as in case 2. From (8.17) one can see that the assumption $\sigma(x) = 0$ for all $x \in [0, 1]$ is equivalent to

$$V_x(2,x) = 1/3 + V(1,x) - V(2,x).$$

Moreover, by substituting this into (8.16), the terms involving u cancel out (as they should) and we obtain

$$V(2,x) = 2[1/3 + V(1,x) - V(2,x)].$$

The last two equations, together with the initial conditions $V(1,0) = 4/15$ and $V(2,0) = 2/5$, have a unique solution, given by

$$V(1,x) = -\frac{1}{3} + \frac{3}{5}e^{x/2} \quad \text{and} \quad V(2,x) = \frac{2}{5}e^{x/2}.$$

Note that $\sigma(x) = 0$ holds by construction and that (8.14) is also true. We still have to determine the function $\phi(2,x)$. To this end, substitute the

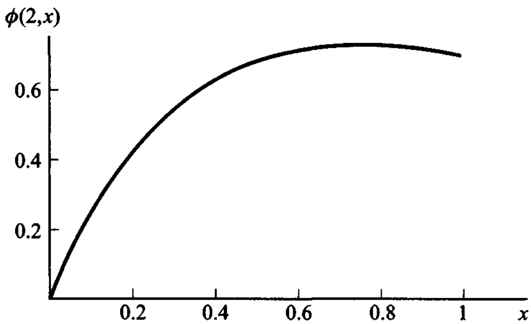

Figure 8.1 The graph of the function $\phi(2, \cdot)$ defined in (8.19)

solutions $V(1, x)$ and $V(2, x)$ into (8.15) and solve the resulting equation for $\phi(2, x)$. This yields

$$\phi(2, x) = \frac{6(1 + x - e^{x/2})}{3e^{x/2} - 2}. \tag{8.19}$$

This function is depicted in figure 8.1. It is clearly seen (and easy to prove) that $\phi(2, 0) = 0$ and that $\phi(2, x) \in (0, 1)$ for all $x \in (0, 1)$. Therefore, we have obtained a third subgame perfect, stationary Markovian Nash equilibrium. It is determined by equation (8.10), $\phi^1(1, x) = 0$, and $\phi^1(2, x) = \phi(2, x)$, where the function ϕ is defined in (8.19). An interesting feature of this equilibrium is that the function $\phi(2, \cdot)$ is strictly decreasing for values x close to 1. This seems to be somewhat counterintuitive because one would expect high stock values x to encourage a lot of control effort, to induce a mode-jump as soon as possible. The decline of $\phi(2, x)$ for large x can be understood, however, by noting that $\sigma(x) = 0$ for all x. This means that both agents are indifferent between their feasible actions in the unfavourable mode.

8.1.3 Piecewise open-loop Nash equilibria

In the previous section we discussed stationary Markovian Nash equilibria for piecewise deterministic differential games. In such an equilibrium each player uses a strategy of the form $u^i(t, \xi) = \phi^i(h(t, \xi), x(t, \xi))$. Of course, there are other reasonable strategies for optimal control paths and in the present section we deal with one of them.

To begin with let us briefly point out that an open-loop strategy of the form $u^i(t, \xi) = \phi^i(t)$ is usually not appropriate for a stochastic control model or a stochastic differential game. A strategy represented in this way does not take into account the information that is revealed during the planning period. When the feasible control sets depend on a stochastic process, it can be impossible to represent a control path in open-loop form. For example, if $U^i(h, x, u^{-i})$ and $U^i(k, x, u^{-i})$ are disjoint sets for two modes $h \neq k$, both of which occur with positive probability, then it is impossible to find any function $\phi^i : [0, \infty) \mapsto \mathbb{R}^{m^i}$ such that $\phi(t) \in U^i(h(t), x(t), u^{-i}(t))$ holds for all realizations of the piecewise deterministic process. In addition to the possible nonexistence of feasible open-loop strategies, it is in most cases unrealistic to assume that the decision makers in the model cannot condition their actions on the realizations of the random events. Therefore, open-loop strategies are not relevant in stochastic control models or stochastic differential games.

A better way to capture the idea of open-loop information in a piecewise deterministic setting is to assume that the agents can at least condition their actions on the mode of the system. This leads to so-called piecewise open-loop strategies. To define this concept let us denote by $\ell(t)$ the last jump time of the mode process occurring at or before time t. If there has not been any jump at or before t we set $\ell(t) = 0$. Formally, $\ell(t) = \max\{T_l \mid T_l \leq t, l \in \mathbb{N}\}$, where the maximum of the empty set is taken as 0. A strategy of the form[6]

$$u^j(t) = \phi^j(h(\ell(t)), x(\ell(t)), t - \ell(t)) \tag{8.20}$$

depends on the mode and the state at the last jump time as well as on the time elapsed since the last jump and is called a piecewise open-loop strategy for the control path $u^j(\cdot)$. Note that it does not make any difference if we replace $h(\ell(t))$ in (8.20) by $h(t)$ since $h(\cdot)$ is a right-continuous process which is constant on the interval $[\ell(t), t]$. In other words, $h(t) = h(\ell(t))$ holds for all t by the very definition of $\ell(t)$ and $h(\cdot)$. On the other hand, one cannot replace $x(\ell(t))$ by $x(t)$ since, in general, the state is not constant on $[t, \ell(t)]$.

The reason why piecewise open-loop representations are useful for piecewise deterministic control problems or differential games is that knowledge of $h(\ell(t))$, $x(\ell(t))$, and the vector of controls $u(s)$ for $s \geq \ell(t)$ is sufficient to compute the state at time t and all profit rates between the last mode switch and the present time t. This is so because no uncertainty is revealed between two successive jumps.

[6]Note that all the arguments on the right-hand side of (8.20) are random variables. A rigorous statement of (8.20) would therefore be $u^j(t, \xi) = \phi^j(h(\ell(t, \xi), \xi), x(\ell(t, \xi), \ell), t-\ell(t, \xi))$.

Now assume that all players $j \neq i$ represent their control paths in the form (8.20), where $\phi^j : M \times X \times [0, \infty)$ are given functions. To simplify the notation we make the following definitions:

$$F^i_{\phi^{-i}}(h, x, u^i; y, s)$$
$$= F^i(h, x, \phi^1(h, y, s), \ldots, \phi^{i-1}(h, y, s), u^i, \phi^{i+1}(h, y, s), \ldots, \phi^N(h, y, s)),$$
$$f^i_{\phi^{-i}}(h, x, u^i; y, s)$$
$$= f(h, x, \phi^1(h, y, s), \ldots, \phi^{i-1}(h, y, s), u^i, \phi^{i+1}(h, y, s), \ldots, \phi^N(h, y, s)),$$
$$U^i_{\phi^{-i}}(h, x; y, s)$$
$$= U^i(h, x, \phi^1(h, y, s), \ldots, \phi^{i-1}(h, y, s), \phi^{i+1}(h, y, s), \ldots, \phi^N(h, y, s)),$$
$$q^i_{\phi^{-i},hk}(x, u^i; y, s)$$
$$= q_{hk}(x, \phi^1(h, y, s), \ldots, \phi^{i-1}(h, y, s), u^i, \phi^{i+1}(h, y, s), \ldots, \phi^N(h, y, s)).$$

Given the fixed strategies (8.20) for the players $j \neq i$ we can write player i's objective functional as follows

$$J^i_{\phi^{-i}}(u^i(\cdot)) = E_{u(\cdot)}\left\{ \int_0^\infty e^{-r^i t} F^i_{\phi^{-i}}(h(t), x(t), u^i(t); x(\ell(t)), t - \ell(t))\, dt \right.$$
$$\left. + \sum_{l \in \mathbb{N}} e^{-r^i T_l} S^i_{h(T_l-)h(T_l)}(x(T_l)) \,\middle|\, x(0) = x_0, h(0) = h_0 \right\}. \tag{8.21}$$

The constraints are

$$\dot{x}(t) = f^i_{\phi^{-i}}(h(t), x(t), u^i(t); x(\ell(t)), t - \ell(t)),$$
$$u^i(t) \in U^i_{\phi^{-i}}(h(t), x(t); x(\ell(t)), t - \ell(t)) \tag{8.22}$$

and the switching rates are determined by

$$q^i_{\phi^{-i},hk}(x(t), u^i(t); x(\ell(t)), t - \ell(t)). \tag{8.23}$$

The initial conditions are $x(0) = x_0$ and $h(0) = h_0$. Note that this optimal control model is not of the form discussed in section 8.1.1 because of the appearance of the arguments $x(\ell(t))$ and $t - \ell(t)$ in the various functions. However, it can be shown that under appropriate smoothness and boundedness assumptions the piecewise deterministic process $(h(\cdot), x(\cdot))$ is still well defined. We assume for the rest of the section that these assumptions are satisfied; see section 8.3 for references to rigorous treatments of this subject.

Definition 8.3 The N-tuple $(\phi^1, \phi^2, \ldots, \phi^N)$ of functions $\phi^i : M \times X \times [0, \infty) \mapsto \mathbb{R}^{m^i}$, $i \in \{1, 2, \ldots, N\}$, is called a piecewise open-loop Nash equilibrium if, for each player $i \in \{1, 2, \ldots, N\}$, an optimal control path $u^i(\cdot)$ of the problem (8.21)–(8.23) exists and is given by the piecewise open-loop strategy $u^i(t) = \phi^i(h(\ell(t)), x(\ell(t)), t - \ell(t))$, where $\ell(t)$ denotes the last jump time at or before t.

It is clear that piecewise open-loop Nash equilibria are not Markovian equilibria since the controls $u^i(t)$ depend on the value of the state at the last jump time, $x(\ell(t))$. For the same reason it is not sufficient to define a subgame by the current state $x(t)$ and the mode $h(t)$ as in definition 4.4. We also have to specify how the state $x(t)$ has been reached. This requires that we include $y = x(\ell(t))$ and $s = \ell(t) - t$ in the description of subgames. An appropriate family of subgames would be $\{\Gamma(h, x; y, s) \mid h \in M, x \in X, y \in X, s \in [0, \infty)\}$. It turns out, however, that piecewise open-loop Nash equilibria are not subgame perfect under this information structure.[7]

Because player i's opponents represent their choices by piecewise open-loop strategies, player i's optimal control problem is not autonomous. This makes the verification of piecewise open-loop Nash equilibria more cumbersome than the verification of stationary Markovian equilibria because (i) one has to use a HJB equation that is valid for non-autonomous problems and (ii) the utility functions $F^i_{\phi^{-i}}$, the state dynamics $f^i_{\phi^{-i}}$, the switching rates $q^i_{\phi^{-i},hk}$, and the control sets $U^i_{\phi^{-i}}$ of the optimal control problem (8.21)–(8.23) depend on $x(\ell(t))$ and $t - \ell(t)$. Having mentioned these difficulties we can state the main result of the section.

Theorem 8.3 *Let $(\phi^1, \phi^2, \ldots, \phi^N)$ be a given N-tuple of functions $\phi^i : M \times X \times [0, \infty) \mapsto \mathbb{R}^{m^i}$, $i \in \{1, 2, \ldots, N\}$, and assume that for every pair $(h, y) \in M \times X$ there exists a unique, absolutely continuous solution $x_{hy} : [0, \infty) \mapsto X$ of the initial value problem*

$$\dot{x}_{hy}(s) = f(h, x_{hy}(s), \phi^1(h, y, s), \phi^2(h, y, s), \ldots, \phi^N(h, y, s)), \; x_{hy}(0) = y.$$

Let there exist bounded functions $V^i : M \times X \times X \times [0, \infty) \mapsto \mathbb{R}$, $i = 1, 2, \ldots, N$, such that $V^i(h, x, y, t)$ is continuously differentiable with respect to x and t, and such that the HJB equations

$$\begin{aligned} & r^i V^i(h, x, y, t) - V^i_t(h, x, y, t) \\ &= \max \Big\{ F^i_{\phi^{-i}}(h, x, u^i; y, t) + V^i_x(h, x, y, t) f^i_{\phi^{-i}}(h, x, u^i; y, t) \\ &+ \sum_{k \neq h} q^i_{\phi^{-i},hk}(x, u^i; y, t)[S^i_{hk}(x) + V^i(k, x, x, 0) - V^i(h, x, y, t)] \Big| u^i \in U^i_{\phi^{-i}}(h, x; y, t) \Big\} \end{aligned} \tag{8.24}$$

[7]If we consider the coarser subgame structure $\{\Gamma(h, x; x, 0) | h \in M, x \in X\}$, then piecewise open-loop Nash equilibria turn out to be subgame perfect. This notion of perfectness is, however, too weak to be of any interest.

are satisfied for all $i \in \{1, 2, \ldots, N\}$ *and all* $(h, x, y, t) \in M \times X \times X \times [0, \infty)$. *Denote by* $\Phi^i(h, x, y, t)$ *the set of all* $u^i \in U^i_{\phi^{-i}}(h, x; y, t)$ *which maximize the right-hand side of (8.24). If* $\phi^i(h, y, t) \in \Phi^i(h, x_{hy}(t), y, t)$ *holds for all* $i \in \{1, 2, \ldots, N\}$ *and all* $(h, y, t) \in M \times X \times [0, \infty)$, *then* $(\phi^1, \phi^2, \ldots, \phi^N)$ *is a piecewise open-loop Nash equilibrium.*

We do not prove this theorem but only give the intuition. To this end, recall that the optimal value function is the obvious candidate for the function V^i, provided it is sufficiently smooth. But why does the optimal value function for the present model have four arguments, h, x, y, and t? The variables h and x are, as before, the current mode and the current state of the system. The arguments y and t are the state at the last jump and the time elapsed since the last jump. They have to be included in the optimal value function because the opponents' controls are functions of these variables. Formally, one can think of $V^i(h, x, y, s)$ as the maximal value of

$$E_{u(\cdot)}\left\{\int_s^\infty e^{-r^i(t-s)} F^i_{\phi^{-i}}(h(t), x(t), u^i(t); y, t)\, dt \right. \tag{8.25}$$

$$\left. + \sum_{l \in \mathbb{N}} e^{-r^i(T_l - s)} S^i_{h(T_l-)h(T_l)}(x(T_l)) \,\middle|\, x(s) = x, h(s) = h, T_1 > s \right\}$$

subject to the constraints (8.22) and (8.23) with $\ell(s) = 0$ and $x(0) = y$. In particular, it follows that $V^i(h_0, x_0, x_0, 0)$ denotes the optimal value that can be achieved in (8.21)–(8.23). The principle of embedding is at work again, a principle which was already identified as one of the building blocks of the dynamic programming approach to optimization in section 3.2.

Now compare the HJB equation, (8.24), for a piecewise open-loop Nash equilibrium with its counterpart for a stationary Markovian Nash equilibrium, (8.9). The additional term on the left-hand side of (8.24), $-V_t(h, x, y, t)$, simply reflects the fact that player i's optimal control problem is nonautonomous because his opponents use time dependent strategies (cf. the remark immediately after theorem 8.1). To grasp the intuition of the HJB equation it is, however, more important to look at its right-hand side. The maximand can be interpreted as the total rate at which value is increased at time t when the control value u^i is chosen. In equation (8.9) this total value rate consists of the immediate profit rate, $F^i_{\phi^{-i}}(h, x, u^i)$, plus the rate of change of the state, $f^i_{\phi^{-i}}(h, x, u^i)$, times the shadow price of the state, $V^i_x(h, x)$, and plus the term $\sum_{k \neq h} q^i_{\phi^{-i}, hk}(x, u^i)[S^i_{hk}(x) + V^i(k, x) - V^i(h, x)]$. The last term takes account of possible switches of the system mode. As a matter of fact, a switch from h to k yields an immediate value gain $S^i_{hk}(x)$ but it also yields the gain (or loss, if it is negative) $V^i(k,x) - V^i(h,x)$ because the system stops being in mode h and starts being in mode k. This gain due to switching is multiplied by $q^i_{\phi^{-i},hk}(x,u^i)$ because this is the rate at which the switching takes place. In the HJB equation (8.24) the interpretation is as before, with the exception that we now have the term $S^i_{hk}(x) + V^i(k,x,x,0) - V^i(h,x,y,t)$ representing the gain due to switching. The part $S^i_{hk}(x)$ is again the immediate reward that is achieved by a switch from h to k. The expression $V^i(k,x,x,0) - V^i(h,x,y,t)$ takes into account that, if such a switch occurs, there is not only a change of the mode but also the opponents' strategies change from $\phi^j(h,y,t)$ to $\phi^j(k,x,0)$.

Example 8.1 (continued) The first two stationary Markovian Nash equilibria derived in this example (cases 1 and 2) are degenerate in the sense that they are constant with respect to x. Consequently, the stationary Markovian strategies representing the players' optimal control paths are also piecewise open-loop strategies and the stationary Markovian Nash equilibria from cases 1 and 2 are piecewise open-loop Nash equilibria as well. The Nash equilibrium constructed in case 3 is not given by a piecewise open-loop strategy. However, one can show that there exists a piecewise open-loop Nash equilibrium which is equivalent to the stationary Markovian equilibrium, in the sense that both equilibria generate exactly the same control paths and exactly the same stochastic process. This follows from the fact that player 2's optimal choice in mode 2 is equal to $u^2(t) = 0$ independently of the control value of player 1 and that $u^2(t) = 0$ is already a piecewise open-loop strategy. We shall encounter a similar situation in exercise 2 at the end of the chapter.

The following example shows that stationary Markovian Nash equilibria and piecewise open-loop Nash equilibria lead, in general, to different control paths. The example is one in which the only random variable is the terminal time of the game.

Example 8.2 Let there be two players $i \in \{1,2\}$, two modes $h \in \{1,2\}$, and let the state space be the closed interval $X = [-\bar{x}, \bar{x}]$ for some sufficiently large number $\bar{x}$. Both players have the same sets of feasible controls, given by

$$U^i(h,x,u^{-i}) = \begin{cases} [0,\infty) & \text{if } x = -\bar{x}, \\ \mathbb{R} & \text{if } x \in (-\bar{x},\bar{x}), \\ (-\infty,0] & \text{if } x = \bar{x}, \end{cases}$$

and the same discount rate $r^1 = r^2 = 1/4$. The system dynamics are given by

$$f(h, x, u^1, u^2) = \begin{cases} u^1 + u^2 & \text{if } h = 1, \\ 0 & \text{if } h = 2, \end{cases}$$

and the switching rates are constant and given by

$$q_{12}(x, u^1, u^2) = 1/4 \quad \text{and} \quad q_{21}(x, u^1, u^2) = 0.$$

The profit rates are

$$F^i(h, x, u^1, u^2) = \begin{cases} -(u^i)^2/2 & \text{if } h = 1, \\ 0 & \text{if } h = 2, \end{cases}$$

and

$$S^i_{12}(x) = -3x^2 \quad \text{and} \quad S^i_{21} = 0.$$

Finally, the initial mode is assumed to be $h = 1$.

Let us first try to understand the goals of the players in this game. Since the switching rate q_{21} is equal to 0, the system will remain in mode 2 from the first switch onwards. In mode 2, the system dynamics and the profit rate both equal 0 so that nothing can be gained after the first switch. Essentially, the game ends at the first switching time T_1. The endpoint T_1, however, is a random variable with constant hazard rate $q_{12} = 1/4$. From the remarks made at the end of section 8.1.1 it follows that this is equivalent to saying that T_1 is exponentially distributed with mean $1/q_{12} = 4$. Before the switch from mode 1 to mode 2 occurs, the players have quadratic control costs $(u^i)^2/2$. If they were to act myopically, they would like to exert no control at all, which would imply that the system state remains at its initial value x_0. At the endpoint, however, there is a quadratic cost $3x(T_1)^2$, which becomes larger the further away the state is from 0. Because the players do not know the realization of T_1, but only its distribution, we would expect them to move the state slowly towards 0 in order to avoid high costs at the end of the game.

We wish to find out which information structure causes the movement to $x = 0$ to be faster, the one described by stationary Markovian strategies or the one described by piecewise open-loop strategies. To this end, first note that the game is completely symmetric with respect to the players. We therefore restrict the analysis to Nash equilibria (ϕ^1, ϕ^2) such that $\phi^1 = \phi^2$. In such a symmetric Nash equilibrium the optimal value functions of the two players are identical, i.e., $V^1 = V^2$. We simplify the notation by writing ϕ and V instead of ϕ^i and V^i, respectively.

Let us begin by computing a stationary Markovian Nash equilibrium of the game. The HJB equation from theorem 8.2 is given by

$$\begin{aligned} V(1, x)/4 = \max\{&-u^2/2 + V_x(1, x)[u + \phi(1, x)] \\ &+ [-3x^2 + V(2, x) - V(1, x)]/4 \mid u \in \mathbb{R}\} \end{aligned}$$

for mode 1 and $V(2, x) = 0$ for mode 2. Carrying out the maximization in the equation for mode 1 yields

$$u = \phi(1, x) = V_x(1, x).$$

Substituting this into the right-hand side of the HJB equation, and using $V(2, x) = 0$, we get

$$V(1, x)/4 = (3/2)V_x(1, x)^2 - (3/4)x^2 - V(1, x)/4. \tag{8.26}$$

Because of the quadratic costs, the linear system dynamics, and the constant switching rates, the model has a linear quadratic structure which leads us to guess a value function of the form $V(1, x) = A + Bx + Cx^2$. Substituting this into (8.26) and collecting terms with equal powers of x, we see that the coefficients A and B must be equal to 0 and that C must satisfy the quadratic equation $6C^2 - C/2 - 3/4 = 0$. This equation has a unique negative root $C = (1 - \sqrt{73})/24$. The conditions of theorem 8.2 are therefore satisfied with the functions $V^1(1, x) = V^2(1, x) = Cx^2$ and $V^1(2, x) = V^2(2, x) = 0$. We conclude that the strategies ϕ^1 and ϕ^2 defined by

$$\phi^1(1, x) = \phi^2(1, x) = V_x(1, x) = 2Cx = (1 - \sqrt{73})x/12$$

constitute a stationary Markovian Nash equilibrium. The state trajectory generated by this equilibrium is given by $x(t) = x_0e^{4Ct}$ for all $t \in [0, T_1]$.

Next let us derive a piecewise open-loop Nash equilibrium of the game. The HJB equations from theorem 8.3 are given by

$$\begin{aligned}
&V(1, x, y, t)/4 - V_t(1, x, y, t) \\
&= \max\{-u^2/2 + V_x(1, x, y, t)[u + \phi(1, y, t)] \\
&+ \big[-3x^2 + V(2, x, x, 0) - V(1, x, y, t)\big]/4 \mid u \in \mathbb{R}\}
\end{aligned}$$

and

$$V(2, x, y, t) - V_t(2, x, y, t) = 0.$$

We may choose $V(2, x, y, t) = 0$ to satisfy the second equation. Maximization of the right-hand side of the first equation yields $u = V_x(1, x, y, t)$. Substituting this into the HJB equation for mode 1 we obtain

$$\begin{aligned}
&V(1, x, y, t)/2 - V_t(1, x, y, t) \\
&= V_x(1, x, y, t)^2/2 + V_x(1, x, y, t)\phi(1, y, t) - (3/4)x^2.
\end{aligned} \tag{8.27}$$

Guessing a solution to this equation is harder than for the corresponding equation in the stationary Markovian equilibrium, but can be done in the following way. Using the interpretation of $V(1, x, y, t)$ given in (8.25), and the fact that the endpoint T_1 is exponentially distributed, we see that (ψ^1, ψ^2) with $\psi^1(t) = \psi^2(t) = \phi(1, y, t)$ must be an open-loop Nash equilibrium of the deterministic and symmetric differential game in which the objective functional of player $i \in \{1, 2\}$ is given by

$$\int_0^\infty e^{-t/2}\left\{-(1/2)[u^i(t)]^2 - (3/4)x(t)^2\right\} dt$$

and the system dynamics are

$$\dot{x}(t) = u^1(t) + u^2(t), \quad x(0) = y.$$

This is a linear quadratic open-loop game. Applying the methods from section 7.1 we obtain $\psi^i(t) = \phi(1, y, t) = -(3/4)ye^{-3t/2}$. Using (8.25) once again we see that $V(1, x, y, t)$ must be the optimal value function of the following deterministic optimal control model on the time interval $[t, \infty)$:

$$\text{Maximize } \int_t^\infty e^{-(s-t)/2}\left[-(1/2)u(s)^2 - (3/4)x(s)^2\right] ds$$
$$\text{subject to } \dot{x}(s) = u(s) - (3/4)ye^{-3s/2}, \quad x(t) = x.$$

One can solve this problem by using the maximum principle to obtain

$$u(s) = -(x - ye^{-3t/2})e^{t-s} - (3/4)ye^{-3s/2}$$

as the optimal solution. Computing the corresponding state trajectory and evaluating the objective functional along this solution finally yields

$$V(1, x, y, t) = -\frac{x^2}{2} + \frac{xy}{4}e^{-3t/2} - \frac{5y^2}{112}e^{-3t}. \tag{8.28}$$

This concludes the informed guessing of a solution to equation (8.27). It is straightforward to verify that the guessing led to a correct solution: simply substitute $V(1, x, y, t)$ from (8.28) and $\phi(1, y, t) = -(3/4)ye^{-3t/2}$ into (8.27). To complete the analysis of the piecewise open-loop Nash equilibrium we compute the state trajectory $x_{1y}(t)$ mentioned in theorem 8.3. This trajectory is the solution to the initial value problem

$$\dot{x}_{1y}(t) = 2\phi(1, y, t) = -(3/2)ye^{-3t/2}, \quad x(0) = y,$$

which is given by $x_{1y}(t) = ye^{-3t/2}$. Since the maximum on the right-hand side of the HJB equation for mode 1 is attained at the unique point $u = V_x(1, x, y, t)$, we have to verify $\phi(1, y, t) = V_x(1, x_{1y}(t), y, t)$. But this condition is indeed satisfied, as can easily be checked by substitution. We conclude from theorem 8.3 that (ϕ^1, ϕ^2), with $\phi^i(1, y, t) = -(3/4)ye^{-3t/2}$, $i = 1, 2$, is a piecewise open-loop Nash equilibrium.

Let us finally compare the stationary Markovian Nash equilibrium with the piecewise open-loop Nash equilibrium. The state trajectory for the former is given by $x(t) = x_0 e^{4Ct}$, with $C = (1 - \sqrt{73})/24$, while the state trajectory generated in the latter is $x(t) = x_0 e^{-3t/2}$. Because $-3/2 < 4C < 0$ we see that the convergence of the state to 0 happens faster in the piecewise open-loop Nash equilibrium than in the stationary Markovian Nash equilibrium.

8.2 Differential games with white noise

8.2.1 A control problem with white noise

We now analyse stochastic optimal control models and differential games in which the uncertainty enters, not in the form of a piecewise deterministic process, but in the form of a Wiener process. Wiener processes are also known as Brownian motion or white noise processes and play an important role in many different fields ranging from physics to economics (e.g., finance).

A standardized k-dimensional Wiener process w with time domain $[0, T\rangle$[8] is a continuous-time stochastic process with values in $\mathbb{R}^k$, that is, $w : [0, T\rangle \times \Xi \mapsto \mathbb{R}^k$. Its defining properties are:[9]

(i) $w(0, \xi) = w_0$ for all ξ in a set of probability 1 where $w_0 \in \mathbb{R}^k$ is an arbitrary initial value;

(ii) for any finite sequence of real numbers $(t_1, t_2, \ldots, t_l)$ with $0 \leq t_1 < t_2 < \ldots < t_l \leq T$ it holds that the random variables $w(t_1, \cdot)$ and $w(t_{i+1}, \cdot) - w(t_i, \cdot)$, $i \in \{1, 2, \ldots, l-1\}$, are stochastically independent;

(iii) for all pairs (s, t) of real numbers such that $0 \leq s < t \leq T$, the random variable $w(t, \cdot) - w(s, \cdot)$ has a normal distribution with mean vector $0 \in \mathbb{R}^k$ and covariance matrix $(t - s)I$, where $I \in \mathbb{R}^{k \times k}$ denotes the $k \times k$ unit matrix.

It will turn out that for our purpose the initial value w_0, mentioned in condition (i), has no importance whatsoever and we choose $w_0 = 0 \in \mathbb{R}^k$ for the rest of the chapter. As in the previous sections we simplify the notation by dropping the argument ξ whenever it is not needed. In particular, the realization of the Wiener process at time t will be denoted by $w(t)$; a similar notation is used for the state and control paths.

[8]We use the same notation as in chapter 4, namely, $[0, T\rangle = [0, T]$ if T is a finite number and $[0, T\rangle = [0, \infty)$ if $T = \infty$.

[9]The definition of stochastically independent random variables and of the normal distribution can be found in any text on probability theory.

In the stochastic control models and differential games to be considered here it is assumed that the evolution of the state variable x can be described by a stochastic differential equation of the form

$$dx(t) = f(x(t), u(t), t)\,dt + \sigma(x(t), u(t), t)\,dw(t), \quad x(0) = x_0. \tag{8.29}$$

As before, feasible control values at time t have to be chosen from the set $U(x(t), t)$. Thus, f is a function defined on $\Omega = \{(x, u, t) \mid x \in X, u \in U(x, t), t \in [0, T)\}$ with values in $\mathbb{R}^n$. The function σ is also defined on Ω and takes values in $\mathbb{R}^{n \times k}$. In other words, $\sigma(x, u, t)$ is an $n \times k$ matrix for every $(x, u, t) \in \Omega$. A component $\sigma_{ij}(x, u, t)$ of this matrix measures the direct influence of the jth component of the k-dimensional Wiener process on the evolution of the ith component of the n-dimensional state vector. In many examples in economics and management both k and n are equal to 1 and the function σ is a real-valued function. Note that if $\sigma(x, u, t)$ is identically equal to 0, equation (8.29) can be divided by dt to obtain the deterministic state dynamics $\dot{x}(t) = f(x(t), u(t), t)$ which we encountered in chapter 3. The whole impact of uncertainty is reflected by the second term on the right-hand side of (8.29).

The precise meaning of the statement that $x(\cdot)$ is a solution to the stochastic differential equation (8.29) is that $x(\cdot)$ satisfies the integral equation

$$x(t) = x_0 + \int_0^t f(x(s), u(s), s)\,ds + \int_0^t \sigma(x(s), u(s), s)\,dw(s)$$

for all ξ in a set of probability 1. The first integral on the right-hand side is the usual Riemann integral while the second integral has to be interpreted as the limit

$$\lim_{\delta \to 0} \sum_{l=1}^{L-1} \sigma(x(t_l), u(t_l), t_l)[w(t_{l+1}) - w(t_l)]$$

where $0 = t_1 < t_2 < \ldots < t_L = t$ and $\delta = \max\{|t_{l+1} - t_l| \mid 1 \leq l \leq L - 1\}$. It follows from this definition, and from the properties of a Wiener process, that the solution of a stochastic differential equation does not depend on the initial value of the Wiener process. This shows that it incurs no loss of generality to assume that this initial value is equal to $0 \in \mathbb{R}^k$, as mentioned above.

A basic result in stochastic calculus, which we do not prove here, is Itô's lemma. It can be stated as follows.[10]

[10]If A is an $l \times l$ square matrix with entries a_{ij}, $i, j \in \{1, 2, \ldots, l\}$, then $\mathrm{tr}(A)$ denotes the trace of A, which is defined as $\mathrm{tr}(A) = \sum_{i=1}^{l} a_{ii}$. Furthermore, for any (not necessarily square) matrix B we denote by B' the transpose of B.

Lemma 8.2 *Suppose that $x(\cdot)$ solves the stochastic differential equation (8.29). Let $G : X \times [0, T\rangle \mapsto \mathbb{R}$ be a (nonrandom) function with continuous partial derivatives G_t, G_x, and G_{xx}. Then the function $g(t) = G(x(t), t)$ satisfies the stochastic differential equation*

$$\begin{aligned} dg(t) = \Big\{&G_t(x(t), t) + G_x(x(t), t)f(x(t), u(t), t) \\ &+ (1/2)\mathrm{tr}\big[G_{xx}(x(t), t)\sigma(x(t), u(t), t)\sigma(x(t), u(t), t)'\big]\Big\}\, dt \\ &+ G_x(x(t), t)\sigma(x(t), u(t), t)\, dw(t). \end{aligned}$$

We emphasize that any solution to the stochastic differential equation (8.29) is a stochastic process depending on the realizations of $\xi \in \Xi$. We do not make this dependence explicit but simply write $x(t)$ instead of the correct, but more cumbersome, notation $x(t, \xi)$. The reader should keep in mind that the value of x at time t cannot be known in advance without knowing the realization of ξ. A similar remark applies to the control paths. As mentioned in section 8.1.1, controls usually depend on the most recent information and are random variables, too. However, we simply write $u(t)$ instead of $u(t, \xi)$.

As in the case of piecewise deterministic optimal control models, we consider an expected value of an integral as the objective functional. Again, we write $E_{u(\cdot)}$ for the expectation operator because the distribution of the state $x(\cdot)$ depends on the control path $u(\cdot)$, through the differential equation (8.29).

More specifically, consider the problem of maximizing the objective functional

$$J(u(\cdot)) = E_{u(\cdot)}\left\{\int_0^T e^{-rt}F(x(t), u(t), t)\, dt + e^{-rT}S(x(T))\right\} \tag{8.30}$$

subject to the state equation (8.29) and the constraints $u(t) \in U(x(t), t)$. Here, the problem horizon may be finite or infinite such that $S(x) = 0$ for all $x \in X$ if $T = \infty$. As in the previous sections of this chapter, we say that a control path $u(\cdot)$ is nonanticipating if its value at time t does not depend on any uncertainty revealed after time t. In particular, $u(t, \xi)$ must not depend on realizations of the random variables $w(t + \tau, \xi)$ or $x(t + \tau, \xi)$ for any strictly positive number τ. We can now define feasible and optimal paths, respectively, for a stochastic optimal control problem.

Definition 8.4 A control path $u : [0, T\rangle \times \Xi \mapsto \mathbb{R}^m$ is feasible for the stochastic optimal control problem stated above if it is nonanticipating, if there exists a unique solution $x(\cdot)$ to the stochastic differential equation

(8.29), if the constraints $x(t) \in X$ and $u(t) \in U(x(t), t)$ are satisfied with probability 1 for all t, and if the integral in (8.30) is well defined. If T is finite then the control path $u(\cdot)$ is optimal if it is feasible and if $J(u(\cdot)) \geq J(\tilde{u}(\cdot))$ holds for all feasible control paths $\tilde{u}(\cdot)$. If T is infinite then a control path $u(\cdot)$ is a catching up optimal path if it is feasible and if $\liminf_{t\to\infty}[J_t(u(\cdot)) - J_t(\tilde{u}(\cdot))] \geq 0$ holds for all feasible control paths $\tilde{u}(\cdot)$. Here, the t-truncations J_t are defined analogously to definition 3.2, taking into account that we now maximize expected values of integrals.

It is clear that in the infinite horizon case one could also consider other optimality criteria (see definition 3.2). The reader will not find it difficult to adapt the following optimality conditions to other criteria. As in the deterministic model, the only thing that has to be changed is the transversality condition.

We now state sufficient optimality conditions for stochastic optimal control problems of the form defined above. These conditions are based on the HJB equation and they differ from the conditions of theorem 3.1 only by an additional term which has to be included owing to the stochastic nature of the problem.

Theorem 8.4 *Let $V : X \times [0, T\rangle \mapsto \mathbb{R}$ be a function with continuous partial derivatives V_t, V_x, and V_{xx} and assume that V satisfies the HJB equation*

$$rV(x, t) - V_t(x, t) = \max\Big\{F(x, u, t) + V_x(x, t)f(x, u, t) + (1/2)\mathrm{tr}\big[V_{xx}(x, t)\sigma(x, u, t)\sigma(x, u, t)'\big] \Big| u \in U(x, t)\Big\} \tag{8.31}$$

for all $(x, t) \in X \times [0, T\rangle$. Let $\Phi(x, t)$ denote the set of controls $u \in U(x, t)$ maximizing the right-hand side of (8.31) and let $u(\cdot)$ be a feasible control path, with corresponding state trajectory $x(\cdot)$, such that $u(t) \in \Phi(x(t), t)$ holds with probability 1 for almost all $t \in [0, T\rangle$.

(i) *If $T < \infty$ and if the boundary condition $V(x, T) = S(x)$ holds for all $x \in X$ then $u(\cdot)$ is an optimal control path.*
(ii) *If $T = \infty$ and if either V is bounded and $r > 0$, or V is bounded below and $\limsup_{t\to\infty} e^{-rt}E_{u(\cdot)}V(x(t), t) \leq 0$ holds, then $u(\cdot)$ is a catching up optimal control path.*

Proof The strategy of the proof is the same as in all our other sufficiency theorems based on an HJB equation. We shall therefore be rather brief and leave the details to the reader. Fix a realization $\xi \in \Xi$ and consider the feasible path $\tilde{u}(\cdot)$ with corresponding state trajectory $\tilde{x}(\cdot)$. Applying lemma 8.2, with $x(\cdot)$ and $u(\cdot)$ replaced by $\tilde{x}(\cdot)$ and $\tilde{u}(\cdot)$, respectively, and with $G(x, t) = e^{-rt}V(x, t)$, we obtain

$$-dg(t) = e^{-rt}\Big\{rV(\tilde{x}(t), t) - V_t(\tilde{x}(t), t) - V_x(\tilde{x}(t), t)f(\tilde{x}(t), \tilde{u}(t), t)$$
$$- (1/2)\mathrm{tr}\big[V_{xx}(\tilde{x}(t), t)\sigma(\tilde{x}(t), \tilde{u}(t), t)\sigma(\tilde{x}(t), \tilde{u}(t), t)'\big]\Big\}\,dt$$
$$- e^{-rt}V_x(\tilde{x}(t), \tilde{u}(t), t)\sigma(\tilde{x}(t), \tilde{u}(t), t)\,dw(t).$$

Because of the HJB equation the term in curly brackets on the right-hand side of this equation is greater than or equal to $F(\tilde{x}(t), \tilde{u}(t), t)$. Thus, by integration we obtain

$$-e^{-rT}V(\tilde{x}(T), T) \geq -V(x_0, 0) + \int_0^T e^{-rt}F(\tilde{x}(t), \tilde{u}(t), t)\,dt$$
$$- \int_0^T e^{-rt}V_x(\tilde{x}(t), t)\sigma(\tilde{x}(t), \tilde{u}(t), t)\,dw(t).$$

Now apply the expectation operator $E_{\tilde{u}(\cdot)}$ to both sides of this inequality. It follows from the definition of the Wiener process that the expectation of the stochastic integral on the right-hand side is equal to 0 whenever $\tilde{u}(\cdot)$ is nonanticipating.[11] Therefore, we obtain

$$E_{\tilde{u}(\cdot)}\left\{\int_0^T e^{-rt}F(\tilde{x}(t), \tilde{u}(t), t)\,dt\right\} \leq V(x_0, 0) - e^{-rT}E_{\tilde{u}(\cdot)}[V(\tilde{x}(T), T)].$$

A similar calculation can be made with $\tilde{u}(\cdot)$ and $\tilde{x}(\cdot)$ replaced by $u(\cdot)$ and $x(\cdot)$, respectively. In that case the inequality sign can be replaced by an equality sign because of the condition $u(t) \in \Phi(x(t), t)$. Taking all these observations together, and using the transversality conditions, one easily arrives at the conclusion of the theorem. ■

8.2.2 Markovian Nash equilibria

In this section we generalize the stochastic optimal control problem from section 8.2.1 to a stochastic differential game in which the state equation contains a white noise term. Let there be N players and denote by $u^i(t)$ the control value chosen by player $i \in \{1, 2, \ldots, N\}$ at time t. As usual, we denote the vector of controls used by the opponents of player i by

$$u^{-i}(t) = (u^1(t), u^2(t), \ldots, u^{i-1}(t), u^{i+1}(t), \ldots, u^N(t)).$$

[11]This can be seen by using the definition of an integral as a limit of finite sums, and recalling that the expectation of a product of two independent random variables is equal to the product of the expectations of the two random variables.

If the state of the system at time $t \in [0, T\rangle$ is equal to $x(t) \in X$, player i's set of feasible controls is given by $U^i(x(t), u^{-i}(t), t) \subseteq \mathbb{R}^{m^i}$. The state equation for the game is

$$dx(t) = f(x(t), u^1(t), u^2(t), \ldots, u^N(t), t)\,dt + \sigma(x(t), u^1(t), u^2(t), \ldots, u^N(t), t)\,dw(t),$$

where $w(\cdot)$ is a k-dimensional standardized Wiener process and f and σ are functions defined on

$$\Omega = \{(x, u^1, u^2, \ldots, u^N, t) \mid x \in X, u^i \in U^i(x, u^{-i}, t), t \in [0, T\rangle\}$$

with values in $\mathbb{R}^n$ and $\mathbb{R}^{n \times k}$, respectively. Player i's goal is to maximize

$$J^i(u^i(\cdot)) = E_{u(\cdot)}\left\{\int_0^T e^{-r^i t} F^i(x(t), u^1(t), u^2(t), \ldots, u^N(t), t)\,dt + e^{-r^i T} S^i(x(T)) \,\Big|\, x(0) = x_0\right\},$$

where F^i is a real-valued utility function defined on Ω, S^i is a real-valued scrap value function defined on X, and r^i is the discount rate of player i. The deterministic vector $x_0 \in X$ is the initial state. We denote the stochastic differential game defined in this way by $\Gamma(x_0, 0)$. The differential game $\Gamma(x, t)$ is defined by replacing the time domain $[0, T\rangle$ by $[t, T\rangle$ and the initial state x_0 by x. Note that the Wiener process for the game $\Gamma(x, t)$ may start at the initial value $w(t) = 0$ since the particular choice of the initial value of w is irrelevant for the differential game.

In the piecewise deterministic games discussed in the first part of this chapter uncertainty is resolved only at the discrete jump times $T_1, T_2, \ldots$. Both the Markovian strategies and the piecewise open-loop strategies make use of the latest information about the random process by conditioning the controls on the current state of the mode process $h(t) = h(\ell(t))$. In the present framework of a system which is disturbed by a white noise process, uncertainty is resolved continuously. Hence only an information structure at least as fine as the Markovian one is up to date with the most recent information about the uncertainty of the system. Consequently, we confine the analysis of equilibria in the present model to Markovian Nash equilibria. These equilibria are subgame perfect in a sense made precise in definition 8.5 below.

The exact definition of a Markovian Nash equilibrium for the stochastic game defined above is almost identical to the corresponding definition of the deterministic game in section 4.1. Assume that all players except for player i determine their actions by the Markovian strategies

$u^j(t) = \phi^j(x(t), t), j \neq i$. Then, player i's decision problem can be rewritten as

$$\text{Maximize } J^i_{\phi^{-i}}(u^i(\cdot))$$
$$= E_{u(\cdot)}\left\{\int_0^T e^{-r^i t} F^i_{\phi^{-i}}(x(t), u^i(t), t)\, dt + e^{-r^i T} S^i(x(T)) \,\Big|\, x(0) = x_0\right\}$$

$$\text{subject to } dx(t) = f^i_{\phi^{-i}}(x(t), u^i(t), t)\, dt + \sigma^i_{\phi^{-i}}(x(t), u^i(t), t)\, dw(t) \qquad (8.32)$$
$$x(0) = x_0$$
$$u^i(t) \in U^i_{\phi^{-i}}(x(t), t),$$

where

$$F^i_{\phi^{-i}}(x, u^i, t) = F^i\big(x, \phi^1(x, t), \ldots, \phi^{i-1}(x, t), u^i, \phi^{i+1}(x, t), \ldots, \phi^N(x, t), t\big),$$
$$f^i_{\phi^{-i}}(x, u^i, t) = f\big(x, \phi^1(x, t), \ldots, \phi^{i-1}(x, t), u^i, \phi^{i+1}(x, t), \ldots, \phi^N(x, t), t\big),$$
$$\sigma^i_{\phi^{-i}}(x, u^i, t) = \sigma\big(x, \phi^1(x, t), \ldots, \phi^{i-1}(x, t), u^i, \phi^{i+1}(x, t), \ldots, \phi^N(x, t), t\big),$$
$$U^i_{\phi^{-i}}(x, t) = U^i\big(x, \phi^1(x, t), \ldots, \phi^{i-1}(x, t), \phi^{i+1}(x, t), \ldots, \phi^N(x, t), t\big).$$

For any given $(N-1)$-tuple $\phi^{-i} = (\phi^1, \ldots, \phi^{i-1}, \phi^{i+1}, \ldots, \phi^N)$ of functions $\phi^j : X \times [0, T\rangle \mapsto \mathbb{R}^{m^j}, j \neq i$, the problem (8.32) is an optimal control model of the form described in the previous section. Markovian Nash equilibria for this differential game are defined in the following, obvious way.

Definition 8.5 The N-tuple $(\phi^1, \phi^2, \ldots, \phi^N)$ of functions $\phi^i : X \times [0, T\rangle \mapsto \mathbb{R}^{m^i}$, $i \in \{1, 2, \ldots, N\}$, is a Markovian Nash equilibrium for the stochastic game defined above if, for each $i \in \{1, 2, \ldots, N\}$, a (catching up) optimal control path $u^i(\cdot)$ of the problem (8.32) exists and is defined by the Markovian strategy $u^i(t) = \phi^i(x(t), t)$. The equilibrium is said to be subgame perfect if, for every pair $(x, t) \in X \times [0, T\rangle$, there exists a Markovian Nash equilibrium $(\psi^1, \psi^2, \ldots, \psi^N)$ of the game $\Gamma(x, t)$ such that $\psi^i(y, s) = \phi^i(y, s)$ holds for all $(y, s) \in X \times [t, T\rangle$.

In the following theorem we state conditions which ensure that a given N-tuple of functions is a subgame perfect Markovian Nash equilibrium. These conditions are very similar to those of theorem 4.4.

Theorem 8.5 *Let $(\phi^1, \phi^2, \ldots, \phi^N)$ be a given N-tuple of functions $\phi^i : X \times [0, T\rangle \mapsto \mathbb{R}^{m^i}$ and make the following assumptions:*

(i) for every pair $(y, s) \in X \times [0, T\rangle$ there exists a unique solution $x_{y,s} : [s, T\rangle \times \Xi \mapsto X$ of the stochastic initial value problem

$$dx(t) = f(x(t), \phi^1(x(t), t), \phi^2(x(t), t), \dots, \phi^N(x(t), t), t)\, dt$$
$$+ \sigma(x(t), \phi^1(x(t), t), \phi^2(x(t), t) \dots, \phi^N(x(t), t), t)\, dw(t)$$
$$x(s) = y,$$

(ii) *for all* $i \in \{1, 2, \dots, N\}$ *there exists a function* $V^i : X \times [0, T\rangle \mapsto \mathbb{R}$, *having continuous partial derivatives* V^i_t, V^i_x, *and* V^i_{xx}, *such that the HJB equations*

$$r^i V^i(x, t) - V^i_t(x, t) = \max \Big\{ F^i_{\phi^{-i}}(x, u^i, t) + V^i_x(x, t) f^i_{\phi^{-i}}(x, u^i, t) \quad (8.33)$$
$$+ (1/2)\mathrm{tr}\Big[V^i_{xx}(x, t)\sigma^i_{\phi^{-i}}(x, u^i, t)\sigma^i_{\phi^{-i}}(x, u^i, t)' \Big] \mid u^i \in U^i_{\phi^{-i}}(x, t) \Big\}$$

are satisfied for all $i \in \{1, 2, \dots, N\}$ *and all* $(x, t) \in X \times [0, T\rangle$,

(iii) *if* $T < \infty$ *then* $V^i(x, T) = S^i(x)$ *for all* $i \in \{1, 2, \dots, N\}$ *and all* $x \in X$,

(iv) *if* $T = \infty$ *then for all* $i \in \{1, 2, \dots, N\}$ *either* V^i *is a bounded function and* $r^i > 0$, *or* V^i *is bounded below and* $\limsup_{t\to\infty} e^{-r^i t} E_{u(\cdot)} V^i(x_{y,s}(t), t) \leq 0$ *along all state trajectories* $x_{y,s}(\cdot)$ *mentioned in condition (i) above.*[12]

Denote by $\Phi^i(x, t)$ *the set of all* $u^i \in U^i_{\phi^{-i}}(x, t)$ *which maximize the right-hand side of (8.33). If* $\phi^i(x, t) \in \Phi^i(x, t)$ *holds for all* $i \in \{1, 2, \dots, N\}$ *and all* $(x, t) \in X \times [0, T\rangle$ *then* $(\phi^1, \phi^2, \dots, \phi^N)$ *is a subgame perfect Markovian Nash equilibrium. (If* $T = \infty$, *optimality is understood in the sense of catching up optimality.)*

Proof Apply theorem 8.4 to the individual stochastic optimal control problems (8.32) derived from the game $\Gamma(x_0, 0)$ as well as to the corresponding individual stochastic optimal control problems derived from the game $\Gamma(x, t)$ for all $(x, t) \in X \times [0, T\rangle$. ∎

Owing to the second order term $V^i_{xx}(x, t)$ occurring in the HJB equations (8.33), it is usually very difficult to find closed-form solutions for the optimal value functions V^i. Two classes of stochastic games, however, are known in which the solution of the HJB equations is not more difficult than in the deterministic case.

The first class of games are stochastic games in which the HJB equations of the corresponding deterministic game, obtained by setting $\sigma(x, u^1, u^2, \dots, u^N, t) = 0$ in (8.33), have solutions V^i which are linear in the state x. If this is the case, the same functions V^i are also solutions to the original HJB equations (8.33) because the term

[12] $E_{u(\cdot)}$ denotes the expectation with respect to the probability distribution generated by the control paths $u(\cdot)$ corresponding to the Markovian strategies ϕ^i, $i \in \{1, 2, \dots, N\}$.

$$\mathrm{tr}\left[V^i_{xx}(x,t)\sigma^i_{\phi^{-i}}(x,u^i,t)\sigma^i_{\phi^{-i}}(x,u^i,t)'\right]$$

vanishes whenever V^i is linear with respect to x. An example of this sort will be discussed in exercise 5. Linearity of the optimal value function with respect to the state is also a feature of linear state games as defined in section 7.2.

The second class of differential games with white noise in which the solution of the HJB equations is not more difficult than in the deterministic case is linear quadratic games. In the stochastic framework of this section, a game is called linear quadratic if the functions F^i, S^i, f, and U^i have the properties required in section 7.1 and if the function σ is linear with respect to the state x and the controls u^i, $i = 1, 2, \dots, N$. In a game with these properties, the usual guess of quadratic value functions leads to solutions of the HJB equations. We now illustrate this class of games by an example.

Example 8.3 Consider the stochastic differential game, with N players and finite time horizon T, defined by

$$\begin{aligned}
F^i(x, u^1, u^2, \dots, u^N, t) &= -x^2/2 - (u^i)^2/2,\\
S^i(x) &= 0,\\
f(x, u^1, u^2, \dots, u^N, t) &= \sum_{i=1}^N u^i,\\
\sigma(x, u^1, u^2, \dots, u^n, t) &= e^{-\alpha t},\\
U^i(x, u^{-i}, t) &= \mathbb{R},\\
r^i &= r
\end{aligned}$$

for all $i \in \{1, 2, \dots, N\}$. Here, α is a real constant and $r > 0$ is the common discount rate of all players. Since the game is completely symmetric with respect to the players we look for symmetric equilibria $(\phi^1, \phi^2, \dots, \phi^N)$ where $\phi^i = \phi$ holds for all $i \in \{1, 2, \dots, N\}$ and some function $\phi : X \times [0, T] \mapsto \mathbb{R}$. In such an equilibrium the optimal value functions of the N players must be identical and we can write V instead of V^i. With these remarks in mind we state the HJB equation as follows:

$$\begin{aligned}
rV(x,t) - V_t(x,t) = \max\{&-[x^2 + (u^i)^2]/2 + V_x(x,t)\left[u^i + (N-1)\phi(x,t)\right]\\
&+ (e^{-2\alpha t}/2)V_{xx}(x,t) \mid u^i \in \mathbb{R}\}.
\end{aligned}$$

The right-hand side is maximized by $u^i = \phi(x,t) = V_x(x,t)$. Hence the equation can be rewritten as

$$rV(x,t) - V_t(x,t) = -x^2/2 + (2N-1)V_x(x,t)^2/2 + (e^{-2\alpha t}/2)V_{xx}(x,t). \tag{8.34}$$

Let us guess a solution of the form $V(x,t) = A(t)x^2 + B(t)$. Because of the transversality condition (iii) of theorem 8.5, we must have $A(T) = B(T) = 0$. Furthermore, by substituting $V(x,t) = A(t)x^2 + B(t)$ into (8.34), and collecting terms with equal powers of x, we see that, in order for (8.34) to be satisfied, it must hold that

$$rA(t) - \dot{A}(t) = -(1/2) + 2(2N-1)A(t)^2, \qquad A(T) = 0;$$
$$rB(t) - \dot{B}(t) = e^{-2\alpha t}A(t), \qquad B(T) = 0.$$

The first of these equations is a Riccati differential equation for the function $A(\cdot)$. Its unique solution is given by

$$A(t) = \frac{1 - e^{C(t-T)}}{(r-C)e^{C(t-T)} - r - C}$$

with $C = \sqrt{r^2 + 8N - 4}$. The second equation is a linear differential equation which has the unique solution

$$B(t) = e^{rt}\int_t^T e^{-(r+2\alpha)s}A(s)\,ds.$$

We have found functions $V^i(x,t) = V(x,t) = A(t)x^2 + B(t)$ such that conditions (ii) and (iii) of theorem 8.5 are satisfied. The strategies $\phi^i(x,t) = \phi(x,t)$ must maximize the right-hand side of the HJB equation, which implies that $\phi(x,t) = V_x(x,t) = 2A(t)x$. Thus the stochastic differential equations in condition (i) of theorem 8.5 are given by

$$dx(t) = 2NA(t)x(t)\,dt + e^{-\alpha t}\,dw(t), \qquad x(s) = y.$$

These linear stochastic differential equations have the solutions

$$x_{y,s}(t) = ye^{2N\int_s^t A(\tau)\,d\tau} + \int_s^t \exp\left\{-\alpha\tau + 2N\int_\tau^t A(z)\,dz\right\} dw(\tau).$$

Therefore, condition (i) of the theorem is satisfied as well. This concludes the solution of this example.

As a final remark let us mention that, in this particular example, the uncertainty does not affect the equilibrium strategies (although it affects the optimal values). More specifically, the strategies do not depend on the parameter α, which is inversely related to the variance of the disturbances. As α tends to $+\infty$ the uncertainty in the model vanishes and it follows that the equilibrium strategies of the stochastic game are the same as those of the corresponding deterministic game in which the Wiener process is replaced by a constant. This is an instance of the so called certainty equivalence principle.

8.3 Further reading

Piecewise deterministic control systems have been studied by Rishel [205, 206], Davis [37], and Vermes [238]. An early application of these techniques is the analysis of a capacity expansion model in Davis et al. [39]. More recent accounts of the literature are Rishel [207] and Davis [38]. The special case in which there is no state variable and only one possible switch between modes has been considered in a number of papers and can easily be reformulated in the form of a standard deterministic optimal control problem with the survivor function as a state variable. Examples of this approach are the machine maintenance problem in Kamien and Schwartz [145] and the entry model in Kamien and Schwartz [144].

Differential games with general piecewise deterministic processes have been considered in only a few papers; see Başar and Haurie [3], Breton and Haurie [11], and Haurie [121, 122]. Another example is the model from Harris and Vickers [116], which we shall discuss in section 10.3. The special case of piecewise deterministic differential games without a state equation occurs in the papers by, for example, Reinganum [197, 198] and Dockner et al. [51, 57]. See also section 10.2 below.

Wiener processes, stochastic differential equations, and Itô's lemma are discussed in, for example, Arnold [1] and Gihman and Skorohod [110]. A good account of the HJB equation for optimal control problems with white noise can be found in Fleming and Rishel [96]. Another useful reference for stochastic differential equations and optimal control problems with white noise is Malliaris and Brock [168] which, however, focuses on the stochastic maximum principle, an optimality condition which we do not discuss in this book.

Differential games with white noise are treated in Başar and Olsder [4]. It is also possible to consider stochastic differential games which contain both piecewise deterministic processes and white noise. Such games are discussed in Haurie [121, 122].

8.4 Exercises

1. Consider the piecewise deterministic optimal control model with modes $M = \{1, 2\}$, state space $X = [0, \bar{x}]$ for some $\bar{x} > 0$, control set $U(h, x) = [0, 1]$, discount rate $r > 1$, system dynamics $f(h, x, u) = -ux$, switching rates $q_{12}(x, u) = u$ and $q_{21}(x, u) = 0$, and utility functions

$F(1, x, u) = -u$, $F(2, x, u) = u$, $S_{12}(x) = S_{21}(x) = x$. Prove that the following rule generates an optimal control path:

$$u(t) = \begin{cases} 0 & \text{if } h(t) = 1 \text{ and } x(t) \in [0, 1 - 1/r), \\ 1 & \text{if } h(t) = 1 \text{ and } x(t) \in [1 - 1/r, \bar{x}], \\ 1 & \text{if } h(t) = 2. \end{cases}$$

2. Consider the piecewise deterministic differential game defined by $N = 2$, $M = \{1, 2\}$, $X = [-\bar{x}, \bar{x}]$, with $\bar{x}$ sufficiently large, discount rates $r^1 = r^2 = 1$, control sets

$$U^i(h, x, u^{-i}) = \begin{cases} [0, \infty) & \text{if } x = -\bar{x}, \\ \mathbb{R} & \text{if } x \in (-\bar{x}, \bar{x}), \\ (-\infty, 0] & \text{if } x = \bar{x}, \end{cases}$$

system dynamics

$$f(h, x, u^1, u^2) = \begin{cases} u^1 & \text{if } h = 1, \\ u^2 & \text{if } h = 2, \end{cases}$$

constant switching rates $q_{hk}(x, u^1, u^2) = 2$, profit rates $F^i(h, x, u^1, u^2) = -(u^i)^2/2$, and

$$S^i_{hk}(x) = \begin{cases} 0 & \text{if } i = h \neq k, \\ -5(x-1)^2/8 & \text{if } i = k = 1, h = 2, \\ -5(x+1)^2/8 & \text{if } i = k = 2, h = 1. \end{cases}$$

Show that if the system is in mode 1 and state x then player 1 is in exactly the same situation as player 2 when the mode is 2 and the state is $-x$. Explain why this suggests considering symmetric stationary Markovian Nash equilibria (ϕ^1, ϕ^2) of the form

$$\phi^2(2, x) = -\phi^1(1, -x) \quad \text{and} \quad \phi^2(1, x) = -\phi^1(2, -x)$$

or symmetric piecewise open-loop Nash equilibria (ψ^1, ψ^2) of the form

$$\psi^2(2, y, t) = -\psi^1(1, -y, t) \quad \text{and} \quad \psi^2(1, y, t) = -\psi^1(2, -y, t).$$

Prove that the pair (ϕ^1, ϕ^2) defined by

$$\phi^1(h, x) = \begin{cases} \dfrac{10}{13} - \dfrac{x}{2} & \text{if } h = 1, \\ 0 & \text{if } h = 2, \end{cases} \qquad \phi^2(h, x) = \begin{cases} 0 & \text{if } h = 1, \\ -\left(\dfrac{10}{13} + \dfrac{x}{2}\right) & \text{if } h = 2 \end{cases}$$

is a stationary Markovian Nash equilibrium of the game.
Prove that the pair (ψ^1, ψ^2) defined by

$$\psi^1(h,y,t)=\begin{cases}\left(\frac{10}{13}-\frac{y}{2}\right)e^{-t/2} & \text{if } h=1,\\ 0 & \text{if } h=2,\end{cases}$$

$$\psi^2(h,y,t)=\begin{cases}0 & \text{if } h=1,\\ -\left(\frac{10}{13}+\frac{y}{2}\right)e^{-t/2} & \text{if } h=2\end{cases}$$

is a piecewise open-loop Nash equilibrium of the game.
Assume that the state at jump time T_l is equal to y and compute the trajectory $x_{hy}(s)$ of the state during the interval $[T_l, T_{l+1})$ as a function of the mode h, the state at the last jump time $y = x(T_l)$, and $s = t - T_l$. Make this computation for the stationary Markovian Nash equilibrium and for the piecewise open-loop Nash equilibrium. Compare the results. Prove that the piecewise open-loop Nash equilibrium and the stationary Markovian Nash equilibrium are equivalent, in the sense that they generate the same control paths.

3. Consider the infinite horizon optimal control problem with white noise disturbances defined by $X = [0, \infty)$, $F(x, u, t) = u^\alpha$, $f(x, u, t) = \beta x - u$, $\sigma(x, u, t) = \gamma x$, and

$$U(x,t)=\begin{cases}\{0\} & \text{if } x=0,\\ [0,\infty) & \text{if } x>0.\end{cases}$$

Assume that the discount rate r and the parameters α, β, and γ satisfy the relations $0 < \alpha < 1$, $\max\{0, \alpha\beta - \alpha(1-\alpha)\gamma^2/2\} < r$. State the HJB equation for this problem and find a constant $A \in \mathbb{R}$ such that the function $V(x, t) = Ax^\alpha$ solves this equation. Show that for this value function V, the right-hand side of the HJB equation is maximized at $u = \phi(x, t) = x[r - \alpha\beta + \alpha(1-\alpha)\gamma^2/2]/(1-\alpha)$. Show that the state dynamics generated by this strategy are given by

$$dx(t) = \frac{\beta - r - \alpha(1-\alpha)\gamma^2/2}{1-\alpha}x(t)\,dt + \gamma x(t)\,dw(t)$$

and that the process

$$x(t) = x(0)\exp\left\{\left[\frac{\beta - r}{1-\alpha} - (1+\alpha)\frac{\gamma^2}{2}\right]t + \gamma w(t)\right\}$$

solves the stochastic differential equation for the state dynamics. Using the fact that $w(t)$ has a normal distribution with mean 0 and variance t compute the expectation of $V(x(t), t) = Ax(t)^\alpha$. Finally, show that condition (ii) of theorem 8.4 holds; that is, verify

$$\limsup_{t\to\infty} e^{-rt} E_{u(\cdot)} V(x(t), t) \le 0.$$

4. Generalize the previous exercise to a symmetric N-player differential game; that is, consider the game with white noise disturbances in which X, $\sigma(\cdot)$, and $U^i(x, t) = U(x, t)$ are as in exercise 3, $F^i(x, u, t) = (u^i)^\alpha$, and $f(x, u^1, u^2, \ldots, u^N, t) = \beta x - \sum_{i=1}^N u^i$. Assume that all players have the same discount rate r. Try to find the parameter restrictions which allow you to follow the same steps as in exercise 3.
5. Consider a stochastic variation on the model in exercise 4.2, replacing the deterministic state equation $\dot{x}(t) = -\sqrt{x(t)} \sum_{i=1}^N u^i(t)$ by the stochastic equation $dx(t) = -\sqrt{x(t)} \sum_{i=1}^N u^i(t)\, dt + \alpha x(t)\, dw(t)$, where α is a real constant. Show that the Markovian Nash equilibrium for the deterministic game derived in exercise 4.2 also qualifies as a Markovian Nash equilibrium for the stochastic game.
6. Reconsider the model of example 8.3 with $\sigma(x, u^1, u^2, \ldots, u^N, t) = \alpha x$ instead of $\sigma(x, u^1, u^2, \ldots, u^N, t) = e^{-\alpha t}$. Find a symmetric Markovian Nash equilibrium of this game. Compute the stochastic process which describes the state in this equilibrium.

Part II Applications

9 Capital accumulation games

One of the driving forces in a market economy is the growth of firms and industries. While traditionally economists have analysed firm and industry growth under the assumption of perfectly competitive product markets (i.e., firms are assumed to be price takers in the output market) more recent research has focused on game theoretic models of growth and capital accumulation. Therefore the aim of this chapter is to present a number of differential games in which two or more firms invest strategically in a physical capital stock and the output market is organized by oligopolistic competition. In such a setting the dynamic evolution of a firm and that of an industry differ substantially from the predictions of perfectly competitive firm models. In particular, the strategic interactions among rival firms give rise to a number of interesting conclusions, such as overaccumulation of capital, preemption, and entry (mobility) deterrence.

We begin with the description of a general Cournot model in which two firms invest in their capital stocks. If they face capacity constraints, oligopolistic product market competition allows us to derive a reduced form profit function for each firm that is dependent on the firm's own capital stock as well as on the capital stock of the rival. Investment is costly and can give rise to internal adjustment costs. In this setting we study two alternative games, one in which firms employ open-loop strategies and one in which strategic interactions are explicitly taken into account through state dependent (nondegenerate) Markovian strategies. In case of open-loop strategies firms play a precommitment game and we find out that the steady state equilibrium coincides with the outcome of the single-period Cournot game. While in some cases precommitment is a useful property, in the case of capital accumulation games it ignores the strategic interactions among firms that can give rise to interesting industry evolutions. Therefore, in the second case, we derive a Markov perfect Nash equilibrium and relate its characteristics to the outcome of the precommitment game. It turns out that the Markov perfect Nash equilibrium gives rise to a number of interesting predictions. Firms have an incentive to produce more in case of a Markov perfect Nash equilibrium, which leads to a more competitive outcome than the static Cournot game. This observation leads us to conclude that, in a Markov perfect Nash equilibrium, strategically investing firms recognize the preemptive role of capital and use it to credibly deter entry in an industry. The equilibrium outcomes, however, do crucially depend on the assumption of whether or not investment is reversible or irreversible. Therefore we study both cases, with and without adjustment costs.

While in the traditional models of capital accumulation firms invest in a privately owned capital stock, we also present a game in which two agents invest in a public stock of capital. There are at least two possible economic interpretations of such a game. One relates to the process of transboundary pollution accumulation (see chapter 12) and the other to the case of knowledge accumulation in an economy where the stock of knowledge is a pure public good. In the knowledge accumulation game we are primarily interested in whether or not dynamic equilibria result in a free rider problem. It turns out that, while for both alternative games (the game in which agents employ open-loop strategies and the game in which they employ state dependent decision rules) a free rider problem exists, it is more severe in the case of a Markov perfect Nash equilibrium than in the open-loop game.

9.1 The structure of capital accumulation games

In this section we present a general model that incorporates the features of a strategic capital accumulation game. Without loss of generality we restrict the analysis to the case of two firms and refer the reader interested in N-player games to the references given in section 9.6. Throughout the chapter we make use of the following notation. Let $K^i(t)$ be the physical capital stock of firm i at time t. Each firm in the industry accumulates capital according to the equation

$$\dot{K}^i(t) = I^i(t) - \delta^i K^i(t), \tag{9.1}$$

where $I^i(t)$ is gross investment of firm i at time t and $\delta^i \geq 0$ is the constant rate of depreciation. According to equation (9.1) we can distinguish two alternative specifications. In the case where $\delta^i = 0$ and $I^i(t)$ is restricted to be nonnegative, investment is irreversible. This implies that once firm i has accumulated capital up to a level $\bar{K}^i$ it is locked in and capital cannot be adjusted to some lower level $K^i(t) < \bar{K}^i$. This case will be important later on when we discuss the strategic implications of investment over time. In the case where $\delta^i > 0$ or $I^i(t)$ is unrestricted investment is reversible. Both firms are assumed to be equipped with initial stocks of capital equal to $K^i(0) = K_0^i \geq 0$.[1]

Firms operate in oligopolistic output markets and choose either prices or quantities as their strategic variables (in addition to their investments). As an example, consider the simple case in which firms face a Cournot output market (cf. chapter 2). Denote the output of firm i at time t by $Q^i(t)$ and industry output by $Q(t) = Q^i(t) + Q^j(t)$ and assume that the market price at time t is $P(Q(t))$ where $P(\cdot)$ is the inverse demand function. Production costs at time t are given by $m^i(Q^i(t))$. If firms are operating in the product market subject to a capacity constraint $Q^i(t) \leq K^i(t)$, competition implies that each firm chooses an equilibrium output that is a function both of its own level of capital stock and of the capital stock of the rival firm. This observation allows us to write the profit rate at time t of each firm as a function of the two capital stocks $K^i(t)$ and $K^j(t)$. This reduced-form profit function is

$$\pi^i(K^i, K^j) = P(Q(K^i, K^j))Q^i(K^i, K^j) - m^i(Q^i(K^i, K^j)),$$

where, with a slight abuse of notation, we have written Q^i and Q as functions of the capital stocks. In what follows we will only work with the reduced profit function π^i.

Investment is costly and may result in additional adjustment costs. If the price of a unit of investment is constant and given by $\rho > 0$ the instantaneous net profit rate of firm i at time t is given by

$$\pi^i(K^i(t), K^j(t)) - \rho I^i(t) - C^i(I^i(t)),$$

where $C^i(I^i(t))$ denotes the internal adjustment costs. These costs occur since not only must firms pay the price for investment but they also have to adjust current plant so that the new capital can be used.

Firms choose an investment path so as to maximize the present value of future profits

$$J^i = \int_0^T e^{-rt}[\pi^i(K^i(t), K^j(t)) - \rho I^i(t) - C^i(I^i(t))]\, dt \tag{9.2}$$

over a given finite ($T < \infty$) or infinite ($T = \infty$) planning period and subject to the accumulation equation

$$\dot{K}^i(t) = I^i(t) - \delta^i K^i(t), \quad K^i(0) = K_0^i, \tag{9.3}$$

with $r \geq 0$ as the constant rate of discount.

Throughout this chapter we make use of a number of assumptions.

[1]Strictly speaking, if $K_0^i = 0$ firm i is not equipped.

Assumption 9.1 The profit functions $\pi^i(K^i, K^j)$, $i \in \{1, 2\}$, $i \neq j$, are twice continuously differentiable, increasing and strictly concave in K^i, and decreasing in K^j. Moreover $\pi^i_{K^i}(K^i, K^j)$ is bounded from above and $\pi^i_{K^iK^j}(K^i, K^j) < 0$.

It is easily seen that, if the reduced-form profit function is derived from oligopolistic product market competition, assumption 9.1 is satisfied. Moreover, since the inequality $\pi^i_{K^iK^j}(K^i, K^j) < 0$ holds, the products are strategic substitutes and the firms face downward sloping reaction functions in the (K^1, K^2) space.

Assumption 9.2 The adjustment cost functions $C^i(I^i)$ are twice continuously differentiable and strictly convex, with $C^i(0) = 0$ and $C^i_{I^i}(0) = 0$.

The general capital accumulation game (9.2)–(9.3) can be used to study the set of questions that we raised in the introduction to this chapter. In particular, we want to explore the relationship between Markov equilibria of the dynamic game and the unique equilibrium of a corresponding static game, the stability properties of the long-run equilibrium, and the preemptive role of capital, i.e., whether a firm that has a head start can deter mobility by overaccumulating capital.

9.2 Qualitative properties of equilibrium strategies

Let us start with a general analysis of the qualitative properties of Markovian equilibria of the game (9.2)–(9.3). First, we note that although it is possible to prove existence of open-loop Nash equilibria for both the finite as well as the infinite horizon model, a general existence proof for Markov perfect Nash equilibria is not available in the literature. Therefore a detailed analysis of Markov perfect Nash equilibria requires more specific assumptions about the profit and the cost functions. For now, we simply assume that Markov perfect equilibria exist and try to shed some light on their qualitative characteristics.

For that purpose let us assume that the two firms operate in a homogeneous product market and play a Cournot game with capacity constraints and internal adjustment costs. We set the price of a unit of investment and the rate of depreciation equal to zero, that is, $\rho = \delta^1 = \delta^2 = 0$, and allow investment to take on any value in the interval $(-\infty, \infty)$ so that investment is perfectly reversible. Moreover, we consider an infinite horizon model, that is, $T = \infty$. In this setting, the capital accumulation game can be interpreted as an infinite horizon Cournot game with output adjustment costs.

The reduced-form profit functions are now given by

$$\pi^i(K^i, K^j) = P(K^i + K^j)K^i - m^i(K^i) \tag{9.4}$$

so that the objective functions become

$$J^i = \int_0^\infty e^{-rt}[P(K^i(t) + K^j(t))K^i(t) - m^i(K^i(t)) - C^i(I^i(t))]\,dt. \tag{9.5}$$

These objective functions are to be maximized subject to the constraints

$$\dot{K}^i(t) = I^i(t) \tag{9.6}$$

and the initial conditions.

Since our interest is to compare the qualitative properties of the dynamic game with those of the one-shot game let us briefly recall the equilibrium of the Cournot game with profit functions (9.4). First, it is easy to show that the single period Cournot game admits a unique equilibrium that is characterized by the first order conditions $P'(K^i + K^j)K^i + P(K^i + K^j) - m^i_{K^i}(K^i) = 0$. These conditions can be rewritten as

$$P(Q)\left(1 + \frac{S^i}{\eta}\right) = m^i_{K^i}(K^i), \tag{9.7}$$

where $Q = K^1 + K^2$ is industry output, $S^i = K^i/Q$ is the market share of firm i, and $\eta = P(Q)/[QP'(Q)]$ is the price elasticity of industry demand.

In deriving the equilibrium equation (9.7) we made use of the Cournot reaction pattern. This means that firm i, when it changes the output, assumes that its rival firm j will not react to this change. While this assumption is not consistent with actual reaction behaviour out of equilibrium, it is true at the equilibrium. An alternative to the Cournot assumption is to allow firms to have nonzero conjectural variations. In this case firm i assumes that, if it changes its output, firm j will react according to $dK^j/dK^i = \xi$. This implies that the first order conditions become $P'(K^i + K^j)K^i(1 + \xi) + P(K^i + K^j) - m^i_{K^i}(K^i) = 0$. Thus, a conjectural variations equilibrium is characterized by the equation

$$P(Q)\left[1 + \frac{S^i}{\eta}(1 + \xi)\right] = m^i_{K^i}(K^i). \tag{9.8}$$

The equilibria determined by (9.7) or (9.8) will serve as benchmarks for the dynamic game with adjustment costs. In particular, we are interested in the circumstances in which the static Cournot outcome is a good prediction of the equilibrium of the infinite horizon dynamic game. As we will see shortly, the static equilibrium is identical to the steady state outcome of the dynamic equilibrium if firms use open-loop strategies. Recall from chapter 4 that in such a game the firms choose a time profile of actions at the beginning of the game and commit themselves to retain these preannounced profiles for the rest of the game. In this sense the dynamic game mimics many features of the static Cournot model and hence the result should not come as a surprise. On the other hand, if firms choose state dependent decision rules as their strategies, we should expect an equilibrium outcome of the dynamic game that is different from the static Cournot result. In particular, it turns out that the steady state of the Markov perfect equilibrium corresponds to a conjectural variations equilibrium of the corresponding static game with negative conjectures.[2]

Let us now work out the details of these results. We first look at the dynamic equilibrium when firms play an open-loop game. Formulate the Hamiltonian of player i

$$H^i(K^i, K^j, I^i, \lambda_i^i, \lambda_j^i, t) = P(K^i + K^j)K^i - m^i(K^i) - C^i(I^i) + \lambda_i^i I^i + \lambda_j^i I^j(t),$$

where $\lambda_i^i(t)$ and $\lambda_j^i(t)$ are the current value adjoint variables. These variables satisfy the adjoint equations

$$\dot{\lambda}_i^i(t) = r\lambda_i^i(t) - P'(K^i(t) + K^j(t))K^i(t) - P(K^i(t) + K^j(t)) + m_{K^i}^i(K^i(t)) \tag{9.9}$$

and

$$\dot{\lambda}_j^i(t) = r\lambda_j^i(t) - P'(K^i(t) + K^j(t))K^i(t). \tag{9.10}$$

Since $I^i(t)$ is unconstrained, the maximum condition for each player is given by

$$C_{I^i}^i(I^i(t)) = \lambda_i^i(t). \tag{9.11}$$

A steady state of the open-loop equilibrium is defined as the solution to the following system of equations

$$\dot{K}^i(t) = I^i(t) = \dot{\lambda}_i^i(t) = \dot{\lambda}_j^i(t) = 0. \tag{9.12}$$

Assumption 9.2 together with the maximum condition (9.11) implies $\hat{\lambda}_i^i = 0$, where the hat notation refers to the steady state level. Together with the adjoint equation (9.9) this shows that

[2]From the definition of conjectural variations it is clear that negative conjectures imply that, whenever firm i increases its output, it conjectures that its rival firm will reduce its output. This reaction pattern implies more competitive behaviour so that we can expect an outcome of the game that is closer to the competitive level.

$$-P'(\hat{K}^i + \hat{K}^j)\hat{K}^i - P(\hat{K}^i + \hat{K}^j) + m^i_{K^i}(\hat{K}^i) = 0.$$

This can be written as

$$P(\hat{Q})\left(1 + \frac{\hat{S}^i}{\eta}\right) = m^i_{K^i}(\hat{K}^i).$$

This result, however, is the equilibrium equation (9.7) of the static Cournot model.

We have already pointed out that this result is not surprising since an open-loop Cournot game has many features of a static Cournot model. But we can only argue that the dynamic game mimics the static Cournot equilibrium if the open-loop equilibrium satisfies a stability property. A possible stability property is that the open-loop equilibrium possesses a unique steady state and that every equilibrium trajectory with an arbitrary initial capital stock converges to the steady state value, in which case we call the steady state globally asymptotically stable. For the open-loop model both properties apply, as we will demonstrate in the next section.

Now we consider nondegenerate Markovian equilibria. In this setting the firms choose investment strategies that are functions of the capital stocks of both firms in the industry, say $I^i(t) = \phi^i(K^i(t), K^j(t))$. As pointed out in several other places in this book, a state dependent decision rule is capable of capturing many features that are present in strategic competition. For our game, state dependent Markovian strategies imply that, whenever firm i makes a decision that results in a change in its capital stock, firm j immediately reacts. This action and reaction pattern lets us expect a steady state outcome of the game that is quite different from that of the single-period Cournot game or the open-loop game.

In case of an equilibrium (ϕ^1, ϕ^2) with stationary strategies as described above, the Hamiltonian functions are

$$H^i(K^i, K^j, I^i, \lambda^i_i, \lambda^i_j) = P(K^i + K^j)K^i - m^i(K^i)$$
$$- C^i(I^i) + \lambda^i_i I^i + \lambda^i_j \phi^j(K^j, K^i).$$

Hence, instead of (9.9) and (9.10), the adjoint equations are now

$$\dot{\lambda}^i_i(t) = r\lambda^i_i(t) - P'(K^i(t) + K^j(t))K^i(t) - P(K^i(t) + K^j(t)) + m^i_{K^i}(K^i(t))$$
$$- \lambda^i_j(t)\phi^j_{K^i}(K^j(t), K^i(t)) \qquad (9.13)$$

and

$$\dot{\lambda}^i_j(t) = r\lambda^i_j(t) - P'(K^i(t) + K^j(t))K^i(t) - \lambda^i_j(t)\phi^j_{K^j}(K^j(t), K^i(t)). \qquad (9.14)$$

The partial derivatives $\phi^j_{K^i}$ and $\phi^j_{K^j}$ capture the reactions of firm j to an increase in the two capital stocks.

As before, the maximum conditions are given by (9.11) so that at a steady state equilibrium where $\dot{\lambda}_i^i(t) = \dot{\lambda}_j^i(t) = 0$ holds we must have

$$\hat{\lambda}_j^i = \frac{P'(\hat{Q})\hat{K}^i}{r - \phi^j_{K^j}(\hat{K}^j, \hat{K}^i)}$$

and

$$P'(\hat{Q})\hat{K}^i\left[1 + \frac{\phi^j_{K^i}(\hat{K}^j, \hat{K}^i)}{r - \phi^j_{K^j}(\hat{K}^j, \hat{K}^i)}\right] + P(\hat{Q}) = m^i_{K^i}(\hat{K}^i),$$

which in turn can be rewritten as

$$P(\hat{Q})\left\{1 + \frac{\hat{S}^i}{\eta}\left[1 + \frac{\phi^j_{K^i}(\hat{K}^j, \hat{K}^i)}{r - \phi^j_{K^j}(\hat{K}^j, \hat{K}^i)}\right]\right\} = m^i_{K^i}(\hat{K}^i).$$

If we define

$$\xi = \frac{\phi^j_{K^i}(\hat{K}^j, \hat{K}^i)}{r - \phi^j_{K^j}(\hat{K}^j, \hat{K}^i)} \tag{9.15}$$

as the conjectural variation that firm i has about the reaction of its rival it follows that the steady state of the Markov perfect equilibrium coincides with a conjectural variations equilibrium of a corresponding static game.

This result has two important implications: first, it implies that Markov perfect Nash equilibria generate long-run behaviour different from that generated by open-loop Nash equilibria and, second, it demonstrates that any steady state of a Markov perfect equilibrium of the infinite horizon differential game can be viewed as a conjectural variations equilibrium of a corresponding static game. This demonstrates a close relationship between the equilibrium of a static game and the steady state of a Markov perfect equilibrium and justifies a static conjectural variations equilibrium approach to mimic dynamic competition. It is important to note, however, that the equilibrium conjectures derived in equation (9.15) are the result of equilibrium play by both firms in the industry and hence are not open to the critique frequently put forward in static conjectural variations equilibrium analyses.

In order to show whether strategic interactions in a Markovian game result in more or less competitive behaviour than in a Cournot game, it is necessary to derive the sign of the conjecture (9.15). If it is negative, the state dependent Markov equilibrium of the accumulation game results in more competitive behaviour than the static Cournot outcome. If it is positive, it results in more collusive behaviour.

Example 9.1 We now study a specific example to derive the closed form of the conjecture (9.15). For that matter, we assume that the inverse demand function is linear and given by $P(Q) = a - Q$ and that both cost curves are quadratic, i.e., $m^i(K^i) = \bar{m}^i K^i + (b/2)(K^i)^2$ and $C^i(I^i) = (k/2)(I^i)^2$, where a, b, $\bar{m}^i$, and k are positive parameters and $a > \bar{m}^i$. With these specifications, the differential game becomes one of the linear quadratic type for which we are able to derive a closed-form solution even when firms use state dependent Markov strategies. (Recall the discussion of linear quadratic games in chapter 7.)

We make use of the value function approach to derive a Markov perfect Nash equilibrium. The value functions have to satisfy the HJB equations

$$\begin{aligned} rV^i(K^i, K^j) = \ & \max\{(a - K^i - K^j)K^i - \bar{m}^i K^i - (b/2)(K^i)^2 - (k/2)(I^i)^2 \\ & + V^i_{K^i}(K^i, K^j)I^i + V^i_{K^j}(K^i, K^j)\phi^j(K^j, K^i) \mid I^i \in \mathbb{R}\}. \end{aligned} \tag{9.16}$$

Maximization of the right-hand side of (9.16) yields

$$I^i = \phi^i(K^i, K^j) = (1/k)V^i_{K^i}(K^i, K^j). \tag{9.17}$$

Substitution of (9.17) into (9.16) provides us with a system of partial differential equations for V^i and V^j. Since our problem is of the linear quadratic type, we guess quadratic value functions of the form

$$V^i(K^i, K^j) = \alpha + \beta_i K^i + \gamma K^j + (\delta/2)(K^i)^2 + (\epsilon/2)(K^j)^2 + \sigma K^i K^j. \tag{9.18}$$

Equation (9.18) shows that the value functions are symmetric except for β_i. It can be shown that the value functions (9.18) solve the partial differential equation system (9.16) if and only if the parameters satisfy the equations

$$\begin{aligned} 0 &= -r\alpha + \frac{1}{2k}\beta_i^2 + \frac{1}{k}\gamma\beta_j, \\ 0 &= -r\beta_i + a - \bar{m}^i + \frac{\beta_i}{k}\delta + \frac{1}{k}\sigma\beta_j + \frac{1}{k}\sigma\gamma, \\ 0 &= -r\gamma + \frac{\beta_i\sigma}{k} + \frac{\epsilon}{k}\beta_j + \frac{\delta}{k}\gamma, \\ 0 &= -\frac{r}{2}\delta - 1 - \frac{b}{2} + \frac{\delta^2}{2k} + \frac{\sigma^2}{k}, \\ 0 &= -\sigma r - 1 + \frac{2}{k}\delta\sigma + \frac{\epsilon\sigma}{k}, \\ 0 &= -\frac{r}{2}\epsilon + \frac{\sigma^2}{2k} + \frac{\epsilon\delta}{k}. \end{aligned}$$

Multiplying the last three equations through by k and defining $s = r/2$ results in

$$-ks\delta - k - \frac{bk}{2} + \frac{\delta^2}{2} + \sigma^2 = 0,$$
$$-2ks\sigma - k + 2\delta\sigma + \epsilon\sigma = 0,$$
$$-ks\epsilon + \frac{\sigma^2}{2} + \epsilon\delta = 0.$$

The first of these three equations is quadratic and has the solution

$$\delta = sk \pm \sqrt{s^2k^2 + 2k + bk - 2\sigma^2},$$

whereas the last one yields

$$\epsilon = \frac{\sigma^2}{2ks - 2\delta}.$$

If we now introduce the variables $\psi = \sigma^2$ and $f = s^2k + b + 2$, we get after some algebraic manipulations the cubic equation

$$81\psi^3 - 72kf\psi^2 + 8k^2(2f^2 + 1)\psi - 4k^3f = 0.$$

A solution to this equation provides the candidates

$$\phi^i(K^i, K^j) = (1/k)(\beta_i + \delta K^i + \sigma K^j)$$

for a Markov perfect Nash equilibrium. Altogether we get six candidates for a Markov perfect Nash equilibrium. Of these six candidates, we are interested only in those that generate a globally asymptotically stable equilibrium, since in that case the limiting transversality conditions are satisfied. Each candidate for a Markov perfect Nash equilibrium results in a system of state equations given by

$$\dot{K}^i(t) = \frac{1}{k}[\beta_i + \delta K^i(t) + \sigma K^j(t)]. \tag{9.19}$$

System (9.19) is globally asymptotically stable if and only if $(1/k)(\delta + \sigma) < 0$ and $(1/k)(\delta - \sigma) < 0$. In the present case this requires that we select the following solutions for δ and σ:

$$\delta = sk - \sqrt{s^2k^2 + 2k + bk - 2\sigma^2},$$
$$\sigma = -\sqrt{4k\left[2f - \sqrt{4f^2 - 6}\cos(\theta/3)\right]/27},$$

where $\theta = \arctan(\sqrt{-D}/(-m))$, $D = 432f^6(-64f^4 + 107f^2 + 128)/9^9 < 0$, and $m = 4k^3(32f^3 - 99f)/3^9$. Thus we have found a Markov perfect Nash equilibrium.

The linear quadratic model can be used to derive an explicit form of the dynamic conjecture derived in (9.15). Making use of the linear strategies given by

$$\phi^i(K^i, K^j) = \frac{1}{k}(\beta_i + \delta K^i + \sigma K^j) \tag{9.20}$$

we find that the conjecture is

$$\xi = \frac{\sigma}{rk - \delta}.$$

This formula implies two important results. First, it shows that the globally asymptotically stable steady state of a Markov perfect Nash equilibrium corresponds to a conjectural variations equilibrium of a static game with constant conjectures. Second, the conjectures ξ are negative. The last claim is easily derived from the stability conditions stated below (9.19). We know that $\sigma < 0$ and $\delta - \sigma < 0$ hold. Therefore we get $\delta - rk < \delta < \sigma < 0$, so that $\xi < 0$.

The decision rules (9.20) can now be used to relate the static to the dynamic reaction functions. Figure 9.1 presents the general result that the dynamic equilibrium results in more competitive behaviour than static Cournot competition. Hence both firms overinvest in capital. This overinvestment is chosen so as to preempt the rival firm.

9.3 Stability properties of equilibria

The analysis of the preceding section concentrated on some qualitative aspects of the steady states generated by the equilibria of the capital accumulation game with adjustment costs. In particular, we related the outcome of the open-loop game to the static Cournot model and the Markov perfect Nash equilibrium to a static conjectural variations equilibrium. We will now discuss the stability properties of these equilibrium steady states. In order to argue that the equilibrium outcome of a static game is reproduced by the equilibrium of a dynamic game we need to show that the steady state of the dynamic equilibrium is globally asymptotically stable. Only in this case is it legitimate to say that long-run behaviour in the dynamic game corresponds to the static Cournot outcome.

In deriving the Markov perfect Nash equilibrium, we have already made use of a stability argument. Out of the six solution candidates

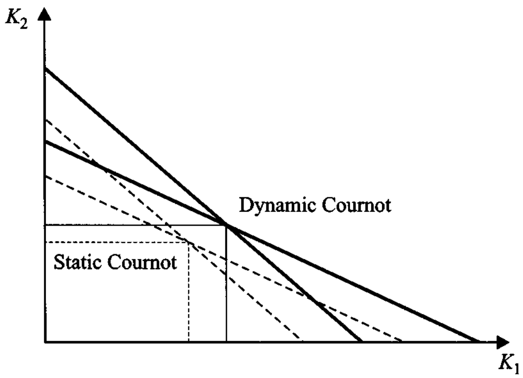

Figure 9.1 Reaction functions in the static and dynamic Cournot models

that satisfied the HJB equations, we have chosen the one that generates a globally asymptotically stable steady state. In this section we will look at the stability properties for the open-loop game.

In contrast to the preceding section, where we used the case of no depreciation but unbounded investment, we now look at the model with depreciation $\delta^i > 0$, $\rho = 0$, and nonnegative investment.

The game is therefore described by the objective functionals

$$J^i = \int_0^\infty e^{-rt}[\pi^i(K^i(t), K^j(t)) - C^i(I^i(t))]\, dt,$$

the accumulation equations

$$\dot{K}^i(t) = I^i(t) - \delta^i K^i(t), \tag{9.21}$$

and the nonnegativity constraints $I^i(t) \geq 0$.[3] Applying the maximum principle, we obtain the equilibrium conditions

$$\dot{\lambda}_i^i(t) = (r + \delta^i)\lambda_i^i(t) - \pi^i_{K^i}(K^i(t), K^j(t)), \tag{9.22}$$

$$C^i_{I^i}(I^i(t)) = \lambda_i^i(t). \tag{9.23}$$

Differentiation of (9.23) with respect to time results in

$$C^i_{I^iI^i}(I^i(t))\dot{I}^i(t) = \dot{\lambda}_i^i(t). \tag{9.24}$$

Let us now define

$$W^i(\dot{K}^i(t), K^i(t)) := -C^i(\dot{K}^i(t) + \delta^i K^i(t)). \tag{9.25}$$

[3]Throughout this section we restrict attention to interior maxima. This can be justified by making appropriate assumptions on the cost functions $C^i(\cdot)$.

Differentiation of the state equation (9.21) with respect to time yields (supposing that $K(t)$ is sufficiently smooth)

$$\ddot{K}^i(t) = \dot{I}^i(t) - \delta^i \dot{K}^i(t)$$

which, after substitution of (9.22), (9.24), (9.25), and $I^i(t) = \dot{K}^i(t) + \delta^i K^i(t)$, results in the following second order differential equation in the two capital stocks $(K^i(t), K^j(t))$:

$$\begin{aligned} &W^i_{\dot{K}^i\dot{K}^i}(\dot{K}^i(t), K^i(t))\ddot{K}^i(t) + W^i_{\dot{K}^i K^i}(\dot{K}^i(t), K^i(t))\dot{K}^i(t) \\ &= \pi^i_{K^i}(K^i(t), K^j(t)) + rW^i_{\dot{K}^i}(\dot{K}^i(t), K^i(t)) + W^i_{K^i}(\dot{K}^i(t), K^i(t)). \end{aligned}$$

In order to check the local stability properties of this equation, we linearize the system around the steady state $(\hat{K}^i, \hat{K}^j)$ and get the differential equation system

$$A\ddot{y}(t) - rA\dot{y}(t) - (rB + C)y(t) = 0, \tag{9.26}$$

where $y(t) = (K^i(t) - \hat{K}^i, K^j(t) - \hat{K}^j)'$ and

$$A = -\begin{pmatrix} C^i_{I^iI^i}(\delta^i\hat{K}^i) & 0 \\ 0 & C^j_{I^jI^j}(\delta^j\hat{K}^j) \end{pmatrix},$$

$$B = -\begin{pmatrix} \delta^i C^i_{I^iI^i}(\delta^i\hat{K}^i) & 0 \\ 0 & \delta^j C^j_{I^jI^j}(\delta^j\hat{K}^j) \end{pmatrix},$$

$$\begin{aligned} C = &\begin{pmatrix} \pi^i_{K^iK^i}(\hat{K}^i, \hat{K}^j) & \pi^i_{K^iK^j}(\hat{K}^i, \hat{K}^j) \\ \pi^j_{K^jK^i}(\hat{K}^j, \hat{K}^i) & \pi^j_{K^jK^j}(\hat{K}^j, \hat{K}^i) \end{pmatrix} \\ &+\begin{pmatrix} -(\delta^i)^2 C^i_{I^iI^i}(\delta^i\hat{K}^i) & 0 \\ 0 & -(\delta^j)^2 C^j_{I^jI^j}(\delta^j\hat{K}^j) \end{pmatrix}. \end{aligned}$$

From these definitions it is clear that both A and B are diagonal matrices that are nonsingular and the matrix C is the sum of a matrix of some second order partial derivatives of the profit functions π^i and π^j and a diagonal matrix with negative elements.

Even without any assumptions on the matrices A, B, and C we are able to get important insights into the stability properties of the dynamical system (9.26) that characterizes the equilibrium trajectory of the open-loop game. In particular, it can be shown that if ρ is a root of the characteristic equation of the linearized system (9.26), then $r - \rho$ is a root as well. To see this, proceed as follows. Using matrix notation, the characteristic equation of system (9.26) can be written as

$$\det\left[A\rho^2 - rA\rho - (C + rB)\right] = 0. \tag{9.27}$$

If we replace ρ by $r-\rho$ in equation (9.27), we get

$$\det[A(r-\rho)^2 - rA(r-\rho) - (C+rB)] = \det[A\rho^2 - rA\rho - (C+rB)] = 0,$$

hence $r-\rho$ is also a characteristic root.

This general result states that for equilibria of the open-loop capital accumulation game the best stability behaviour that we can expect is saddle point stability, i.e., having two roots with negative real parts and two roots with positive real parts. We rule out that the system has purely imaginary roots by assuming that the profit functions π^i and π^j satisfy certain regularity conditions.

Definition 9.1. An $N \times N$ matrix $F = (f_{ij})$ is said to have a (column) dominant diagonal if there exist positive constants, d_j, $j = 1, 2, \ldots, N$, such that

$$d_j|f_{jj}| > \sum_{i \neq j} d_i|f_{ij}|$$

holds for all $j \in \{1, 2, \ldots, N\}$.

The definition of a dominant diagonal matrix is used in the next assumption that is crucial for saddle point stability. Saddle point stability of system (9.21) and (9.22) or (9.26) means that there exists a stable manifold of dimension equal to the number of state variables so that, if the initial conditions of the costate variables are appropriately chosen, the solution of the canonical system (9.21) and (9.22) starts from the stable manifold and converges to the steady state.

Assumption 9.3 Let

$$\Pi = \begin{pmatrix} \pi^1_{K^1K^1}(\hat{K}^1, \hat{K}^2) & \pi^1_{K^1K^2}(\hat{K}^1, \hat{K}^2) \\ \pi^2_{K^2K^1}(\hat{K}^2, \hat{K}^1) & \pi^2_{K^2K^2}(\hat{K}^2, \hat{K}^1) \end{pmatrix}$$

be a matrix of second order partial derivatives of the profit functions. Either the matrix Π or its transpose has a dominant diagonal.

The assumption of a dominant diagonal of Π has a very intuitive interpretation. In the case of a column dominant diagonal it implies that a change of the capital stock of firm i has a larger impact on its own marginal profit than on the sum of the marginal profits of the rival firms. In the case of a row dominant diagonal the assumption implies that a change in the capital stock of firm i has a larger impact on its own marginal profit than a change in the capital stock of the rival.

With the use of assumption 9.3 we are now able to rule out purely imaginary roots of (9.27) if the discount rate $r = 0$. Assume that we find a purely imaginary root $\rho = i\beta$. This root must satisfy the characteristic equation $\det[A\rho^2 - C] = 0$ or, equivalently,

$$\det\left[-A\beta^2 - C\right] = 0. \tag{9.28}$$

We know that the matrix A is a diagonal matrix and that the signs of its elements are the same as those of C. More specifically, these signs are negative. Thus, with the use of assumption 9.3 the matrix $A\beta^2 + C$ has a dominant diagonal. From linear algebra we know that a matrix with a dominant diagonal is nonsingular. Hence $A\beta^2 + C$ is nonsingular, which contradicts equation (9.28).

The above results imply the following stability properties of the open-loop equilibria of the capital accumulation game with adjustment costs. We have found that, with zero discounting, there are no purely imaginary roots of (9.27) and that for every root ρ it holds that $-\rho$ is also a root. Thus, there must be as many roots with a positive real part as there are roots with a negative real part. By continuity of the eigenvalues of a matrix this property carries over to the situation where the discount rate r is positive but sufficiently small. Hence, for sufficiently small discount rates, the game has the saddle point property in the sense that one can always choose the initial conditions for the adjoint variables such that the equilibrium trajectory starts on the stable manifold and, consequently, converges to the steady state.

9.4 Games without adjustment costs

So far we have looked only at games in which firms face internal adjustment costs and investment is (partially) reversible. A large body of literature deals with the case of zero adjustment costs and irreversible investment. In that case the dynamic game has the objective functionals

$$J^i = \int_0^\infty e^{-rt}[\pi^i(K^i(t), K^j(t)) - I^i(t)]\, dt \tag{9.29}$$

and the state equations

$$\dot{K}^i(t) = I^i(t), \tag{9.30}$$

where we assume that the price of one unit of investment is normalized to $p = 1$. Moreover, investment is constrained by $0 \leq I^i(t) \leq \bar{I}$. The game given by (9.29)–(9.30) has two important features that make it quite distinct from the other games discussed in this chapter. First, investment in this model is irreversible. This implies that firms face a lock-in effect. A capital stock that was built up in the past cannot be reduced.[4] Second, the game is linear in the control variables $I^i(t)$, which changes the way we derive the equilibrium solution.

The linearity property of the game allows us to substitute (9.30) into the objective function (9.29) to get

$$J^i = \int_0^\infty e^{-rt}[\pi^i(K^i(t), K^j(t)) - \dot{K}^i(t)]\,dt.$$

Integrating by parts yields

$$J^i = \int_0^\infty e^{-rt}[\pi^i(K^i(t), K^j(t)) - rK^i(t)]\,dt + K^i(0),$$

which has to be maximized with respect to the choice of an investment strategy.

Before deriving an open-loop equilibrium we introduce an additional assumption. We assume that the firms are characterized by differential entry times, which means that firm i starts investing at time $t_i > 0$ and has no capital stock before that time. Thus the discounted streams of profits are

$$J^i = \int_{t_i}^\infty e^{-rt}[\pi^i(K^i(t), K^j(t)) - rK^i(t)]\,dt.$$

Now let us focus upon one of the two firms (firm i) and examine its optimal investment policy, taking the investment path of the rival firm j as given. For a given investment path the capital stock of the rival firm j at time t is $\tilde{K}^j(t)$. Since investment is irreversible, $\tilde{K}^j(\cdot)$ must be nondecreasing. Because of the assumptions that $\pi^i_{K^iK^j}(K^i, K^j) < 0$ and that $\pi^i(K^i, K^j)$ is strictly concave in the capital stock K^i, the function

$$K^i \mapsto \pi^i(K^i, \tilde{K}^j(t)) - rK^i$$

has a unique maximizer for each t, and this maximizer is a nonincreasing function of t. (Of course, the maximizer is constant on intervals on which firm j does not invest.) Let us denote this maximizer by $\hat{K}^i(t)$. A possible path of $\hat{K}^i(t)$ is depicted in figure 9.2.

Given the investment constraint and the investment path of the rival firm $\tilde{K}^j(t)$ we define the maximum expansion path, $K^{i*}(t)$, by $\dot{K}^{i*}(t) = \bar{I}$ and $K^{i*}(t_i) = 0$. On this path investment equals its upper bound. The

[4]Note that we assume $\delta^i = 0$ and $I^i(t) \geq 0$. Hence, if a firm has overinvested at some point in time it needs to bear the corresponding consequences in all future periods.

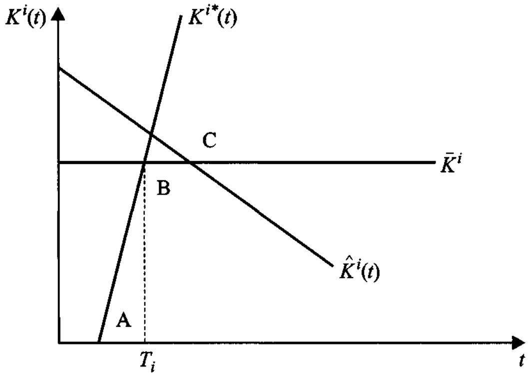

Figure 9.2 Maximum expansion path

maximum expansion path is depicted in figure 9.2. Following this path, however, is not the optimal strategy of firm i given the behaviour of its rival. Since the objective of firm i is to maximize the present value of future profits, J^i, the optimal policy is easy to establish. The integrand of $\pi^i(K^i, \tilde{K}^j(t)) - rK^i$ is increasing in K^i for $K^i \in [0, \hat{K}^i(t))$ and falling otherwise. Since investment is irreversible, firm i will stop investing once a level $K^i(t) \geq \hat{K}^i(t)$ is reached. Suppose that the optimal investment path of firm i crosses $\hat{K}^i(t)$ at point C on figure 9.2. The corresponding capital stock is $\bar{K}^i$. The investment problem of firm i then reduces to the following: what is the best way to get to point C and remain at the capital stock $\bar{K}^i$ forever after? The answer is to invest as rapidly as possible until $\bar{K}^i$ is reached and then stop investing. This results in the path denoted by ABC in figure 9.2. Any alternative path that reaches point C must lie below ABC, but this implies that $\pi^i(K^i(t), \tilde{K}^j(t)) - rK^i(t)$ is less than or equal to its level along ABC. Hence the most rapid approach path becomes the optimal strategy given a capital stock of the rival. This implies, however, that the optimal strategy of firm i is fully characterized by the level $\bar{K}^i$ or by the stopping time T_i at which point B is reached. Writing the objective functional of player i in terms of T_i and differentiating with respect to that parameter we obtain the first order condition

$$\int_{T_i}^{\infty} e^{-rt}[\pi^i_{K^i}((T_i - t_i)\bar{I}, K^j(t)) - r]\,dt = 0.$$

Note that $(T_i - t_i)\bar{I} = \bar{K}^i$. The above maximization condition implies that the firm should stop investing when the present value of marginal profit is 0. Because of $\pi^i_{K^iK^j}(K^i, K^j) < 0$ marginal profit $\pi^i_{K^i}(\bar{K}^i, K^j(t)) - r$ is declining or constant with respect to t depending on whether the rival firm is still investing or not. If firm i stops investing before firm j does, then firm i must stop when $\pi^i_{K^i}(K^i(t), K^j(t)) - r$ is still positive (because it anticipates a further decline of marginal profit). Hence the optimal strategy of player i can be written as

$$I^{i*}(t) = \begin{cases} 0 & \text{if } t < t_i, \\ \bar{I} & \text{if } t_i \leq t \leq T_i, \\ 0 & \text{if } t \geq T_i. \end{cases}$$

This means that a bang-bang strategy (i.e., a switch between the lower and upper bound of investment) is the optimal policy. It remains to determine simultaneously the capital stocks $\bar{K}^i$ and $\bar{K}^j$ at which the firms stop investing. In general, these levels do not satisfy the conditions

$$\pi^i_{K^i}(\bar{K}^i, \bar{K}^j) = r$$

for $i, j \in \{1, 2\}$, $i \neq j$, so that they may very well be different from the equilibrium levels in the static Cournot model.

9.5 Knowledge as a public good

The models of the preceding sections have the property that each firm invests in a private, physical capital stock. Therefore each firm has a capital stock and an accumulation equation. In this section we take a very different perspective. We look at knowledge accumulation, whereby knowledge is modelled as a pure public good; this means that every individual has access to all knowledge in the economy.[5] This implies that there is only a single stock of capital, $K(t)$, into which two or more individuals invest. In the case of two individuals, each choosing an investment level of $I^i(t)$ at time t, the accumulation equation becomes

$$\dot{K}(t) = I^i(t) + I^j(t) - \delta K(t), \quad i \neq j,$$

where $\delta \geq 0$ is a constant rate of depreciation and $K(0) = K_0$ is a given initial stock of capital. Each individual who invests in the public capital stock faces costs of investment which are given by

$$C^i(I^i(t)) = \rho I^i(t) + \frac{1}{2}[I^i(t)]^2.$$

Each individual, investing or not, benefits from the existing level of the capital stock and derives instantaneous revenues equal to

$$\pi^i(K(t)) = K(t)\left[a^i - K(t)\right].$$

[5]See also section 11.3, where we deal with an advertising model of similar structure.

These assumptions imply that the accumulation game has the objective functionals

$$J^i = \int_0^\infty e^{-rt}\{K(t)[a^i - K(t)] - \rho I^i(t) - (1/2)[I^i(t)]^2\}\,dt$$

and the state equation

$$\dot{K}(t) = I^i(t) + I^j(t) - \delta K(t), \quad i \neq j. \tag{9.31}$$

We are interested in both the open-loop Nash equilibrium and the Markov perfect Nash equilibria of this game and how these decentralized solutions (i.e., the noncooperative outcomes) relate to the collusive solution (i.e., the outcome of joint maximization).

We start by deriving a Markov perfect Nash equilibrium for the game. Since the game structure is of the linear quadratic type we can apply the methods discussed in chapter 7. The HJB equations are given by

$$\begin{aligned} rV^i(K) = \max\{&K(a^i - K) - \rho I^i - (1/2)(I^i)^2 \\ &+ V^i_K(K)[I^i + \phi^j(K) - \delta K] \mid I^i \in \mathbb{R}\}. \end{aligned} \tag{9.32}$$

Our routine in solving differential games suggests that we conjecture quadratic value functions of the type

$$V^i(K) = (\alpha/2)K^2 + \beta^i K + \gamma^i.$$

Maximizing the right-hand side of (9.32) with respect to I^i and making use of the quadratic value functions results in linear investment strategies given by

$$I^i = \phi^i(K) = -\rho + \alpha K + \beta^i. \tag{9.33}$$

These linear investment strategies constitute a Markov perfect Nash equilibrium if the constants α, β^i, and γ^i satisfy the equations

$$\frac{r}{2}\alpha = -1 + \frac{3}{2}\alpha^2 - \delta\alpha, \tag{9.34}$$

$$\beta^i r = a^i + 2\alpha\beta^i - 2\alpha\rho - \beta^i\delta + \alpha\beta^j, \tag{9.35}$$

$$\gamma^i r = \frac{\rho^2}{2} + \frac{1}{2}(\beta^i)^2 - 2\beta^i\rho + \beta^i\beta^j. \tag{9.36}$$

We see that (9.34) is a quadratic equation in α and (9.35) is a system of linear equations in (β^i, β^j) that can be solved easily. The roots of (9.34) are given by

$$\alpha = \frac{r + 2\delta}{6} \pm \sqrt{\left(\frac{r + 2\delta}{6}\right)^2 + \frac{2}{3}}. \tag{9.37}$$

With the linear decision rules (9.33), the state dynamics of the game become

$$\dot{K}(t) = -2\rho + (2\alpha - \delta)K(t) + \beta^1 + \beta^2, \tag{9.38}$$

which has a globally asymptotically stable steady state if $2\alpha - \delta < 0$. This inequality is satisfied if we choose the negative root of (9.37). Solving (9.35) yields

$$\beta^i = (1/\Delta_c)[2\alpha^2\rho + (r + \delta)(a^i - 2\alpha\rho) - \alpha(2a^i - a^j)],$$

where $\Delta_c = (r + \delta - 2\alpha)^2 - \alpha^2$.

The unique steady state level of the accumulation game is given by

$$\hat{K}_c = -\frac{\beta^1 + \beta^2 - 2\rho}{2\alpha - \delta}$$

and the capital stock in the Markov perfect Nash equilibrium evolves according to $K_c(t) = (K_0 - \hat{K}_c)e^{(2\alpha-\delta)t} + \hat{K}_c$, which shows that the capital stock converges either from above or from below to the steady state level. The speed of adjustment is given by $2\alpha - \delta$.

The structure of the Markov perfect Nash equilibrium is our point of reference. In a next step we derive an open-loop Nash equilibrium. We formulate the current value Hamiltonian which is given by

$$H^i(K, I^i, \lambda^i, t) = K(a^i - K) - \rho I^i - (1/2)(I^i)^2 + \lambda^i[I^i + I^j(t) - \delta K],$$

where λ^i is the current value adjoint variable of player i. Assuming, for simplicity, the case of symmetric agents (i.e., $a^i = a^j = a$), the costate variables of both players coincide and the equilibrium conditions result in the following differential equation system for the state variable and the common costate variable λ

$$\dot{K}(t) = -2\rho + 2\lambda(t) - \delta K(t), \tag{9.39}$$

$$\dot{\lambda}(t) = (r + \delta)\lambda(t) - a + 2K(t). \tag{9.40}$$

This system, consisting of two linear differential equations, has the saddle point property. This claim can be verified by noting that the trace of the Jacobian matrix of the system is positive and equal to r and that the determinant is negative and given by $\Delta_o = -[\delta(\delta + r) + 4]$. The stable root of the characteristic equation of system (9.39) and (9.40) is given by

$$\mu_o = \frac{r}{2} - \sqrt{\left(\frac{r}{2}\right)^2 - \Delta_o}.$$

The dynamics as well as the long-run equilibrium of the open-loop game are completely determined by a solution of the system (9.39)–(9.40). The steady state values are given by

$$\hat{K}_o = (2/\Delta_o)[\rho(r+\delta) - a]$$

and

$$\hat{\lambda}_o = (1/\Delta_o)(-\delta a - 4\rho).$$

The evolution of the open-loop capital stock is given by $K_o(t) = (K_0 - \hat{K}_o)e^{\mu_o t} + \hat{K}_o$. Again we observe that the transient behaviour of the capital stock is either increasing or decreasing towards the steady state level depending on the initial condition.

Having characterized both an open-loop and a Markov perfect Nash equilibrium, we are now interested in comparing those two and also in comparing them with the collusive solution. Moreover, we want to investigate whether the decentralized solution (i.e., a noncooperative equilibrium) results in a free rider problem. To do so we need to derive the efficient (collusive) solution of the problem and compare it with the two Nash equilibria.

The collusive problem is to maximize

$$\int_0^\infty e^{-rt}\{2K(t)[a - K(t)] - C^1(I^1(t)) - C^2(I^2(t))\}\, dt$$

subject to the state equation (9.31). This is a standard optimal control problem and can be solved using the techniques presented in chapter 3. The current value Hamiltonian is given by

$$H(K, I^1, I^2, \lambda) = 2K(a - K) - C^1(I^1) - C^2(I^2) + \lambda(I^1 + I^2 - \delta K).$$

The maximum conditions are $H_{I^i}(K(t), I^i(t), I^j(t), \lambda(t)) = 0$, which implies $I^i(t) = -\rho + \lambda(t)$. The adjoint equation is

$$\dot{\lambda}(t) = (r+\delta)\lambda(t) - 2a + 4K(t).$$

The canonical equation system is

$$\dot{K}(t) = -2\rho + 2\lambda(t) - \delta K(t),$$
$$\dot{\lambda}(t) = (r+\rho)\lambda(t) - 2a + 4K(t).$$

It is easily checked that this system possesses the saddle point property with the negative root

$$\mu_e = \frac{r}{2} - \sqrt{\left(\frac{r}{2}\right)^2 - \Delta_e},$$

where $\Delta_e = -\delta(r+\delta) - 8$. The unique steady state of the system is given by

$$\hat{K}_e = (2/\Delta_e)[\rho(r+\delta) - 2a]$$

and

$$\hat{\lambda}_e = (2/\Delta_e)(-\delta a - 4\rho).$$

Hence, the evolution of the capital stock in the collusive solution is governed by

$$K_e(t) = (K_0 - \hat{K}_e)e^{\mu_e t} + \hat{K}_e.$$

We are now in a position to compare the collusive solution with the decentralized ones. The outcomes of the open-loop game as well as that of the Markovian game result in a free rider problem, i.e., the steady state equilibrium stocks of capital are lower than in the collusive solution. The free rider problem is more severe in the case of the Markov perfect Nash equilibrium than in the open-loop game, that is, $\hat{K}_o > \hat{K}_c$.

The economic interpretation of these results is straightforward. The free rider problem is an immediate consequence of the public goods aspect of the capital stock. Since both agents derive utility from the stock $K(t)$ they have an incentive not to invest themselves but to benefit from the investment of the other agent. This behaviour is more severe in the case of a Markov perfect Nash equilibrium than in the case of an open-loop game.

9.6 Further reading

The class of capital accumulation games discussed in section 9.1 originally was proposed in Spence [224]. His paper deals with the case of open-loop equilibria under the assumptions of no discounting and a linear investment cost function. The results of Spence [224] (see also Spence [223]) highlight the preemptive role of capital. Fudenberg and Tirole [105] use the same model. They point out that the open-loop equilibrium has many features of a static Cournot outcome and then derive Markov perfect equilibria for this class of games. In particular, in a Markov perfect equilibrium it is possible that a firm that has a headstart in an industry can deter entry (or at least mobility) by over-investment, which causes the rival to invest less (see also Fudenberg and Tirole [103, 104]).

The game model with reversible investment was first formulated by Dockner [46]. In that paper the relationship between the stationary state of a Markov perfect equilibrium of the dynamic game and a conjectural variations equilibrium of the corresponding static game is derived in detail. The solution technique for deriving the Markov perfect equilibrium in the linear quadratic version follows Reynolds [203], where a two-player game is solved using the approach presented in chapter 7. In Reynolds [204] the N-player case is discussed.

The existence and global asymptotic stability of the open-loop equilibrium for a game with strictly convex adjustment costs was dealt with by Fershtman and Muller [92, 93]. They also prove a series of turnpike theorems much in the spirit of the optimal growth literature (see McKenzie [171]).

The general stability analysis of the open-loop game with convex adjustment costs was first carried out by Dockner and Takahashi [63], for the cases both of differential and difference games. General turnpike results in the case of discrete-time formulations have been presented by Dockner and Takahashi [62, 64].

The model of the dynamic provision of a public good was first formulated by Fershtman and Nitzan [94]. They study the case of linear Markov perfect equilibria. Wirl [244] uses the same linear quadratic version of the game and studies nonlinear Markov perfect equilibria.

9.7 Exercises

1. Consider the capital accumulation game of example 9.1 with the state equation

$$\dot{K}^i(t) = I^i(t) - \delta^i K^i(t),$$

where $I^i(t) \geq 0$ and $\delta^i > 0$. Show that there exists a Markov perfect Nash equilibrium that results in a higher steady state level of capital than the steady state in the open-loop equilibrium.

2. Show that the open-loop equilibrium of the capital accumulation game of exercise 1 is fully characterized by a system of four linear differential equations in the state and costate variables. Discuss an explicit solution to this system of equations for the case of a finite time horizon T and no salvage value at T.

3. Consider two firms that produce a homogeneous product in a dynamic Cournot market. The level of output of firm i at time t is $Q^i(t)$. Production of each firm accumulates a stock of pollutants according to the equation

$$\dot{P}(t) = Q^1(t) + Q^2(t) - \delta P(t),$$

where $P(t)$ is the stock of pollution and δ is the natural rate of purification. Firms face an inverse demand function of the linear type $D(Q^1(t) + Q^2(t)) = 1 - Q^1(t) - Q^2(t)$. Each firm faces constant marginal production costs $c^i > 0$ and quadratic costs of pollution $K(P(t)) = (1/2)P(t)^2$. Firms maximize the discounted stream of profits over an infinite planning period. Formulate a differential game for this Cournot market and derive the open-loop equilibrium. Show that the steady state output of the open-loop equilibrium is lower than that of the static Cournot model.

4. Derive a Markov perfect Nash equilibrium for the game of exercise 3 and demonstrate that the steady state equilibrium output level is higher than in the open-loop game, but lower than in the static Cournot game.
5. Consider the game from exercise 3 and assume that one of the two firms is a public firm that maximizes social welfare, where social welfare is defined as the sum of consumer and producer surplus minus the costs of pollution. Moreover, assume that the remaining private firm does not care about pollution. Formulate the differential game model for this scenario and derive the open-loop Nash equilibrium. Show that the equilibrium quantities are higher than in the static Cournot model.

10 Industrial organization and oligopoly games

This chapter deals with two issues in dynamic oligopoly theory and industrial organization. The first concerns a duopolistic market in which producers determine their output rates but the market price does not adjust instantaneously to the price indicated by the demand function (as it is supposed to do in the static Cournot model of chapter 2). The market price is sticky. We consider two firms that play a linear quadratic differential game (cf. chapter 7). The second issue comes from industrial organization and is that of research and development. R&D activities are aimed at developing new technologies, production processes, or products. Related problems concern the diffusion and adoption of innovations and the transfer of new technologies. In all these areas, game theoretic models have been proposed. In this chapter we confine our interest to a class of R&D differential games where the date of successful completion of the innovation by one of the oligopolists is a random variable with a probability distribution that is known to depend on the oligopolists' R&D efforts. First, we analyse a pure R&D game which subsequently is modified to include the extraction of a nonrenewable resource (cf. chapter 12). In this modification, an importer of a nonrenewable resource (e.g., oil) seeks to develop a new technology the output of which can be substituted for imports of the nonrenewable resource.

10.1 Dynamic duopoly with sticky prices

We consider dynamic duopolistic competition in a market for a homogeneous good. A key feature of the problem is that the market price does not adjust instantaneously to the price indicated by the demand function. The evolution of market price over time is a function of the difference between the current market price and the price specified by the demand function for each level of industry output. There is a lag in the market price adjustment: the price is sticky. This is in contrast to the standard

Cournot model where the market price adjusts instantaneously. The scenario is modelled as a differential game where the dynamics describe the evolution of market price over time. The dynamics include the reasonable feature that, when the parameter measuring the speed of price adjustment tends to infinity, price converges impulsively to its value on the demand function.

The model assumptions are as follows. Demand is linear in price, production cost functions are quadratic in output, and the game is played over an infinite interval of time. Denote by $u_i(t) \geq 0$ the output rate of duopolist i ($i = 1, 2$). A (linear) instantaneous inverse demand function is given by $p(t) = a - [u_1(t) + u_2(t)]$, where $p(t)$ is the market price at time t and $a > 0$ is a constant. To model that the market price does not adjust instantaneously to the price indicated by the demand function we let the rate of change of the market price be a function of the difference between current market price and the price indicated by the linear demand function (for any particular aggregate output). Hence

$$\dot{p}(t) = s\{a - [u_1(t) + u_2(t)] - p(t)\}, \; p(0) = p_0, \tag{10.1}$$

where $s \in (0, \infty)$ is the adjustment speed parameter. For simplicity, the duopolists are assumed to have the same quadratic production cost function $C(u_i) = cu_i + (1/2)u_i^2$, where $c \in (0, a)$ is a fixed parameter. The objective functional of duopolist i is given by

$$J^i(u_1, u_2) = \int_0^\infty e^{-rt}[p(t)u_i(t) - cu_i(t) - (1/2)u_i(t)^2]\,dt, \tag{10.2}$$

in which $r > 0$ is the discount rate. Firm i wishes to maximize its objective (10.2) with respect to $u_i(\cdot)$, subject to (10.1) and $u_i(t) \geq 0$.

The implication of sticky prices can be seen by solving (10.1) with respect to $p(t)$ and substituting the resulting expression into (10.2) to get

$$J^i(u_1, u_2) = \int_0^\infty e^{-rt}\{[a - u_1(t) - u_2(t)]u_i(t) - \dot{p}(t)u_i(t)/s - cu_i(t) - (1/2)u_i(t)^2\}\,dt.$$

This shows that essentially each firm faces a standard linear demand function, but adjustments along this line are retarded (whenever s is finite). For s going to infinity, the market price adjusts instantaneously along the demand function.

The differential game defined by (10.1)–(10.2) is linear-quadratic and admits both open-loop and Markov perfect equilibria under simultaneous play.[1] Thanks to the linear quadratic structure, explicit analytical expressions can be obtained for both kinds of equilibrium strategies. To determine an open-loop Nash equilibrium $(u_1(\cdot), u_2(\cdot))$, define the current-value Hamiltonians

$$H^i(p, u_i, \lambda_i, t) = pu_i - cu_i - \frac{1}{2}u_i^2 + \lambda_i s[a - u_i - u_j(t) - p].$$

At any t for which the output rate u_i in an optimal solution is positive, it is necessary for maximizing the Hamiltonian that

$$p(t) - \lambda_i(t)s = c + u_i(t) \tag{10.3}$$

and the costate equations

$$\dot{\lambda}_i(t) = (r + s)\lambda_i(t) - u_i(t) \tag{10.4}$$

must be satisfied. Equation (10.3) states that each duopolist determines his instantaneous production rate according to the rule that marginal revenue equals marginal cost. In this dynamic setting, marginal revenue consists of two terms. The instantaneous marginal revenue is $p(t)$, and from this revenue one subtracts the product of the costate and the adjustment speed parameter. This product represents the long-run effect of a marginal change in the current production rate since the costate has the interpretation of a shadow price (an imputed value) of the state variable. The right-hand side of (10.3) is the instantaneous marginal cost.

Differentiation with respect to time in (10.3) and using (10.1), (10.3), and (10.4) yields

$$\dot{u}_i(t) = s[a - u_j(t) - p(t)] - (r + s)[p(t) - c - u_i(t)]. \tag{10.5}$$

Equations (10.1) and (10.5) provide a system of three ordinary differential equations, defined in the feasible region of the (p, u_1, u_2) space.

In view of the symmetric structure of the game we look for a symmetric equilibrium in which $u_1(t) = u_2(t) = u(t)$. At a steady state market price p_{SS} we must have $\dot{p}(t) = \dot{u}(t) = 0$ and therefore

$$p_{SS} = \frac{a(2s + r) + 2c(r + s)}{3r + 4s}. \tag{10.6}$$

This price equals the static Cournot duopoly price $(a + c)/2$ if the discount rate r equals zero, or if the adjustment speed parameter s tends to infinity. On the other hand, if the discount rate tends to infinity, or the adjustment speed parameter tends to zero, the steady state price p_{SS}

[1]Strictly speaking, the linear quadratic structure is lost if control nonnegativity constraints are active.

equals the perfect competition price $(a+2c)/3$. The latter occurs when both firms set their output rates at the point where marginal cost equals price.

To obtain an equilibrium state trajectory, differentiate with respect to time in the dynamics (10.1) and substitute into this expression the quadruple $(\lambda(t), \dot{\lambda}(t), u(t), \dot{u}(t))$ given by the necessary conditions (and use (10.1)). The result is the following linear second order differential equation for the state $p(t)$:

$$\ddot{p}(t) + A\dot{p}(t) + Bp(t) = R,$$

where $A = s - r$, $B = -s^2 - 3s(s+r)$, and $R = -[s^2 a + s(2c+a)(s+r)]$. This equation has an infinity of solutions that satisfy the initial condition $p(0) = p_0$. Among these we look for a solution which remains bounded and restrict the parameters so as to obtain such a solution. Using standard procedures for solving a linear second order differential equation yields

$$p^*(t) = p_{SS} + (p_0 - p_{SS})e^{-kt},$$

in which k is a positive constant and the steady state market price p_{SS} is given by (10.6). The price trajectory $p^*(t)$ converges to the steady state level p_{SS} for any value of the initial price p_0. The steady state is globally asymptotically stable.

At this stage of the analysis we still cannot conclude that by $p^*(t)$ and the conditions stated in (10.3) and (10.4) we have an open-loop Nash equilibrium of the duopoly game. What is needed is to verify that the candidate output rate $u(t)$ remains positive for all t; otherwise the solution must be modified accordingly. We also need to verify a sufficient optimality condition, that is, show that the solution is indeed optimal and not just a candidate. The sufficient conditions of chapter 3 are not applicable since in the model at hand we cannot verify the concavity requirements. (See section 10.4 for a remark on the issue of sufficiency.)

As mentioned in chapter 4, when one analyses an autonomous game defined on an infinite time interval, one often restricts attention to stationary Nash equilibria (not forgetting that the game may have nonstationary equilibria as well). A stationary open-loop equilibrium will be constant with respect to time. Inserting p_{SS} from (10.6) into (10.5) and equating the left-hand side of (10.5) to zero yields

$$u_{SS} = \frac{(a-c)(s+r)}{3r+4s}.$$

Then, by the assumption $a > c$, u_{SS} is strictly positive.

We proceed by deriving a symmetric and stationary Markov perfect equilibrium. The motivation for confining the analysis to a symmetric and stationary equilibrium is the same as in the open-loop case (symmetric game structure and an autonomous problem over an infinite horizon). Denote the equilibrium strategies again by $u_1(\cdot)$ and $u_2(\cdot)$ but note that, in contrast to the open-loop case, these functions are now defined on the state space instead of the time domain. The HJB equations are given by

$$rV^i(p) = \max\left\{(p-c)u_i - \frac{1}{2}u_i^2 + \frac{dV^i(p)}{dp}s\{a - [u_i + u_j(p)] - p\} \,\middle|\, u_i \geq 0\right\}, \tag{10.7}$$

where $i, j \in \{1, 2\}$ and $i \neq j$. Because of symmetry and stationarity, the value functions are identical and depend on p only. We can therefore omit the player indices. Performing the maximization indicated in the HJB equation yields a unique Markovian output strategy

$$\begin{array}{lll} \text{(A):} & u(p) = 0 & \text{if } p - c - sV'(p) \leq 0, \\ \text{(B):} & u(p) = p - c - sV'(p) & \text{if } p - c - sV'(p) > 0. \end{array} \tag{10.8}$$

Substituting from (10.8) into the term in curly brackets on the right-hand side of the HJB equation yields the following differential equations for the value function:

$$\begin{aligned} &\text{(A): } rV(p) = sV'(p)(a-p), \\ &\text{(B): } rV(p) = (p-c)[p-c-sV'(p)] - (1/2)[p-c-sV'(p)]^2 \\ &\qquad\qquad + V'(p)s\{a - p - 2[p-c-sV'(p)]\}. \end{aligned} \tag{10.9}$$

Starting with case (B), we conjecture that the quadratic value function

$$V(p) = \frac{K}{2}p^2 - Ep + G \tag{10.10}$$

solves the HJB equation. Here, K, E, and G are constants to be determined. Substituting the value function given by (10.10), as well as its first order derivative, into the HJB equation in case (B) provides a set of conditions that must be satisfied by the three constants K, E, and G.[2] Determining these constants is a straightforward algebraic exercise which leads to

[2]For more details of this procedure, see chapter 7.

$$K = \frac{r + 6s - \sqrt{(r + 6s)^2 - 12s^2}}{6s^2} > 0,$$
$$E = \frac{-asK + c - 2scK}{r - 3s^2K + 3s},$$
$$G = \frac{c^2 + 3s^2E^2 - 2sE(2c + a)}{2r}.$$

Doing this exercise reveals that the equation for K is a quadratic equation. To verify boundedness of the value function (cf. theorem 4.4) we need convergence of the state trajectory, and for this purpose we choose the smaller root of the second degree equation. Doing this implies $K < 1/(3s)$.

Using (10.8) and (10.10) for case (B) yields the symmetric output strategies

$$u(p) = (1 - sK)p + sE - c. \tag{10.11}$$

Recall that the results for case (B) are valid only if output is positive, that is, if $(1 - sK)p + sE - c > 0$ which is equivalent to

$$p \geq \hat{p} := \frac{c - sE}{1 - sK}.$$

To ensure that the optimal output given by (10.8) remains positive we need to show that $\hat{p}$ is positive. As mentioned above, it holds that $K < 1/(3s)$ and hence $1 - sK > 0$. It is readily verified that $c - sE > 0$ and hence $\hat{p}$ is strictly positive. Define $\hat{V} = V(\hat{p})$.

The equilibrium price trajectory in case (B) is given by

$$p(t) = p_{SS} + (p_0 - p_{SS})e^{Dt},$$

where $D = s[2(sK - 1) - 1] < 0$. From (10.1) and (10.11) we obtain the steady state price for case (B):

$$p_{SS} = \frac{a + 2c - 2sE}{3 - 2sK}. \tag{10.12}$$

We turn to the determination of the value function in case (A) where optimal outputs are zero. (The intuition is that the market price is too low to warrant any production.) Solving the HJB equation in case (A) in (10.9) yields

$$V(p) = D_0(a - p)^{-r/s},$$

where D_0 is a constant of integration. Continuity of the value function implies the boundary condition $\hat{V} = D_0(a - \hat{p})^{-r/s}$. Hence, for case (A) we obtain the value function

$$V(p) = \hat{V}\left(\frac{a-\hat{p}}{a-p}\right)^{r/s} \quad \text{for } p \leq \hat{p}.$$

Recall that we assumed $a > c$. Exploiting this assumption one can show that $a > \hat{p}$ and, since the Nash equilibrium price is increasing over time, this price remains in the interval $[0, \hat{p}]$ only for a finite period of time. As of some instant of time, the price will exceed the level $\hat{p}$ and the strategy of zero output is switched to the positive output strategy. This means that the only possible steady state is the one given by (10.12). In this steady state production rates are positive.

In summary, the value function which has been constructed for cases (A) and (B) is continuous and continuously differentiable. To see the latter, note that for $p < \hat{p}$ we have $V'(p) = (r/s)V(p)/(a-p)$ and for $p > \hat{p}$ we have $V'(p) = Kp - E$. What is left to prove is that $r\hat{V}/[s(a-\hat{p})] = \lim_{p \to \hat{p}}(Kp - E)$. The Markovian Nash equilibrium that has been derived is a Markov perfect Nash equilibrium. It follows from the last remark of chapter 4 that any stationary Markovian Nash equilibrium is subgame perfect, provided that the equilibrium is independent of the initial state. We have seen that the latter is true.

From an economic point of view it is worth noticing that, as the adjustment speed parameter s tends to infinity, the steady state price in case (B) (cf. equation (10.12)) is smaller than the static Cournot price $(a+c)/2$. The static Cournot price is the exact limit in the open-loop case. The difference arises from the players' ability to precommit to their output strategies.

10.2 A game of R&D competition

The problem in this section concerns the area of research and development. The game theoretic approach to R&D views innovations (new products, new production processes, new technologies, new services) as being developed in competitive environments, not in Robinson Crusoe-like isolation. Sometimes one thinks of the R&D activities of competing firms as a race of being the first to reach a technological breakthrough. The efforts and resources that are devoted to R&D in a firm influence the probability that the firm is successful and completes the innovation before its rivals (wins the race). Once a firm has won the race, a standard assumption is that the winning firm acquires a monopolistic position, that is, rival firms are kept out of the market (for instance, by patent protection). Thus, a basic idea behind this view on innovative behaviour is that it is the profit from being first that spurs entrepreneurial activity and brings forward a stream of new products, processes, and technologies. Dynamic game models of R&D often rest on the following three assumptions. First, no firm knows in advance exactly how much it must spend to develop the innovation. Second, there are several possible paths to the successful development of the innovation. Third, R&D activities are costly but lead to the accumulation of know-how, which influences positively the probability of winning the race.

Consider the following situation. There are N firms competing for the completion of a particular R&D project (e.g., a new electronic device). The time it will take for firm $i \in \{1, 2, \ldots, N\}$ to complete the project is a random variable τ_i having a probability distribution $F_i(t) = \text{Prob}(\tau_i \leq t)$. Assuming no spillovers of knowledge among the firms it is plausible to suppose that the random variables τ_i are stochastically independent. Denote the time instant at which one of the firms makes the innovation by $\tau = \min\{\tau_i \mid i = 1, 2, \ldots, N\}$. The firm, number k say, which has $\tau_k = \tau$ is called the innovator. By stochastic independence it holds that

$$\text{Prob}(\tau \leq t) = 1 - \prod_{i=1}^{N}[1 - F_i(t)]. \tag{10.13}$$

Let $u_i(t) \geq 0$ denote the rate of R&D efforts that firm i devotes to its project. The hazard rate corresponding to the distribution $F_i(t)$ is assumed to be proportional to effort $u_i(t)$. Then we have

$$\dot{F}_i(t) = \lambda u_i(t)[1 - F_i(t)], \qquad F_i(0) = 0 \tag{10.14}$$

in which λ is a positive constant. The hazard rate can be thought of as the (conditional) probability that a breakthrough will be made at time t, given that this has not happened before time t.

Define P_I as the present value (i.e., the value as of the initial instant of time) of the net benefits that will accrue to the innovator at the instant t of completion. This present value is assumed to be constant and hence it is independent of the instant of completion. Denote by P_F the present value of the net benefits that will accrue at time t to any competitor other than the innovator. This present value is also assumed constant. In particular, assuming that P_I is constant may be critical. The assumption means that the present value of the innovator's benefits is the same regardless of the date of success. Strategically, the assumption eliminates a possible influence of growth or decline in the value of the benefits upon the firm's R&D efforts and serves to focus on the possibility of a rival firm's prior success. The assumptions of constant P_I and P_F can be relaxed by, for example, multiplying each of the two constants by e^{gt}, where $g > 0$ ($g < 0$) is the rate of growth (decline) of the respective benefits. (See section 10.4 for more details.)

Assume $P_I > P_F$ and let the game be played over a fixed and finite horizon T. The costs of R&D efforts are quadratic in the effort rate. The objective functional of firm i is its expected present value profit, which is given by

$$\int_0^T \left\{ P_I \dot{F}_i(t) \prod_{j \neq i} [1 - F_j(t)] + P_F \sum_{j \neq i} \dot{F}_j(t) \prod_{k \neq j} [1 - F_k(t)] \right. \\ \left. - \frac{e^{-rt}}{2} u_i(t)^2 \prod_{j=1}^{N} [1 - F_j(t)] \right\} dt. \tag{10.15}$$

This profit consists of three terms. The first reflects the fact that firm i's value of net benefits (in present value terms) is equal to P_I if this firm succeeds in becoming the innovator. The second term shows that firm i's value of net benefits (in present value terms) is equal to P_F if this firm loses the innovation race. The third term represents the present value of the cost of R&D effort at time t. All R&D activities on the project are assumed to terminate at the instant of time where a firm makes the innovation. All three monetary components are weighted by their corresponding probabilities. To simplify (10.15), the right-hand side of (10.14) can be substituted in the two first terms in the integrand of (10.15). This yields the payoff

$$\int_0^T \left[\lambda P_I u_i(t) + \lambda P_F \sum_{j \neq i} u_j(t) - \frac{e^{-rt}}{2} u_i(t)^2 \right] \prod_{j=1}^{N} [1 - F_j(t)] \, dt. \tag{10.16}$$

By (10.14) and (10.16) we have defined a deterministic N-player differential game with state variables $F_1, F_2, \ldots, F_N$ and controls $u_1, u_2, \ldots, u_N$. This game is a deterministic one since the stochastic variable τ has been eliminated by using the expected profits as the objectives, cf. (10.16).

Originally (that is, before using the expected payoff values), the game is a stochastic differential game and belongs to the class of piecewise deterministic games discussed in chapter 8. In the terminology of chapter 8, the innovation game has $N + 1$ modes. We may think of the mode of the system as an additional state variable which randomly assumes one of its $N + 1$ values. The system is in mode 0 before the innovation is made and, if firm i succeeds in making the innovation, the system switches to mode $i \in \{1, 2, \ldots N\}$. Thus, in the innovation game there is at most one switch in the system mode, and this switch is from 0 to i. After a switch the system stays in mode i for ever. The switching time is τ, which is a random variable with the probability distribution given by (10.13). Dividing by the survival probability $1 - F_i(t)$ on both sides of the equality sign in (10.14) yields the hazard rate $\dot{F}_i(t)/[1 - F_i(t)]$. The hazard rate is the switching rate in the terminology of chapter 8. From (10.14) we see that the hazard rate is only a function of the effort rate $u_i(t)$ of firm i. In general, the hazard rate may depend on the controls of all firms, the state vector, and time (see section 8.1).

The deterministic differential game given by (10.14) and (10.16) is a linear state game (cf. section 7.2). To analyse the game it is convenient to introduce the state transformation

$$-\ln(1 - F_i(t)) = \lambda z_i(t). \tag{10.17}$$

Differentiation in (10.17) with respect to time and using (10.14) yields the state equations for the transformed game

$$\dot{z}_i(t) = u_i(t)\,,\ z_i(0) = 0 \tag{10.18}$$

and the corresponding payoffs are given by

$$J^i = \int_0^T \exp\left(-\lambda \sum_{j=1}^{N} z_j(t)\right)\left[\lambda P_L u_i(t) + \lambda P_F \sum_{j \neq i} u_j(t) - \frac{e^{-rt}}{2} u_i(t)^2\right] dt. \tag{10.19}$$

Now the game has been transformed into an exponential game (cf. section 7.3).

The exponential game defined by (10.18) and (10.19) can be given the following interpretation. Let the control $u_i(t)$ represent firm i's rate of acquisition of know-how at time t. Then, by (10.18), the state variable $z_i(t)$ is firm i's accumulated know-how by time t. It is a stock of know-how (human capital) in the sense of chapter 9.

We look for a Markovian Nash equilibrium in the innovation game given by (10.18)–(10.19). In a Markovian Nash equilibrium, the assumption is that all firms observe and base their strategies upon the N-dimensional state vector $(z_1, z_2, \ldots, z_N)$. Recall that we can interpret these state variables as the firms' respective stocks of know-how, having been acquired through their R&D efforts. However, although we suppose that the firms know the full state vector $(z_1, z_2, \ldots, z_N)$, only a function of it, namely the one-dimensional state variable

$$y(t) = \exp\left(-\lambda \sum_{i=1}^{N} z_i(t)\right) \tag{10.20}$$

is payoff-relevant, cf. (10.19). The criterion of payoff-relevance states that players condition their actions only upon variables that influence their payoffs. In the R&D game this suggests that the players would not condition their strategies on $(z_1, z_2, \ldots, z_N)$ but rather on $y(t)$. Note that the state variable $y(t)$ represents the aggregate (i.e., industry-wide) stock of know-how. Also note that if the vector $(z_1, z_2, \ldots, z_N)$ is known, y can be calculated. But, as mentioned, we only need to assume that players know $y(t)$ for any t. Differentiation of (10.20) with respect to time provides a single state equation

$$\dot{y}(t) = -\lambda y(t) \sum_{i=1}^{N} u_i(t), \qquad y(0) = 1, \tag{10.21}$$

which is the one employed in the sequel.

We look for a Nash equilibrium with open-loop effort strategies. Define for $i = 1, 2, \ldots, N$ the present-value Hamiltonians

$$H^i(y, u_i, \mu_i, t) = y\left[\lambda P_I u_i + \lambda P_F \sum_{j \neq i} u_j(t) - \frac{e^{-rt}}{2} u_i^2\right] - \mu_i \lambda y\left[u_i + \sum_{j \neq i}^{N} u_j(t)\right],$$

in which μ_i are present-value costate variables.[3] Assuming that the equilibrium effort rates are strictly positive, the necessary conditions for Hamiltonian maximization provide the candidate strategies

$$u_i(t) = \lambda e^{rt}[P_I - \mu_i(t)], \tag{10.22}$$

where we note that a candidate effort rate depends only on the costate, and not on the state y. The costates must satisfy the differential equations

$$\dot{\mu}_i(t) = -\frac{\partial H^i(y(t), u_i(t), \mu_i(t), t)}{\partial y}, \quad \mu_i(T) = 0. \tag{10.23}$$

The state equation (10.21) and the first order necessary conditions in (10.22) and (10.23) are $2N + 1$ equations in the same number of variables, namely, y, u_i, μ_i $(i = 1, 2, \ldots, N)$. Substituting the candidate strategies from (10.22) into the right-hand sides of (10.21) and (10.23) yields a two-point boundary value problem in $N + 1$ variables y, μ_i, $(i = 1, 2, \ldots, N)$. As indicated in (10.21) and (10.23), we have $N + 1$ boundary conditions at our disposal. The solution of this boundary value problem is facilitated by the fact that the state variable y does not appear on the right-hand side of (10.23). This is true because the Hamiltonian H^i is linear in y. The solution of the costate equation (10.23) then can be obtained independently of the solution of the state equation (10.21). Moreover, the candidate effort rates in (10.22) depend on neither the state variable nor the initial condition of (10.21). Referring to chapter 4, we can conclude that the open-loop equilibrium is Markov perfect. This result is due to the special, exponential structure of the game. The reader should consult chapter 7 for a general exposition of exponential games.

[3]See the final paragraph of section 3.3 for the distinction between present-value and current-value Hamiltonians.

To solve the boundary value problem we start by supposing that

$$u_i(t) = -b(t)\lambda e^{rt}. \tag{10.24}$$

Using (10.22) yields an expression for the costate

$$\mu_i(t) = P_I + b(t) \tag{10.25}$$

and substituting our conjectured solution (10.24) into (10.21) provides the state equation

$$\dot{y}(t) = \lambda^2 N b(t) y(t) e^{rt}, \qquad y(0) = 1.$$

Now, for the conjectured solution to hold, the function b must satisfy the differential equation

$$-\frac{\lambda^2 e^{rt}}{2}\left[(2N-1)b(t)^2 + 2b(t)(1-N)(P_F - P_I)\right] = \dot{b}(t) \tag{10.26}$$

and the terminal condition $b(T) = -P_I$. The differential equation (10.26) is a Riccati equation, the solution $b(t)$ of which can be found by the transformation $g(t) = -1/b(t)$.

Substituting the solution of the Riccati equation into (10.24) yields the (candidate) Markovian strategies. Not surprisingly, these candidates are symmetric: all firms use identical effort strategies. Straightforward calculations show that they are given by

$$u(t) = \frac{2\lambda P_I(P_I - P_F)(N-1)e^{rt}}{(2N-1)P_I - [P_I + 2(N-1)P_F]e^{m(t)}}$$

in which

$$m(t) = \frac{1}{r}(P_I - P_F)(N-1)\lambda^2(e^{rt} - e^{rT}).$$

Note that $m(t) \leq 0$ holds.

Next, we have to verify that the candidates constitute an equilibrium by, for instance, assessing concavity of the maximized Hamiltonians. This is trivial in our case because the Hamiltonians are linear in the state y and the candidate strategies do not depend on y, which allows us to conclude that the maximized Hamiltonians are concave in the state variable y.

10.3 A game of R&D and resource extraction

This game combines the previous section's problem of determining optimal R&D efforts with the extraction of an exhaustible resource (see chapter 12 for details of resource games). As in the R&D example of section 10.2, we rely on the hazard rate formalism.

The differential game is played between a country which exports a nonrenewable resource and an importing country which seeks to invent a substitute technology in order to become less dependent (in fact independent) of the imports of the resource. The time of the importing country's innovation is uncertain but can be affected by the country's R&D efforts. A key issue in the problem is the fact that the resource exporter must take into account the incentives of the importing country to try to develop a backstop technology which can be substituted for the depletable imported resource. The resource extractor is assumed to be a monopolist supplier of the resource. The strategic interdependence between the two countries lies in the fact that the producer's extraction policy has an impact on the importer's innovation efforts, and vice versa.

Denote the resource-extracting country by P and the importing country (or a homogeneous group of importing countries) by C. If C did not have the possibility of inventing a substitute technology, an optimal depletion policy for P would follow the Hotelling rule of exhaustible resource extraction: marginal revenues should grow over time at a rate equal to the interest rate until the stock is depleted. The Hotelling rule is an important result in the theory of exhaustible resources and is discussed in chapter 12.

Countries P and C seek to maximize their respective national welfare functions. The monopolist P wishes to choose an extraction policy over an infinite time horizon. Denote the extraction rate at time t by $q(t)$, being constrained by $q(t) \in [0, \bar{q}]$, where $\bar{q}$ denotes a fixed upper bound on P's extraction rate. Denote the stock remaining at time t by $s(t)$. The resource stock evolves according to the differential equation (cf. chapter 12)

$$\dot{s}(t) = -q(t)\,, \; s(0) = s_0 > 0. \tag{10.27}$$

The unit extraction cost of country P is constant and equals $c > 0$. Country C is assumed to be the sole importer of P's resource. Country C's inverse demand function for the resource does not change over time and is denoted by $p(q)$. This function is twice continuously differentiable and downward sloping for $q \in [0, \bar{q}]$. It satisfies $p(0) > c$ and $p(\bar{q}) = 0$.

Country C attempts to invent a technology that would be a perfect substitute for the exhaustible resource. The unit production cost to be incurred by the substitute technology is constant, equal to b. We assume that $b \le c$, which has the implication that, should the backstop technology be invented, there will be no demand at all from country C for the resource. Thus, if the consuming country succeeds with the innovation, it becomes completely independent of the producing country. Denote by

$u(t)$ the R&D effort rate of country C. The hazard rate of the stochastic process leading to the innovation is given by (cf. equation (10.14))

$$\frac{\dot{F}(t)}{1 - F(t)} = u(t), \quad F(0) = 0, \tag{10.28}$$

in which $F(t)$ is the probability that country C invents by time t. Denote by τ the random instant of time at which the invention occurs. Hence, $\text{Prob}(\tau \leq t) = F(t)$.

The objective functional of country P is expected profit

$$J^P(q) = E_{u(\cdot)} \int_0^{\tau} e^{-rt}[p(q(t)) - c]q(t)\,dt, \tag{10.29}$$

where we recall that country P receives zero profit for $t > \tau$ (since demand vanishes). Assume that the instantaneous profit function $\pi(q) = [p(q) - c]q$ is strictly concave for $q \in [0, \bar{q}]$. The objective functional of country C is expected welfare, given by

$$J^C(u) = E_{u(\cdot)}\left\{\int_0^{\tau} e^{-rt}[\sigma(q(t)) - f(u(t))]\,dt + e^{-r\tau}\sigma(p^{-1}(b))/r\right\}, \tag{10.30}$$

in which

$$\sigma(q) = \int_0^q p(y)\,dy - p(q)q$$

is consumers' surplus (if imports are q) in country C. The function f in (10.30) represents country C's cost of R&D efforts. This cost function is twice continuously differentiable, strictly increasing, convex, and $f(0) = f'(0) = 0$. The latter assumption has the implication that zero R&D effort is suboptimal and for the strategy u we can confine our interest to interior solutions.

The differential game defined by (10.27)–(10.30) belongs to the class of piecewise deterministic games discussed in section 8.1.2. The game has two modes only: mode 0 is active before country C has made the innovation and mode 1 becomes active if C succeeds in making the innovation. Thus, as in the model of the previous section, there can be at most one switch of mode. The switching time is the random variable τ and the probability distribution of τ is given by F (cf. equation (10.28)). The game has only one state variable s, the dynamics of which are given by (10.27).

We look for a stationary Markov perfect Nash equilibrium. Country P's stationary Markovian extraction strategy $Q(h, s)$ is a mapping

$Q : M \times (0, \infty) \mapsto [0, \bar{q}]$, where $M = \{0, 1\}$ is the set of modes. Country C's stationary Markovian effort strategy $U(h, s)$ is a mapping $U : M \times (0, \infty) \mapsto [0, \infty)$. Refer to the theory developed in chapter 8, in particular definition 8.2 and theorem 8.2. The piecewise deterministic dynamics are given by $\dot{s}(t) = -Q(h(t), s(t))$ and the switching rate (the hazard rate) is given by $k_{01} = U(h(t), s(t))$.[4] Initial conditions are $h(t) = 0$ and $s(0) = s_0 > 0$.

Define the value functions $V^P(h, s)$ and $V^C(h, s)$ for $h \in M = \{0, 1\}$ and $s \in (0, \infty)$. We need to construct a pair of value functions that are bounded and continuously differentiable, and that satisfy the HJB equations

$$rV^P(0, s) = \max\Big\{\pi(q) - qV_s^P(0, s) + U(0, s)[V^P(1, s) - V^P(0, s)] \,\Big|\, q \in [0, \bar{q}]\Big\}, \tag{10.31}$$

$$rV^P(1, s) = 0, \tag{10.32}$$

$$rV^C(0, s) = \max\Big\{\sigma(Q(0, s)) - f(u) - Q(0, s)V_s^C(0, s) + u[V^C(1, s) - V^C(0, s)] \,\Big|\, u \in [0, \infty)\Big\}, \tag{10.33}$$

$$rV^C(1, s) = \sigma(p^{-1}(b)) \tag{10.34}$$

for all $s \in (0, \infty)$.

The HJB equations in mode 1 state the following. Equation (10.32) reflects our assumption that demand for the resource vanishes if the innovation is made and then the producing country makes no profits at all. In (10.34) we note that $V^C(1, s) = \sigma(p^{-1}(b))/r =: \overline{V^C}$ is the present value of a constant stream of consumer surplus over an infinite interval of time. This value is the payoff to country C from making the innovation. We add the following constraints on the value functions:

$$V^P(h, 0) = 0 \text{ for all } h \in M, \quad V^C(0, 0) = \underline{V^C}. \tag{10.35}$$

The first constraint states that P's profit is zero, irrespective of the system mode, if there is no remaining resource to exploit. In the second condition, $\underline{V^C}$ represents the optimal expected payoff to C in the absence of any extraction by P. Thus, $\underline{V^C}$ is defined as the optimal value of a one-player stochastic control problem with country C as the decision maker. If P does not extract any resource, consumer surplus in country C is zero as long as the country has not yet made the innovation.

[4] In order to avoid a possibly confusing notation we have denoted the switching rate q of chapter 8 by k.

Denote by $\Phi^P(s)$ and $\Phi^C(s)$ the sets of all $q \in [0, \bar{q}]$ and $u \in [0, \infty)$, respectively, that are maximizers in (10.31) and (10.33). If $Q(0, s) \in \Phi^P(s)$ and $U(0, s) \in \Phi^C(s)$ for all $s \in (0, \infty)$ then theorem 8.2 tells us that the strategy pair $(Q(h, s), U(h, s))$ is a stationary Markov perfect Nash equilibrium (provided, of course, that value functions can be found that satisfy the HJB equations).

Maximization in (10.31) and (10.33) yields a unique maximum in both cases. In (10.31) it follows from the assumption of strict concavity of the function $\pi(q)$ and from the plausible hypothesis that the shadow price $V_s^P(0, s)$ is strictly positive (we shall see that in equilibrium this hypothesis is satisfied). In (10.33) it follows from the fact that the cost of R&D efforts is a strictly convex function. It is convenient to rewrite (10.31) and (10.33). Define for mode 0

$$G^P(z) = \max\{\pi(q) - qz \mid 0 \le q \le \bar{q}\} \text{ and } G^C(z) = \max\{zu - f(u) \mid 0 \le u\}.$$

Then we can rewrite (10.31) and (10.33) as

$$rV^P(0, s) = G^P(V_s^P(0, s)) - U(0, s)V^P(0, s) \tag{10.36}$$

$$rV^C(0, s) = \sigma(Q(0, s)) - Q(0, s)V_s^C(0, s) + G^C[\overline{V^C} - V^C(0, s)] \tag{10.37}$$

in which

$$Q(0, s) = \operatorname{argmax}\{\pi(q) - qV_s^P(0, s) \mid 0 \le q \le \bar{q}\} \tag{10.38}$$

$$U(0, s) = \operatorname{argmax}\{[\overline{V^C} - V^C(0, s)]u - f(u) \mid 0 \le u\}. \tag{10.39}$$

By (10.36)–(10.39) we have a system of autonomous differential equations to determine the two value functions. This system is not explicitly solvable since the demand function and the R&D cost function are not specified. We have to resort to a qualitative analysis and wish to make this analysis in the payoff space, that is, in the (V^P, V^C) plane. For this purpose, let $\pi_m = \max\{\pi(q) \mid q \in [0, \bar{q}]\}$ be the monopoly profits. Let $q_m = \operatorname{argmax}\{\pi(q) \mid q \in [0, \bar{q}]\}$ be P's monopoly production rate and define $\hat{V}^C = \sigma(\bar{q})/r$.

Figure 10.1 depicts a phase diagram in the (V^P, V^C) plane under the assumption that V^P is positive. The following properties can be established.

(i) The isoclines $V_s^P = 0$ and $V_s^C = 0$ intersect at a unique point E_∞. This point can be found from the system (10.36)–(10.37).

(ii) The $V_s^P = 0$ isocline is upward sloping for $0 < V^P < \pi_m/r$ and becomes vertical for $V^P = \pi_m/r$. If V^P goes to zero, the isocline approaches the V^C axis.

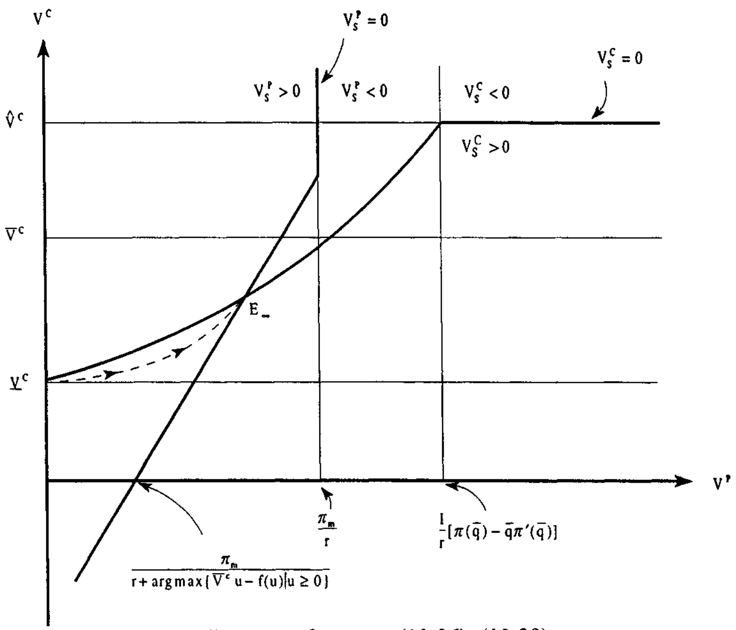

Figure 10.1 Phase diagram of system (10.36)–(10.39)

(iii) The $V_s^C = 0$ isocline is upward sloping for $\underline{V}^C < V^C < \hat{V}^C$ and goes to zero if V^C approaches $\underline{V}^C$. The isocline becomes horizontal at $V^C = \hat{V}^C$.

Taking the initial conditions (10.35) into account, we need to find a path in the phase diagram that starts in the point $(V^P = 0, V^C = \underline{V}^C)$ and converges toward E_∞. One can show that such a path exists and it holds that $V_s^P > 0$, $V_s^C > 0$ along the path. Hence, both value functions are strictly increasing functions of the remaining stock. It can also be shown that the value functions are bounded and continuously differentiable for $s \in [0, \infty)$. The equilibrium strategies are constructed by inserting the value functions so identified into (10.38) and (10.39). The equilibrium resource path can be found from (10.27).

The equilibrium effort rate is decreasing in the stock of the resource. This follows from the strict convexity of cost function f and the fact that in equilibrium it holds that $f'(U(0, s)) = \overline{V}^C - V^C(0, s)$ which, in turn, implies $f''(U(0, s))U_s(0, s) = -V_s^C(0, s) < 0$. The equilibrium output rate of the producer satisfies the conditions $\lim_{s\to 0} Q(0, s) = 0$ and $\lim_{s\to\infty} Q(0, s) = q_m$ but output is not necessarily monotonic in s. Finally, it can be shown that there exists a finite instant of time after which either country C has made the innovation (and some stock of resource is still remaining) or the resource has been completely used up.

10.4 Further reading

The results on second order differential equations and Riccati differential equations employed in sections 10.1 and 10.2 can be found in, for instance, Brock and Malliaris [12].

The model of section 10.1 was studied by Simaan and Takayama [218] under the assumption that the speed of adjustment equals 1. Fershtman and Kamien [89] dealt with the open-loop and Markovian Nash equilibria in the model of section 10.1. Fershtman and Kamien [90] analysed the case of nonstationary Markovian strategies in a finite horizon model. Further analyses of the infinite horizon problem are provided by Tsutsui and Mino [233] and Tsutsui [232]. These two papers also address the general problem of how to generate a solution to the HJB equation when a state constraint or a control constraint is explicitly taken into account. The papers use the technique of differentiating the HJB equation with respect to the state variable (see also section 4.2). In some games (among them, linear-quadratic ones) this yields an equation which may be more readily analysed than the HJB equation itself. Tsutsui [232] considers explicit lower and upper bounds on the production rates, in which case strategies and value functions must be synthesized from solution segments corresponding to minimal, maximal, or interior output rates. Dockner [44] extends the problem in Fershtman and Kamien [89] to an N-player oligopoly and investigates the limit game occurring when N goes to infinity. Driskill and McCafferty [67] include an adjustment cost (incurred when the production rate is changed) in the Fershtman and Kamien [89] model.

We saw that the sufficient optimality conditions of chapter 3 failed to be applicable in the model of section 10.1. Other sufficiency theorems then can be explored, for instance, Dockner et al. [53], and Leitmann and Stalford [159].

The idea of letting the hazard rate of a stochastic process depend on the control variables goes back to Kamien and Schwartz [146]. Applications of this idea in the area of R&D and technological innovation are, for instance, Reinganum [197, 198, 200, 201], Kamien and Schwartz [147], and Feichtinger [81]. The R&D model of section 10.2 is due to Reinganum [197, 198]. See also Mehlmann [51] and Dockner et al. [51]. A solution of the R&D game of section 10.2 with time-varying P_I and P_F can be found in Reinganum [201] for the case of exponentially growing P_I and P_F and the special case of $e^{gt} = e^{rt}$. The game of R&D

and exhaustible resource extraction in section 10.3 is due to Harris and Vickers [116].

10.5 Exercises

1. In the model of section 10.1, verify the expressions for the constants K, E, and G in the value function of case (B).
2. Determine a Markovian Nash equilibrium for the game of section 10.1 under the assumption that the game is played over a fixed and finite horizon $[0, T]$.
3. Analyse the innovation game in section 10.2 by using the HJB approach.
4. Consider the innovation game in section 10.2 and determine the optimal strategies if $P_F = 0$ (which may be interpreted to mean a case of perfect patent protection). What does this imply for the optimal R&D strategies? What would be the consequence of increasing the number of firms in the market?
5. In the game of section 10.2, change the dynamics (10.24) to

$$\dot{y}(t) = -\alpha\lambda y(t) \sum_{i=1}^{N} u_i(t)$$

   and apply the maximum principle approach to determine a Markovian Nash equilibrium. The constant α in the dynamics can be thought of as a scaling factor which changes the unit of measurement of the state variables z_i. The change of dynamics amounts to changing the state variable from y to y^α. Then show that the Markovian strategies no longer are independent of the state y.
6. In the game of section 10.2, consider a cooperative scenario in which the firms agree to maximize their joint profits and employ coordinated R&D strategies. To simplify, consider the case of $N = 2$ and assume that $P_F = 0$ (winner takes all). Suppose that the firms exchange know-how, which can be interpreted to mean that a certain fraction of a firm's R&D knowledge spills over to the other firm. To accommodate this extension, change the dynamics in (10.18) to

$$\dot{z}_i(t) = u_i(t) + \gamma u_j(t), \qquad i = 1, 2, \ i \neq j,$$

   where $\gamma \in [0, 1]$ and recall that the problem is symmetric. Determine the Markovian Nash R&D effort strategies and compare the results of the cooperative game with those obtained in section 10.2.

11 Differential games in marketing

Essential issues that must be addressed when designing a firm's marketing strategy are the strategic intertemporal interdependencies between the firm and its competitors in the market, the firm and the various agents in the firm's distribution channel, and the firm and its customers. Here, we confine our interest to the first two issues (which, however, does not mean that the third one is less important).

A firm often tries to attain its marketing objectives (e.g., sales revenue, market share) vis-à-vis its competitors through a range of direct marketing efforts (including, for example, price, advertising, quality, and distribution), but cost-reducing efforts or strategic investments can also increase a firm's competitive position. Any marketing strategy needs to be coordinated with investment and capacity plans (see chapter 9), financial planning, and R&D policy (see chapter 10).

The chapter proceeds as follows. First we present a game in which two duopolistic firms use their advertising efforts to compete for market share. Then we study an oligopoly pricing game of new product diffusion. Third, we introduce a game between the two parties in a vertical distribution channel: a manufacturer and a retailer. The area of marketing has been rather popular for differential game applications, and in section 11.4 we list a series of other applications. With respect to modelling, we introduce the following categories: market share models (section 11.1), cumulative sales models (section 11.2), and advertising goodwill models (section 11.3).

11.1 A duopolistic advertising game

Some market share models are capable of dealing with an oligopolistic setting. The Lanchester model, which is used in this section, is designed for duopolistic competition. Originally developed to model military combat, the Lanchester model has been applied to a number of studies of battles for market share occurring in duopolistic markets. Denote by $x(t)$ the market share of firm 1 at time t and assume that the size of the total market is constant over time. Normalizing the total market to 1, we obtain $1 - x(t)$ as the market share of firm 2 at time t. The standard Lanchester dynamics are

$$\dot{x}(t) = u^1(t)[1 - x(t)] - u^2(t)x(t), \quad x(0) = x_0 \in [0, 1], \tag{11.1}$$

in which $u^i(t)$ is the rate of advertising expenditure of firm $i \in \{1, 2\}$ at time t. The right-hand side of (11.1) reflects the assumption that a firm's advertising efforts only influence the rival firm's customers, not the firm's own customers. One can suppose that a firm's advertising efforts are particularly designed to draw away customers from the rival firm. Basically, the model assumption is that consumers are disloyal and drift toward the firm that spends substantial amounts of money on advertising.

In the sequel we consider a modification of the Lanchester model (11.1), namely

$$\dot{x} = u^1(t)\sqrt{1 - x(t)} - u^2(t)\sqrt{x(t)}. \tag{11.2}$$

To motivate the modification, approximate the two square root terms in (11.2) by the quadratic terms $1 - x(t) + x(t)[1 - x(t)]$ and $x(t) + x(t)[1 - x(t)]$, respectively. Then we can rewrite (11.2) as

$$\begin{aligned}\dot{x}(t) &= u^1(t)\{1 - x(t) + x(t)[1 - x(t)]\} - u^2(t)\{x(t) + x(t)[1 - x(t)]\} \\ &= u^1(t)[1 - x(t)] - u^2(t)x(t) + [u^1(t) - u^2(t)]x(t)[1 - x(t)],\end{aligned}$$

which shows that (11.2) includes an imitation term $x(t)[1 - x(t)]$ that intends to model social interactions between a firm's own customers and the rival firm's customers. Moreover, the Lanchester model (11.1) has been extended with an excess advertising (advertising differential) term, $u^1(t) - u^2(t)$. This term reflects a hypothesis that firm 1's market share increases if the firm's advertising rate exceeds that of the competitor. Thus, consumers compare the two firms' advertising efforts and drift to the firm with the higher advertising effort. The excess advertising effect is largest if $x(t)$ is neither very small nor very large. For other applications of the excess advertising idea, see exercises 3 and 4.

For notational economy, define $x_1(t) = x(t)$ and $x_2(t) = 1 - x(t)$ so that $x_i(t)$ denotes firm i's market share at time t. The firms' payoffs are

$$J^i = \int_0^T e^{-r_i t}\{q_i x_i(t) - (c_i/2)[u^i(t)]^2\}\, dt + e^{-r_i T} S_i x_i(T),$$

where $r_i > 0$ represents firm i's discount rate and $q_i > 0$ is firm i's revenue per unit of market share. We have assumed convex advertising cost functions, $c_i > 0$, and linear salvage value functions in which $S_i > 0$ represents firm i's evaluation of a unit of its terminal market share. Firm i wishes to choose its advertising effort $u^i(t)$, $t \in [0, T]$, such that the objective J^i is maximized subject to $u^i(t) \geq 0$ for all $t \in [0, T]$. The state space is specified by the constraint $0 \leq x_i(t) \leq 1$, which must be satisfied for all $t \in [0, T]$.

We wish to identify two equilibria, an open-loop Nash equilibrium and a Markov perfect Nash equilibrium. Playing open-loop strategies means that the firms use time functions as their advertising strategies while the use of nondegenerate Markovian strategies means that a firm conditions its choice of its current advertising rate on time and its current market share.

For an open-loop Nash equilibrium $(u^1(\cdot), u^2(\cdot))$, define a current-value Hamiltonian of each firm. Denote by $\lambda^i(t)$ the current-value costate variable associated with the state variable $x_i(t)$. Hence

$$H^i(x_i, u^i, \lambda^i, t) = q_i x_i - \frac{c_i}{2}(u^i)^2 + \lambda^i\left[u^i\sqrt{1 - x_i} - u^j(t)\sqrt{x_i}\right].$$

To get a first insight, we note that zero advertising is desirable in terms of costs, but the dynamics show that this would make firm i's market share decrease as long as the rival firm maintains a positive amount of advertising effort. On the other hand, a large value of advertising effort is desirable as far as growth of market share is concerned, but large advertising efforts are heavily penalized by the convex advertising cost function. An optimal solution is likely to be a compromise between such extreme control policies. Note that the costate variable has an interpretation as a shadow price of firm i's market share. As market share is a good stock, intuition suggests that this shadow price is positive.

Refer to theorem 4.2 and the discussion of open-loop Nash equilibria in section 4.2. In the advertising model, the sets of feasible controls are independent of the state variable. Condition (i) of the theorem requires that, given $(x_i(t), \lambda^i(t), t)$, the Hamiltonian H^i attains its maximum with respect to u^i at $u^i(t)$. Conditions (ii)–(iii) require that the costate equation and the transversality condition are satisfied. Hence, for $i, j = 1, 2$, $i \neq j$:

$$u^i(t) = \max\left\{0, [\lambda^i(t)/c_i]\sqrt{1 - x_i(t)}\right\},$$

$$\dot{\lambda}^i(t) = r_i\lambda^i(t) - q_i + \frac{\lambda^i(t)u^i(t)}{2\sqrt{1 - x_i(t)}} + \frac{\lambda^i(t)u^j(t)}{2\sqrt{x_i(t)}}, \ \lambda^i(T) = S_i. \qquad (11.3)$$

Using the costate equation yields $\dot{\lambda}^i(t)|_{\lambda^i(t)=0} = -q_i < 0$, which shows (by using the transversality condition) that $\lambda^i(t) > 0$ for all $t \in [0, T]$. This is what we expected. Therefore, the control $u^i(t)$ is strictly positive for all t and is given by

$$u^i(t) = \frac{\lambda^i(t)}{c_i}\sqrt{1 - x_i(t)}\,, \; i = 1, 2. \tag{11.4}$$

Inserting the right-hand side of (11.4) in the costate equations (11.3) yields the following system of (coupled) differential equations for the two costates:

$$\dot{\lambda}^i(t) = \frac{\lambda^i(t)^2}{2c_i} + \left[\frac{\lambda^j(t)}{2c_j} + r_i\right]\lambda^i(t) - q_i\,, \; i, j = 1, 2. \tag{11.5}$$

The state $x_i(t)$ does not appear on the right-hand side of (11.5), but the control $u^i(t)$ in (11.4) depends on $x_i(t)$. Thus, the game is not a linear state game in the sense of chapter 7.

Denote by $\eta^1(t)$ and $\eta^2(t)$ the unique solution to (11.5), satisfying the transversality conditions $\eta^1(T) = S_1$ and $\eta^2(T) = S_2$. Denote by $y(t)$ the unique solution of the state equation

$$\dot{x}(t) = \frac{\eta^1(t)}{c_1}[1 - x(t)] - \frac{\eta^2(t)}{c_2}x(t)\,, \; y(0) = x_0. \tag{11.6}$$

Collecting our results, the following (unique) equilibrium candidate strategies have been obtained:

$$u^1(t) = \frac{\eta^1(t)}{c_1}\sqrt{1 - y(t)}\,, \; u^2(t) = \frac{\eta^2(t)}{c_2}\sqrt{y(t)}. \tag{11.7}$$

Refer to the discussion in section 3.5 concerning different representations of optimal paths and note that the firms also could achieve their equilibrium profits by using the nondegenerate Markovian strategies $u^i(t) = \phi^i(x_i(t), t)$, where

$$\phi^i(x_i, t) = \frac{\eta^i(t)}{c_i}\sqrt{1 - x_i}\,, \; i = 1, 2. \tag{11.8}$$

The concavity condition from theorem 4.2 is not satisfied in the present model, because the maximized Hamiltonian

$$H^{i*}(x_i, \lambda^i, t) = q_i x_i + \frac{(\lambda^i)^2}{2c_i}(1 - x_i) - \lambda^i u^j(t)\sqrt{x_i}$$

is strictly convex with respect to x_i. Fortunately, one can show that, in the model under consideration, the conditions of the maximum principle are necessarily satisfied by an open-loop Nash equilibrium and that such an equilibrium exists (see the original reference cited in section 11.4). Since we have identified the unique solution of the equilibrium conditions, it follows that this solution is indeed the unique open-loop Nash equilibrium of the game.

The open-loop Nash equilibrium that we have derived is (as are all open-loop equilibria) also a Markovian Nash equilibrium, but the open-loop equilibrium at hand is not a Markov perfect Nash equilibrium. To derive a Markov perfect Nash equilibrium we turn to theorem 4.4, which is based on the HJB equations. (Recall that theorem 4.2 was based on the maximum principle.)

Markovian strategies are functions $\phi^i : [0,1] \times [0,T] \mapsto [0,\infty)$, where the interval $[0,1]$ is the state space. Let $\phi^i(x_i, t)$ denote a Markovian strategy of player i and consider the value functions $V^i(x_i, t)$, $i = 1, 2$. Condition (ii) of theorem 4.4 requires

$$V^i(x_i, T) = S_i x_i, \qquad i \in 1, 2 \tag{11.9}$$

for all $x_i \in [0,1]$. From (4.10) we obtain for $i, j = \{1, 2\}$ and $x_i \in [0,1]$ the HJB equations

$$\begin{aligned} &r_i V^i(x_i, t) - V^i_t(x_i, t) \qquad (11.10)\\ &= \max\left\{q_i x_i - \frac{c_i}{2}(u^i)^2 + V^i_{x_i}(x_i, t)\left[u^i\sqrt{1-x_i} - \phi^j(1-x_i, t)\sqrt{x_i}\right] \,\middle|\, u_i \geq 0\right\}. \end{aligned}$$

Maximization on the right-hand side of (11.10) yields

$$u^i = \frac{1}{c_i} V^i_{x_i}(x_i, t)\sqrt{1-x_i}\,, \; i \in \{1, 2\}. \tag{11.11}$$

To determine value functions that are solutions to the HJB equations we use the informed guessing approach (see chapter 3). Let us conjecture that the value function V^i is linear in the state variable x_i, that is,

$$V^i(x_i, t) = \beta^i(t) x_i + \nu_i(t)\,, \; i \in \{1, 2\}. \tag{11.12}$$

According to (11.9) and (11.10), the functions $\beta^i(t)$ in (11.12) must satisfy $\beta^i(T) = S_i$ and the differential equations

$$\dot{\beta}^i(t) = r_i \beta^i(t) - q_i + \frac{\beta^i(t)^2}{2c_i} + \frac{\beta^i(t)\beta^j(t)}{c_j}\,, \; i, j = \{1, 2\}. \tag{11.13}$$

The functions $\nu_i(t)$ in (11.12) must satisfy the boundary conditions $\nu_i(T) = 0$ and the linear differential equations

$$\dot{\nu}_i(t) = r_i \nu_i(t) - \frac{[\beta^i(t)]^2}{2c_i}\,, \; i \in \{1, 2\}. \tag{11.14}$$

The next step is to see if the conjectured value functions in (11.12) satisfy the requirements of theorem 4.4. The value functions in (11.12) are continuously differentiable in both arguments and it is easy to calculate

$$V_t^i(x_i, t) = \dot{\beta}^i(t)x_i + \dot{\nu}_i(t), \quad V_{x_i}^i(x_i, t) = \beta^i(t), \quad i \in \{1, 2\}. \tag{11.15}$$

The functions $\beta^i(t)$ and $\nu_i(t)$ must be chosen so as to satisfy the boundary conditions $\beta^i(T) = S_i$ and $\nu_i(T) = 0$, respectively, to fulfil condition (iii) of theorem 4.1. Substitute the right-hand side of (11.11) into (11.10) and use (11.15). This yields, for $\{i, j\} = \{1, 2\}$, the equation

$$x_i\left[r_i\beta^i(t) - \dot{\beta}^i(t) - q_i + \frac{\beta^i(t)^2}{2c_i} + \frac{\beta^i(t)\beta^j(t)}{c_j}\right] = \dot{\nu}_i(t) - r_i\nu_i(t) + \frac{\beta_i(t)^2}{2c_i}.$$

This equation is valid for all $x_i \in [0, 1]$ if and only if equations (11.13) and (11.14) are satisfied. Solutions $\beta^i(t)$ and $\nu_i(t)$ can be found that satisfy these two equations.

To sum up, a Markov perfect Nash equilibrium is provided by the strategy pair

$$\phi^i(x_i, t) = \frac{\beta^i(t)}{c_i}\sqrt{1 - x_i}, \quad i \in \{1, 2\}, \tag{11.16}$$

where the functions $\beta^i(t)$ solve (11.13) subject to $\beta^i(T) = S_i$. Condition (i) of theorem 4.4 holds because the state trajectory associated with the equilibrium strategies is the unique solution to the initial value problem

$$\dot{x}(t) = \frac{\beta^1(t)}{c_1}[1 - x(t)] - \frac{\beta^2(t)}{c_2}x(t), \quad x(0) = x_0 \in [0, 1].$$

How did we guess that the value functions are linear? In the game at hand, the open-loop Nash equilibrium provides a clue. In this equilibrium, the strategies can be represented in the form of the feedback rule (11.8). This leads us to investigate whether or not this rule would yield a Markovian Nash equilibrium, provided, of course, that the costate functions are chosen appropriately. Comparing (11.8) and (11.11) provides the answer.

In the language of a control engineer, the two important equations (11.8) and (11.16) are feedback rules. But notice that (11.8) is a rule that generates the open-loop Nash equilibrium paths whereas (11.16) describes the Markov perfect Nash equilibrium strategies. The two equations show that optimal advertising policies are of the same structure, but the time functions $\eta^i(\cdot)$ and $\beta^i(\cdot)$ are certainly not the same: the former satisfies (11.5) whereas the latter satisfies (11.13). Despite this difference, the rules show that firms should choose to spend more on advertising the smaller is their market share and vice versa. This kind of advertising policy is rather intuitive and has been observed in other market share games. The common structure of the equilibrium advertising strategies also allows for an additional characterization of optimal advertising strategies. Irrespective of whether ϕ^i denotes the feedback rule generating the open-loop equilibrium paths or the Markov perfect equilibrium strategies it holds that

$$\frac{\partial \phi^i}{\partial x_j}\frac{x_j}{\phi^i} = \frac{1}{2}, \ \{i, j\} = \{1, 2\}.$$

This means that, at any time t, an increase of $(2p)$ per cent of the rival's market share leads to a p per cent increase of a firm's advertising effort. Obviously, the number 2 is specific to the particular model at hand. Nevertheless, both types of Nash equilibria provide a guideline for the design of a firm's optimal advertising policy.

We wish to characterize the qualitative behaviour of equilibrium advertising strategies over time. Equations (11.8) and (11.16) show that the advertising policies depend on time via functions $\eta^i(\cdot)$ and $\beta^i(\cdot)$, respectively, henceforth denoted as costate functions. A study of the behaviour of the costate functions could help to characterize the optimal advertising rates over time. To simplify the notation, note that (11.5) and (11.13) can be subsumed under a common representation. Denote by $m^i(\cdot)$ the costate functions in both open-loop and Markov perfect equilibrium. For $\{i, j\} \in \{1, 2\}$, (11.5) and (11.13) can be represented by

$$\dot{m}^i(t) = r_i m^i(t) - q_i + \frac{m^i(t)^2}{2c_i} + k\frac{m^i(t)m^j(t)}{2c_j}, \ m^i(T) = S_i \quad (11.17)$$

such that one obtains (11.5) if $k = 1$ and (11.13) if $k = 2$. Owing to the nonnegativity of the costates, the relevant region for a phase diagram analysis is the first quadrant of the (m^1, m^2) plane. Putting the left-hand side of (11.17) equal to zero readily shows that the isoclines are hyperbolas, given by

$$m^j = \frac{2c_j}{k}\left[\frac{q_i}{m^i} - r_i - \frac{m^i}{2c_i}\right], \ \{i, j\} = \{1, 2\}.$$

Figure 11.1 depicts the isoclines by two dashed curves. It is straightforward to verify that both isoclines are strictly decreasing and that they intersect at a unique steady state point, say $(\hat{m}^1, \hat{m}^2)$. This point is in the interior of the first quadrant.

The Jacobian matrix J associated with the differential equation system (11.17), evaluated at the steady state, is given by

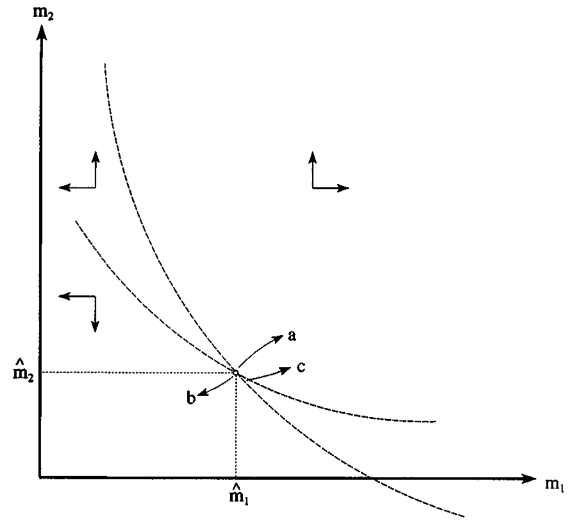

Figure 11.1 Phase diagram of system (11.17)

$$J = \begin{pmatrix} A_1 & B_1 \\ B_2 & A_2 \end{pmatrix},$$

where

$$A_i = r_i + \frac{m^i}{c_i} + k\frac{m^j}{2c_j}\,, \quad B_i = k\frac{m^i}{2c_j}\,, \quad \{i,j\} = \{1,2\}.$$

In the positive quadrant, A_i and B_i are positive. It is straightforward to establish that

$$\operatorname{tr} J = A_1 + A_2 > 0\,, \quad \det J = \left(r_1 + \frac{m^1}{c_1}\right)\left(r_2 + \frac{m^2}{c_2}\right) + \sum_{i=1}^{2} \frac{km^i}{2c_i}\left(r_i + \frac{m^i}{c_i}\right),$$

$$(\operatorname{tr} J)^2 - 4\det J = (A_1 - A_2)^2 + 4B_1B_2 > 0.$$

Denote the eigenvalues of J by θ_i. Using the above expressions for the trace and determinant of J, as well as the formula

$$\theta_{1,2} = \frac{1}{2}\left[\operatorname{tr} J \pm \sqrt{(\operatorname{tr} J)^2 - 4\det J}\right],$$

shows that both eigenvalues of J are strictly positive. Then from the theory of differential equations we know that the steady state is an unstable node. Figure 11.1 illustrates this by the arrows pointing outward from the steady state point. Trajectories $m^i(t)$ start out near the steady state point and reach at time T the point (S_1, S_2) that is prescribed by the transversality conditions. The figure illustrates three possible cases such that each particular case arises for a specific configuration of the salvage value parameters. Consider, for instance, case (a), where both firms have high evaluations of their terminal market shares, in the sense that both S_1 and S_2 are large relative to the steady state point. Both costates $m^i(t)$ increase over time, which is intuitive since market share is a good stock and the costate is the marginal valuation of that stock.

It is important to note that the steady state point is not the same in the two equilibria (open-loop and Markov perfect). Among other things, this means that monotonicity properties of the costate functions may not be the same in the two equilibria. For example, it could happen that in open-loop equilibrium both costates decrease (cf. case (b)) but in Markov perfect equilibrium both costate functions increase (cf. case (a)). It can be shown that the shadow price of market share of at least one firm is lower in the Markov perfect equilibrium than in open-loop equilibrium. This result carries over to the firms' optimal advertising. The intuition is that the state information used in the Markov perfect equilibrium reduces advertising outlays. In open-loop equilibrium, each firm uses a fixed advertising time function and at least one of the firms uses higher advertising expenditures. As to the bottom line, the two firms' profits are not necessarily higher in Markov perfect than in open-loop equilibrium. To settle this question, however, one needs to calculate explicitly the values of the optimal payoffs.

So far we have considered the market share game with a finite planning period T. Now suppose that the planning horizon is infinite. In this case the salvage values S_i are put equal to zero. To study infinite horizon games we must make sure that the objective functionals are well defined, in the sense that they are upper bounded for all feasible strategy pairs and their associated state trajectory (cf. condition (ii) of theorem 4.4). In the game at hand, observe that $x(t) \in [0, 1]$ and that the discount rates r_i are strictly positive. The integrals defining the objectives are upper bounded, which can be seen by noting that

$$\int_0^\infty e^{-r_i t} q_i \, dt = \frac{q_i}{r_i} < \infty, \quad i \in \{1, 2\}.$$

Thus, maximization of the objective functionals makes sense in the infinite horizon version of our model. If this were not true, we would have to resort to weaker notions of optimality, e.g., catching-up optimality, as defined in definition 3.2.

To identify a Markov perfect Nash equilibrium of the infinite horizon game, note that the model is autonomous. Then it is plausible to look for a stationary Markov perfect equilibrium in which advertising strategies and value functions depend only on the state variable. Denote by $(\hat{\beta}^1, \hat{\beta}^2)$ the unique solution in the positive quadrant of the algebraic equations (cf. equation (11.13))

$$0 = r_i\hat{\beta}^i - q_i + \frac{\left(\hat{\beta}^i\right)^2}{2c_i} + \frac{\hat{\beta}^i\hat{\beta}^j}{c_j}, \; \{i,j\} = \{1,2\}. \tag{11.18}$$

We conjecture that the value functions have the linear form

$$V^i(x_i) = \hat{\beta}^i x_i + \hat{v}_i, \; \hat{v}_i = \frac{\left(\hat{\beta}^i\right)^2}{2c_i r_i}, \; i \in \{1,2\}. \tag{11.19}$$

In (11.19) the coefficients in the value functions are constants. Advertising strategies are given by (11.11), in which we have $V^i_{x_i}(x_i) = \hat{\beta}^i$ for $i \in \{1,2\}$. Thus, they can be expressed as

$$\phi^i(x_i) = \frac{\hat{\beta}^i}{c_i}\sqrt{1 - x_i}, \; \{i,j\} = \{1,2\}. \tag{11.20}$$

What remains to be shown is that our conjectured value functions do satisfy the HJB equations. Continuity and continuous differentiability is obvious. The HJB equations are given by (11.10), in which we put $V^i_t = 0$. Inserting from (11.19)–(11.20) into the HJB equations yields (11.18), and since this equation is satisfied by our choice of $(\hat{\beta}^1, \hat{\beta}^2)$ we conclude that the HJB equations hold.

11.2 An oligopoly game of new product pricing

The marketing science literature using differential games deals with a wide range of problems in strategic pricing: new product pricing, limit pricing and entry problems, the effects of cost experience on optimal pricing policies, and pricing in distribution channels. In this section we focus on two important factors that affect the pricing of new products in an oligopolistic setting: cost learning and new product adoption dynamics.

In the manufacture of some products, current productivity has been observed to be related to former experience in production activities. This phenomenon – learning-by-doing – has been documented in, for example, chemical processing, manufacturing of airframes, semiconductors, and memory chips. Using cumulative output as a proxy for production experience, learning-by-doing implies that productivity increases as cumulative output increases. Productivity increases are reflected in decreasing unit costs of production. To formalize, let $x_i(t)$ denote the cumulative output by time t of firm i in an N-firm industry and let $c_i(x_i(t))$ represent the unit cost of production of firm i. The learning-by-doing hypothesis states that $dc_i(x_i)/dx_i = c_i'(x_i)$ is negative. In multi-firm situations, an interesting phenomenon is spillover of learning, that is, some or all of a firm's own learning can be exploited by the rival firms. However, in the example considered in this section there are no such spillovers. In the context of R&D competition (see chapter 10) we noted the possibility of spillover of know-how in R&D activities.

Cumulative sales models (market growth models, new product diffusion models) form a class of marketing models that intend to describe the process by which new products penetrate a market. Quite often it is assumed that new product adoption is driven by three sources: a word-of-mouth (or imitation) process generated by consumers' social interaction, the firms' mass-communication efforts (in particular, advertising), and the pricing policies employed by the firms. In general, the sales dynamics describe a firm's current sales rate as a function of all firms' cumulative sales and the firms' current marketing efforts. By including the firms' cumulative sales rates in the dynamics, one is able to model important dynamic demand effects such as innovation, imitation, and market saturation. The model describes the diffusion (adoption) process of a new product (or a number of new products) among a group of potential buyers. The diffusion process starts out with a majority of potential consumers being uninformed and ignorant of the new product(s). Gradually, awareness is created and eventually some consumers take the step of adopting the product(s), i.e., buy the product(s). These early adopters are innovative and eager to try new products. Firms that introduce new products wish to influence the innovative consumers, not only to obtain the sales to these consumers but also since innovative consumers are thought to be influential for the purchase decisions of non-innovative consumers. Other segments of consumers are more reluctant to respond to new product innovations and adopt the product only later, if ever. The buying decision of these consumers could be a consequence of the firms' marketing policies (for instance, because prices have gradually been lowered) but could also be a product of social interaction and imitation of the behaviour of innovators.

The model which we analyse in this section applies to a single product category of a durable product. (An example might be cellular phones.) Each of N competing firms launches at the same time its own, particular brand. We assume that the number of potential buyers is fixed during the planning period and that each adopter buys one and only one unit during the planning period. To set up the model, let $x_i(t)$ denote the number of individuals who have adopted the product of firm $i = 1, 2, \ldots, N$ by time t. This number equals the cumulative sales of firm i by time t, which (assuming away inventories) equals the cumulative production of firm i by time t. Let $p^i(t)$ represent the unit price quoted by firm i at time t. The time derivative of $x_i(t)$ represents the sales (and production) rate of firm i at time t. Assuming $c_i'(x_i) \leq 0$ means that each firm's unit cost of production is either strictly decreasing (cost learning effect) or constant (no cost learning effect). The diffusion model is stated by the differential equations

$$\dot{x}_i(t) = f^i(x_1(t), x_2(t), \ldots, x_N(t), p^1(t), p^2(t), \ldots, p^N(t)), \quad x_i(0) = 0, \tag{11.21}$$

where the function f^i has nonnegative values and is twice continuously differentiable. The dynamics in (11.21) state that the sales rate at time t of firm i depends on its own cumulative sales and own current price but also on the cumulative sales and current prices of the firm's $N - 1$ competitors. The dependence of current sales rate of a firm on the current prices of all brands in the market is a classic hypothesis in the theory of oligopolistic competition. What is particular to the diffusion model approach is that the sales rate of any firm depends on the cumulative sales of all firms.

We need to introduce some additional assumptions on the functions f^i:

$$\frac{\partial f^i}{\partial p^i} < 0, \quad \frac{\partial f^i}{\partial p^j} > 0, \quad \sum_{j=1}^{N} \frac{\partial f^i}{\partial p^j} < 0, \quad \frac{\partial^2 f^i}{\partial p^i \partial p^j} \leq 0, \tag{11.22}$$

$$1 - \frac{\partial^2 f^i}{\partial p^i \partial p^i} \frac{f^i}{\left(\frac{\partial f^i}{\partial p^i}\right)^2} \geq 0, \quad 1 - \frac{\partial^2 f^i}{\partial p^i \partial p^j} \frac{f^i}{\frac{\partial f^i}{\partial p^i} \frac{\partial f^i}{\partial p^j}} \geq 0. \tag{11.23}$$

In (11.22), the first inequality expresses the usual assumption of a downward-sloping sales function. The second inequality states that brands i and j are substitutes. The third inequality means that if all firms simultaneously raise their prices, firm i experiences a decrease in the sales of its own brand. The fourth inequality represents an assumption saying that it is more difficult to increase the sales rate of firm i by lowering this firm's price when the jth competitor's price is high than when it is low. The first inequality in (11.23) states, roughly speaking, that the sales rate function f^i must not be 'too convex' in p^i. This assumption is satisfied if, for example, f^i is linear or concave in p^i and has the implication that $\partial^2 H^i / \partial (p^i)^2 \leq 0$. The second inequality in (11.23) is not readily interpretable. The reader might regard it as a technicality. It is satisfied if f^i is linear, or if p^i in f^i is additively separated from all other prices $(p^1, \ldots, p^{i-1}, p^{i+1}, \ldots, p^N)$.

Firm i wishes to choose a pricing policy over a fixed and finite horizon $[0, T]$ so as to maximize its objective functional. Suppose that each firm is mainly interested in its profits over the time horizon. Then we can omit a salvage value term and the objective functional of firm i becomes

$$J^i = \int_0^T e^{-r_i t}[p^i(t) - c_i(x_i(t))]\dot{x}_i(t)\,dt,$$

in which the sales rate $\dot{x}_i(t)$ is given by (11.21).

It turns out that playing the game with nondegenerate Markovian strategies leads to an intractable problem. Instead, we look for an open-loop Nash equilibrium, that is, we assume that the firms use pricing strategies that depend on current time only. It may be that firm i is simply unable to observe the vector of cumulative sales of its competitors. Nevertheless, an analysis of an open-loop Nash equilibrium can illustrate how to analyse a differential game in which the functional forms are stated in a general way. The reader will notice that many of the models analysed in this book involve specific functional forms.

To apply theorem 4.2, define current-value Hamiltonians as follows

$$H^i(x_1, \ldots, x_N, p^i, \lambda_1^i, \ldots, \lambda_N^i, t) = [p^i - c_i(x_i) + \lambda_i^i]f^i + \sum_{j \neq i} \lambda_j^i f^j.$$

For notational simplicity we omit the arguments $x_1, \ldots, x_N, p_i$, and $p_j(t)$ for $j \neq i$ of the function f^i. The costate variables must satisfy

$$\dot{\lambda}_i^i(t) = \lambda_i^i(t)[r_i - f_{x_i}^i] + c_i'(x_i(t))f^i - f_{x_i}^i[p^i(t) - c_i(x_i(t))] - \sum_{j=1}^N \lambda_j^i(t) f_{x_i}^j \tag{11.24}$$

$$\dot{\lambda}_j^i(t) = r_i \lambda_j^i(t) - f_{x_j}^i[p^i(t) - c_i(x_i(t)) + \lambda_i^i(t)] - \sum_{k \neq i} \lambda_k^i(t) f_{x_j}^k \tag{11.25}$$

and since there are no salvage values, the transversality conditions are

$$\lambda_j^i(T) = 0. \tag{11.26}$$

We confine our interest to interior equilibria (that is, equilibria admitting prices above unit costs), supposing that such equilibria exist. The Hamiltonian maximization conditions state that for all t it must hold that $H_{p^i}^i = 0$, which implies

$$f^i_{p^i}[p^i(t) - c_i(x_i(t)) + \lambda^i_i(t)] + f^i + \sum_{j \neq i} \lambda^i_j(t) f^j_{p^i} = 0. \tag{11.27}$$

For an economic interpretation of (11.27), introduce the elasticities

$$\eta_i = -f^i_{p^i} \frac{p^i}{f^i}, \quad \rho_{ji} = f^j_{p^i} \frac{p^i}{f^j},$$

where $\eta_i > 0$ is the direct price-elasticity of firm i's sales and $\rho_{ji} > 0$ is the cross-elasticity, representing the percentage increase in firm j's sales caused by a 1 per cent increase of firm i's price. Since the brands are substitutes, ρ_{ji} is positive. The elasticities are functions of the state and control vectors, but to simplify the notation we have left out the arguments. Using the elasticities, (11.27) is rewritten as

$$p^i(t) = \frac{\eta_i}{\eta_i - 1}[c_i(x_i(t)) - \lambda^i_i(t)] + \frac{1}{\eta_i - 1} \sum_{j \neq i} \lambda^i_j(t) \rho_{ji} \frac{f^j}{f^i}. \tag{11.28}$$

For $N = 1$ equation (11.28) yields the monopolistic dynamic pricing rule and the myopic oligopolistic pricing rule if all costates are identically zero.

For the dynamic oligopolistic case, (11.27) or (11.28) does not yield a useful characterization of the optimal pricing policies. Facing such an obstacle, an often-used research strategy is to choose less general functional forms (that is, to trade generality for tractability). Consider the specialization of (11.21) given by

$$\dot{x}^i(t) = k^i(x_i(t)) q^i(p^1(t), p^2(t), \ldots, p^N(t)). \tag{11.29}$$

Compared with (11.21), the dynamics in (11.29) mean that the sales rate of firm i still is influenced by the prices of all the rival firms, but (11.21) has been changed in two respects. Demand diffusion effects are now supposed to be firm-specific, that is, the sales rate of firm i is influenced only by its own cumulative sales, not by the cumulative sales of the competitors. Moreover, demand diffusion effects and prices influence the sales rate of firm i in a multiplicatively separable way. A drawback of such separability is that the elasticities become independent of firm i's cumulative sales. This may be less realistic for some products.

In addition to the simplification made in (11.29), we need the discount rates r_i to be sufficiently low that we can approximate them by zero. One can suppose, for instance, that the planning horizon is relatively short so that discounting does not make a significant difference. The assumptions in (11.22) and (11.23) still have to hold, but we notice that owing to (11.29), the left-hand sides of the inequalities in (11.23) are now independent of x_i. The following result concerning optimal pricing policies can be established. The proof follows.

Theorem 11.1 *Suppose that $dk^i(x_i(t))/dx_i$ is positive (negative) for all $i = 1, 2, \ldots, N$ and for all t in some interval I. Then it holds that an equilibrium price $p^j(t)$ is increasing (decreasing) for $t \in I$ and for all firms $j = 1, 2, \ldots, N$.*[1]

A situation where, for example, all firms have increasing prices could occur in an introductory phase during which all N derivatives dk^i/dx_i most likely are positive. Positivity means that all brands enjoy positive demand diffusion effects: the larger the cumulative sales of a firm, the larger its current sales. If the derivatives initially were negative it would mean that the firms had introduced 'bad' products, a less likely situation. But positive demand diffusion effects do not last for ever and, as of some instant of time, saturation effects (due to the finite size of the market) tend to dominate the positive diffusion effect. (Saturation effects could be offset by repeat sales, but in our model repeat sales are ruled out by an earlier assumption.) If a situation of positive demand diffusion effects occurs in the beginning of the planning period, we thus get what marketing people call penetration pricing. We proceed by proving theorem 11.1.

Proof The first step in the proof is to rewrite condition (11.27). Differentiating totally with respect to t in (11.27) yields

$$\dot{p}^i(t) = c_i'(x_i(t))f^i - \dot{\lambda}_i^i(t) - \frac{d}{dt}\left(\frac{f^i}{f^i_{p^i}}\right) - \sum_{j=1}^{N} \dot{\lambda}_j^i(t)\frac{f^j_{p^i}}{f^i_{p^i}} - \sum_{j=1}^{N} \lambda_j^i(t)\frac{d}{dt}\left(\frac{f^j_{p^i}}{f^i_{p^i}}\right). \tag{11.30}$$

Define the N-vector z and the $N \times N$ matrix A by

$$z = \left(\frac{[q^1]^2 k^1_{x_1}}{q^1_{p^1}}, \frac{[q^2]^2 k^2_{x_2}}{q^2_{p^2}}, \ldots, \frac{[q^N]^2 k^N_{x_N}}{q^N_{p^N}}\right), \quad A = I + J_F,$$

[1] The result is only true if the derivatives dk^i/dx_i have the same sign for all firms. Otherwise, no results can be obtained from (11.29). Mixed cases can be handled if we assume separability with respect to p^i in the term $q^i(p^1, p^2, \ldots, p^N)$ on the right-hand side of (11.29):

$$\dot{x}_i(t) = k^i(x_i(t))g^i(p^i(t))h^i(p^1(t), \ldots, p^{i-1}(t), p^{i+1}(t), \ldots, p^N(t)).$$

Separability obtains if the functions q^i are isoelastic or exponential. With these dynamics one can show that an optimal p^i increases (decreases) if dk^i/dx_i is positive (negative).

in which I is the unit matrix and J_F is a Jacobian matrix. Let the functions F^i be given by $F^i(p^1, p^2, \ldots, p^N) = f^i/f^i_{p^i}$ and introduce the vector $F = (F^1, F^2, \ldots, F^N)$. The matrix J_F is the Jacobian of F and is defined by

$$J_F = \begin{pmatrix} F^1_{p^1} & \ldots & F^1_{p^N} \\ \vdots & \ddots & \vdots \\ F^N_{p^1} & \ldots & F^N_{p^N} \end{pmatrix}.$$

Using (11.24), (11.25), and (11.27) it is straightforward to show that (11.30) can be rewritten as

$$A\dot{p}(t) = z, \tag{11.31}$$

in which $\dot{p}(t) = (\dot{p}^1(t), \dot{p}^2(t), \ldots, \dot{p}^N)$.

The next step is to show that A is a dominant diagonal matrix. With this at hand we can invoke a result about such a matrix, leading us to the desired conclusion. Define the elasticity ϵ_i by

$$\epsilon_i = -\frac{f^i_{p^i p^i} p^i}{f^i_{p^i}}$$

and denote by a_{ii} and a_{ij} the diagonal and off-diagonal elements, respectively, of the matrix A. The diagonal element a_{ii} can be written as

$$a_{ii} = 1 + F^i_{p^i} = 2 - \frac{\epsilon_i}{\eta_i}. \tag{11.32}$$

Using the first inequality in (11.23) it is easy to verify that $a_{ii} > 0$ for all $i = 1, 2, \ldots, N$. Using the fact that

$$a_{ij} = F^i_{p^j} \tag{11.33}$$

and the second inequality in (11.23) shows that $a_{ij} \leq 0$ for all $i, j = 1, 2, \ldots, N$ such that $i \neq j$. Using the assumptions in (11.22) shows that $f^i_{p^i} < -\sum_{j \neq i} f^i_{p^j}$ or, equivalently,

$$1 > -\frac{1}{f^i_{p^i}} \sum_{j \neq i} f^i_{p^j}. \tag{11.34}$$

Differentiate F^i with respect to p^j to obtain

$$F^i_{p^j} = \frac{f^i_{p^j}}{f^i_{p^i}} \left[1 - \frac{f^i_{p^i p^j} f^i}{f^i_{p^i} f^i_{p^j}} \right]. \tag{11.35}$$

Define

$$\delta_{ij} = -\frac{f^i_{p^i p^j} p^i}{f^i_{p^j}}$$

and observe that $\delta_{ij} \geq 0$ by the assumptions in (11.22). We can rewrite (11.35) as

$$F^i_{p^j} = \frac{f^i_{p^j}}{f^i_{p^i}} \left[1 - \frac{\delta_{ij}}{\eta_i}\right] \tag{11.36}$$

and conclude that $\partial F^i / \partial p^j \leq 0$ by the second inequality in (11.23). Using (11.33), (11.34), and (11.36) yields

$$0 \leq -\sum_{j \neq i} F^i_{p^j} = -\frac{1}{f^i_{p^i}} \sum_{j \neq i} f^i_{p^j} \left[1 - \frac{\delta_{ij}}{\eta_i}\right] < 1. \tag{11.37}$$

Finally, using (11.32), (11.33), and (11.37) shows that for all i it holds that

$$|a_{ii}| = 1 + F^i_{p^i} > -\sum_{j \neq i} F^i_{p^j} = \sum_{j \neq i} |a_{ij}|.$$

From this expression we conclude that A is a dominant diagonal matrix. Using this result, and our hypothesis that all derivatives dk^i / dx_i have the same sign, we can apply to system (11.31) a theorem of dominant diagonal matrices. The theorem states the following:

Suppose that $A = (a_{ij})$ is an $N \times N$ matrix with $a_{ii} > 0$ for all i and $a_{ij} \leq 0$ for all $i \neq j$. There exists a unique $y \in \mathbb{R}^N_+$ such that $Ay = c$ for any $c \in \mathbb{R}^N_+$, if and only if A is a dominant diagonal matrix.

Apply this result to (11.31) to see that for all $t \in [0, T]$ there exists a vector of time-derivatives of p^i, being the unique solution of (11.31), and such that each component of the vector is positive (negative), whenever all derivatives dk^i / dx_i are positive (negative). ■

It remains to be shown that the solution is indeed optimal. One could try to verify that the maximized Hamiltonians are concave in the state vector, but, since the model does not specify the functions involved, it is more than doubtful that we would arrive at a definitive result concerning the signs of the principal minors of the Hessian matrices. Also note that the pricing strategies are only implicitly given by the necessary optimality conditions. The upshot is that our calculations would involve third order partial derivatives of the demand functions with respect to prices and no economic knowledge is available to assess the signs of such derivatives. In an exercise we ask the reader in a simplified setup to provide a proof of optimality by using an approach other than concavity of the maximized Hamiltonians.

11.3 Advertising goodwill accumulation

In this section we study a dynamic relationship between a manufacturer and a single retailer in a vertical distribution channel. (The reader could also think of the retailer as being the representative of a coordinated group of retailers, or the retailer could be a representative store of a corporate chain.) We suppose that the two members of the channel make their decisions independently and noncooperatively. This section employs another type of dynamic marketing model than those used in sections 11.1 and 11.2, namely, an advertising goodwill model. Such a model has the same mathematical representation as a capital accumulation model (cf. chapter 9, in particular section 9.5). The hypothesis of this section is that the retail firm has a stock of advertising goodwill which summarizes the effects of past consumer advertising efforts.

The decision variable of each channel member is the consumer advertising effort. Denote by $a_M(t)$ and $a_R(t)$ the two firms' respective consumer advertising effort rates at time $t \in [0, \infty)$. Suppose that consumer advertising adds to the accumulation of a stock of goodwill $G(t)$ at the retail level and that the size of this stock influences current consumer demand. Assuming away any inventories in the channel, consumer demand equals retail sales, which in turn equal the manufacturer's production rate. Let $q(t)$ denote the rate of consumer demand at time t and let $p_R(t)$ be the unit retail price. Consumer demand is supposed to depend on the retail price and the stock of advertising goodwill. Thus, consumer demand depends only indirectly on the advertising efforts, through the impact of these efforts on the goodwill stock. The demand function $q(t) = f(p_R(t), G(t))$ decreases in $p_R(t)$, increases in $G(t)$, and is multiplicatively separable: $q(t) = g(p_R(t))h(G(t))$. The price function $g(p_R(t))$ is linear whereas the goodwill function $h(G(t))$ is quadratic and concave. Thus, the demand function is given by

$$q(t) = \max\{0, [\alpha - \beta p_R(t)][g_1 G(t) - (g_2/2)G(t)^2]\}$$

in which g_1, g_2, α, β are positive constants. This specification means that, for $G \in [0, g_1/g_2]$, demand is shifted upward (downward) by an increase (decrease) in the stock of goodwill. Note that there are decreasing returns to the effects of goodwill on demand (because the function h is concave).

The manufacturer contributes to the accumulation of retailer goodwill by advertising directly his brand to the consumers. The purpose is to increase the consumers' awareness of his brand, which should translate into increasing sales at the retail level and larger quantities ordered by the retailer. The marketing of automobiles is an example of a situation in which both channel members aim at stimulating the final consumer demand through their individual advertising efforts.

For the dynamics of the stock of goodwill we assume (see also example 7.1 and section 9.5)

$$\dot{G}(t) = a_M(t) + a_R(t) - \delta G(t), \quad G(0) = G_0 \geq 0, \tag{11.38}$$

which has an obvious similarity to the standard capital accumulation model of chapter 9. With the dynamics (11.38), the state constraint $G(t) \geq 0$ is satisfied for all $t > 0$ when advertising rates are lower bounded by zero. Assume that both firms face the same quadratic advertising cost function: $wa_j^2/2, j \in \{M, R\}$.[2] Denote by $p_M(t)$ the transfer price charged by the manufacturer per unit of sales to the retailer and let $c > 0$ be the manufacturer's unit production cost. Assume that $\alpha > \beta c$. The objective functional of each firm is its discounted profit stream over an infinite planning horizon. With a common discount rate $r > 0$ we have

$$\begin{aligned} J^M = \int_0^\infty e^{-rt}\{&[p_M(t) - c][\alpha - \beta p_R(t)][g_1 G(t) \\ &-(g_2/2)G(t)^2] - (w/2)a_M(t)^2\}\, dt, \end{aligned} \tag{11.39}$$

$$\begin{aligned} J^R = \int_0^\infty e^{-rt}\{&[p_R(t) - p_M(t)][\alpha - \beta p_R(t)][g_1 G(t) \\ &-(g_2/2)G(t)^2] - (w/2)a_R(t)^2\}\, dt. \end{aligned} \tag{11.40}$$

Neither the transfer price nor the retail price enter into the dynamics (11.38). This means that the prices are determined by maximization of the integrands of the payoff functions at any point (G, t). Owing to the fact that the model is autonomous, the resultant optimal prices are constant. Thus, the firms could as well determine their prices once and for all before playing the advertising game. To settle the pricing issue, the linearity in (11.39) with respect to the transfer price implies that the manufacturer would like to charge the highest possible transfer price, subject to $c \leq p_M \leq \alpha/\beta$. A feasible retail price, on the other hand, is constrained by $p_M \leq p_R \leq \alpha/\beta$. Let us assume the following.

(i) The manufacturer charges a transfer price which is a markup on his unit production cost. He would like to charge his maximal markup

[2] The solution to the problem with quadratic advertising costs and linear dynamics is equivalent to the solution of a problem with linear advertising costs and square root advertising terms in the dynamics.

price, which equals $p_M = c + (\alpha - \beta c)/\beta = \alpha/\beta$ but is restricted by the retailer participation constraint $p_M - c \le p_R - p_M$. Such a constraint means that the retailer will not carry the manufacturer's product if the retailer does not get a minimum instantaneous unit profit (net of advertising expenditures). Among the many choices of such a minimum level, we have chosen the instantaneous unit profit of the manufacturer (net of advertising expenditures).

(ii) For any feasible transfer price, the retailer sets a retail price so as to maximize his instantaneous profit $(p_R - p_M)(\alpha - \beta p_R)$. A maximizing price is easily found to be

$$\tilde{p}_R = (\alpha + \beta p_M)/(2\beta). \tag{11.41}$$

Using the participation constraint and (11.41) yields $p_M \le (\alpha + 2\beta c)/(3\beta)$. Since the manufacturer wishes the transfer price to be as high as possible, he chooses this price as

$$\tilde{p}_M = \frac{\alpha + 2\beta c}{3\beta} = c + \frac{\alpha - \beta c}{3\beta}. \tag{11.42}$$

The term $(\alpha - \beta c)/(3\beta) > 0$ is the manufacturer's markup on his unit production cost. Choosing the transfer price as in (11.42) has the plausible implications $\tilde{p}_M > c$, $\alpha - \beta\tilde{p}_R > 0$, and $\tilde{p}_R > \tilde{p}_M$.

We conclude that with retail and transfer prices given by (11.41) and (11.42), respectively, it holds that

$$(\tilde{p}_M - c)(\alpha - \beta\tilde{p}_R) = (\tilde{p}_R - \tilde{p}_M)(\alpha - \beta\tilde{p}_R) = \theta := (\alpha - \beta c)^2/(9\beta) \tag{11.43}$$

and we see that the game is symmetric (since advertising cost functions are identical). We look for a stationary and symmetric Markov perfect Nash equilibrium of our advertising game. Stationarity of strategies means that each firm's advertising strategy is determined as a function of the goodwill stock only: $a_j(G), j \in \{M, R\}$. From chapter 4 we have the following sufficient condition for symmetric strategies: $a_M(G) = a_R(G) = a(G)$ to yield a stationary Markov perfect Nash equilibrium. There exists a unique and absolutely continuous solution $G : [0, \infty) \mapsto \mathbb{R}_+$ of the initial value problem $\dot{G}(t) = 2a(G(t)) - \delta G(t)$, $G(0) = G_0 \ge 0$. There exist continuously differentiable value functions $V_j(G)$, $G \ge 0$, such that the HJB equations

$$\begin{aligned} rV_M(G) = \max\Bigl\{&\theta\Bigl[g_1 G - \frac{g_2}{2}G^2\Bigr] - \frac{w}{2}a_M^2 \\ &+ \frac{dV_M(G)}{dG}[a_M + a_R(G) - \delta G]\Bigm| 0 \le a_M\Bigr\}, \end{aligned} \tag{11.44}$$

$$rV_R(G) = \max\left\{\theta\left[g_1 G - \frac{g_2}{2}G^2\right] - \frac{w}{2}a_R^2 + \frac{dV_R(G)}{dG}[a_M(G) + a_R - \delta G]\,\Big|\, 0 \le a_R\right\} \quad (11.45)$$

are satisfied for all feasible G, and either V_j is a bounded function or V_j is lower bounded and $\limsup_{t\to\infty} e^{-rt}V_j(G(t)) \le 0$.

To perform the maximizations in (11.44) and (11.45), note that the maximands are concave in a_j. This follows from the convexity of the advertising cost function and the linearity of the dynamics. Writing $dV_j(G)/dG = V_j'(G)$ we see that interior candidates are

$$a_j(G) = \frac{1}{w}V_j'(G) > 0 \qquad \text{if} \qquad V_j'(G) > 0. \quad (11.46)$$

Invoking symmetry, we write $V_M'(G) = V_R'(G) = V'(G)$. Insert from (11.46) into the curly brackets on the right-hand sides of the two HJB equations in (11.44) and (11.45) to obtain a single equation for $V(G)$:

$$\theta\left(g_1 G - \frac{g_2}{2}G^2\right) + \frac{3}{2w}[V'(G)]^2 - \delta G V'(G) = rV(G). \quad (11.47)$$

Because of the linear-quadratic structure of the game, we conjecture that a value function satisfying (11.47) is quadratic, i.e.,

$$V(G) = \frac{u}{2}G^2 + EG + F, \quad (11.48)$$

in which u, E, and F are constants, owing to the stationarity assumption. Substitute the conjectured value function from (11.48), as well as its derivative $V'(G) = uG + E$, into (11.47) and collect terms corresponding to G^0, G^1, and G^2. To satisfy the HJB equation we determine the three constants in (11.48) as

$$u = \frac{w(2\delta + r)}{6} \pm \frac{w}{6}\sqrt{(2\delta + r)^2 + (12/w)\theta g_2}, \; E = \frac{\theta w g_1}{w(\delta + r) - 3u}, \; F = \frac{3E^2}{2wr}. \quad (11.49)$$

In (11.49), obviously $F > 0$. To obtain a globally asymptotically stable steady state we choose the u-solution with the minus sign. In that case we have $u < 0$ and then $E > 0$.

The above derivations are correct if and only if the advertising rate is positive, i.e., $a(G) = (1/w)V'(G) > 0$. Since $V'(G) = uG + E$, our solution to the HJB equation is valid for $G < -E/u =: \bar{G}$ where $\bar{G} > 0$. Define $\bar{V} = V(\bar{G}) = -E^2/(2u) + F$. We see that, given that the advertising rate is positive, the strategy is linearly decreasing in the goodwill stock and is determined by

$$a(G) = \frac{u}{w}G + \frac{E}{w}. \tag{11.50}$$

The intuition of this strategy is that advertising should be positive (but decreasing in the goodwill stock) whenever the goodwill stock does not exceed the threshold $\bar{G}$. Note that the derivative $V'(G) > 0$ has the interpretation of a shadow price of the goodwill stock. Since this shadow price is positive, goodwill is a good stock and it pays to add to the stock, that is, to advertise.

By (11.46), (11.48), and (11.49) we have an equilibrium advertising strategy and a value function which is a solution to the HJB equation on the interval $[0, \bar{G})$. It remains to provide an advertising strategy and a solution of the HJB equation for $G \in [\bar{G}, \infty)$. Recall that the maximand of the HJB equation is concave in a. This means that if $V'(G) < 0$ we should put advertising equal to zero. We conjecture that this is the case if the goodwill stock exceeds the threshold level $\bar{G}$. Thus, the shadow price of the goodwill stock must be negative for $G \in [\bar{G}, \infty)$ to warrant a strategy of zero advertising on this interval. The intuition of a zero advertising policy is that, for goodwill stocks exceeding the threshold $\bar{G}$, advertising is not worthwhile since the goodwill stock is already sufficiently high. Recall that advertising decreases in G for all G below the threshold.

With zero advertising the HJB equation becomes

$$\theta\left(g_1 G - \frac{g_2}{2}G^2\right) - \delta G V'(G) = rV(G). \tag{11.51}$$

To have a continuous value function, (11.51) must be solved with the boundary condition $V(\bar{G}) = \bar{V}$. Dividing in (11.51) by δG yields a linear differential equation which has the solution

$$\begin{aligned} V(G) = &-\frac{G^{-r/\delta}}{2(\delta+r)(2\delta+r)}\left\{\theta g_2(\delta+r)\left[G^{(2\delta+r)/\delta} - \bar{G}^{(2\delta+r)/\delta}\right]\right. \\ &\left. -2\theta g_1(2\delta+r)\left[G^{(\delta+r)/\delta} - \bar{G}^{(\delta+r)/\delta}\right] - 2(\delta+r)(2\delta+r)\bar{V}\bar{G}^{r/\delta}\right\}. \end{aligned} \tag{11.52}$$

Taking the derivative in (11.52) yields

$$\begin{aligned} V'(G) = &-\frac{1}{2(\delta+r)(2\delta+r)}\left\{\theta g_2(\delta+r)\left[2G + \bar{G}^{(2\delta+r)/\delta}\frac{r}{\delta}G^{-r/\delta-1}\right]\right. \\ &- 2\theta g_1(2\delta+r)\left[1 + \bar{G}^{(\delta+r)/\delta}\frac{r}{\delta}G^{-r/\delta-1}\right] \\ &\left. + 2(\delta+r)(2\delta+r)\bar{V}\bar{G}^{r/\delta}\frac{r}{\delta}G^{-r/\delta-1}\right\}. \end{aligned} \tag{11.53}$$

From (11.48) we obtain $\lim_{G \nearrow \bar{G}} V'(G) = 0$. After some tedious calculations, (11.53) yields $\lim_{G \searrow \bar{G}} V'(G) = 0$. In these calculations we exploit the fact that the coefficient u in (11.48) is chosen as the negative root of $(3/w)u^2 - (2\delta + r)u - \theta g_2 = 0$.

Equations (11.48) and (11.53) show that, on their respective intervals of definition, both value functions $V(G)$ are continuously differentiable. We conclude that by (11.48) and (11.52) we have a solution $V(G)$ to the HJB equation for $G \in [0, \infty)$. This solution is by construction continuous and we have just seen that it is continuously differentiable.

Finally, we need to show that the solution to (11.51) satisfies $V'(G) < 0$ for $G > \bar{G}$. A direct evaluation of the derivative in (11.53) does not readily provide an answer. Instead we differentiate with respect to G in (11.51) to obtain

$$(\delta + r)V'(G) + \delta G V''(G) = \theta(g_1 - g_2 G). \tag{11.54}$$

For $G > \bar{G}$ the right-hand side of (11.54) is negative. This follows from the fact that $\bar{G} > g_1/g_2$, which can be established by using the definition $\bar{G} = -E/u$, substituting E from (11.49), and recalling $3u^2/w - (2\delta + r)u - \theta g_2 = 0$. Obviously, $V'(G) = 0$ is not a solution of (11.54). Solutions with $V'(G) > 0$ either do not satisfy (11.54) or fail to be continuously differentiable. Solutions with $V'(G) < 0$ and $V''(G) \geq 0$ are not continuously differentiable. The only remaining solution is one having $V'(G) < 0$ and $V''(G) < 0$. Combining this result with (11.48) shows that the value function is strictly concave for $G \in [0, \infty)$.

11.4 Further reading

This section gives some pointers to the literature on differential game models in marketing competition, particularly in the areas of pricing and advertising. Hanssens et al. [115] cover marketing models of competition in general. Rao [195] and Moorthy [181] survey game theoretic models in marketing. Reviews of differential games of pricing are found in Jørgensen [133], Kalish [143], Rao [194], and Moorthy [181]. Differential game models of advertising competition are reviewed in Jørgensen [132], Erickson [78, 80], Moorthy [181], and Feichtinger et al. [86]. Diffusion models in marketing competition are surveyed in Dolan et al. [65] and Mahajan et al. [167]. Eliashberg and Steinberg [74] deal with models in the interface between marketing and production.

The advertising game of section 11.1 appeared in Sorger [220]. The original Lanchester game is analysed as an advertising differential game in Case [16], see also Erickson [77, 78, 79]. Recently there has also been an interest in using the Lanchester model in empirical studies of advertising competition, cf. Erickson [79], Chintagunta and

Vilcassim [25], and Chintagunta and Jain [22]. Chintagunta and Jain [22] employ Sorger's modification [220] of the Lanchester model in an econometric model. Fruchter and Kalish [102] reconsider the problem studied in Chintagunta and Vilcassim [25] and Erickson [79]. Other applications of the Lanchester model are in Mesak and Darrat [178], who approach the problem of advertising pulsing strategies (alternating between high and low levels of advertising). See also Feichtinger et al. [86].

The pricing model of section 11.2 can be found in Dockner and Jørgensen [54]. Other pricing models are Feichtinger and Dockner [84], who demonstrate various aspects of the phase diagram solution procedure, and Wernerfelt [240], where the dynamics describe the flow of market shares between firms as a function of price differentials (see also Wernerfelt [241]). Wernerfelt [242] suggests a differential game model of an oligopoly in which firms choose prices subject to cost experience curves and consumer brand loyalty. Wernerfelt [243] uses a stochastic dynamic game approach to examine two types of brand loyalty (one being caused by time lags in the awareness of the product's values, the other due to consumers' switching costs). Dockner and Gaunersdorfer [52] study a dynamic duopoly using a diffusion model in which the current sales rate of a firm depends on the remaining market potential and a price function. Eliashberg and Jeuland [71] assume that a firm introduces a new durable good and has a monopoly until another firm enters the market. The after-entry dynamics are similar to those in Dockner and Gaunersdorfer [52]. Xie and Sirbu [246] use a diffusion model to study dynamic price competition among networks. Positive network externalities are involved if a consumer's utility of consumption increases with the number of other users. This is important in, for instance, telecommunication services.

Sales response models specify the rate of change of a firm's sales as a function of current sales and the current values of the marketing instruments of the players. Examples of such models in advertising competition are Jørgensen [131], and Feichtinger and Dockner [83].

A series of differential game models of advertising have used diffusion model dynamics. Teng and Thompson [228] propose a rather elaborate model but have to resort to numerical simulations. Thompson and Teng [230] assume market share dynamics that are a combination of diffusion and sales response models. The assumption is that one firm acts as a price-leader. Dockner and Jørgensen [55] consider an oligopoly differential game of advertising. A general sales function is specified in which current sales depend on the vectors of current advertising and accumulated sales (a similar approach to the one in section 11.2). Horsky and Mate [128] study advertising competition in a two-firm, durable-goods producing industry, using a stochastic process model of the diffusion of a new product.

The game in section 11.3 builds upon the model of Nerlove and Arrow [187], which has a close relationship to the capital accumulation dynamics of chapter 9 (see also Jørgensen and Zaccour [136]). Chintagunta [20] uses the Nerlove–Arrow model with diminishing returns to advertising effort. Sales of a firm depend in a quadratic way on its own goodwill and the rival's goodwill, including a product term of the two goodwill levels. Chintagunta and Jain [21] deal with the manufacturer–retailer relationship in a vertical channel. As opposed to the game in section 11.3, there are two goodwill levels (each evolving according to the Nerlove–Arrow specification). Retail sales depend nonlinearly on the goodwill levels. For a special case of the sales function, a Markovian equilibrium can be derived. Advertising goodwill accumulation is also studied in Fershtman [87] and Fershtman et al. [91], using the Nerlove–Arrow model in an oligopolistic version and extended with pricing. As in section 11.3, prices do not enter into the dynamics and the firms play a static pricing game at each instant of time. Chintagunta et al. [24] study a market for experience goods in which accumulated consumption experience of a brand is an important determinant of the firm's sales or market share. The dynamics are the Nerlove–Arrow model in which accumulated aggregate consumption experience plays the role of the goodwill stock. Chintagunta and Rao [23] use a model of the Nerlove–Arrow type to represent consumer preferences that change over time. Essentially, the setup is stochastic but is transformed into a deterministic differential game. Fornell et al. [97] also exploit the concept of consumer experience. Brand specific consumer experience is a state variable and oligopolists set advertising and promotional expenditures. In Thépot [229] the Nerlove–Arrow advertising model is extended to include pricing and investments in productive capital, using a path-connecting procedure from optimal control theory (for details of this procedure, see, Van Hilten et al. [237]).

Some papers are concerned with more than one marketing instrument. Dockner and Feichtinger [47] study an oligopoly differential game in which the setup is similar to the sticky prices model of section 10.1. Gaugusch [109] studies a differential game between one firm which has price as its decision variable (and does not advertise), and another firm which decides on its advertising expenditure (and has a constant price). Chintagunta and Vilcassim [26] assume that each firm in a duopoly has K marketing instruments. The model is of the Lanchester type, as generalized by Erickson [79], and the paper contains an empirical study for the case of two prescription drugs. See also Chintagunta and Vilcassim [25].

The interface between production, capacity planning, inventory, and marketing has been studied in differential games. (For the case of one-firm planning, see Eliashberg and Steinberg [74].) In Gaimon [106], two firms determine their prices and production rates as well as levels and compositions of capacity. The composition of capacity can be modified by purchasing new capacity that reduces unit costs. Gaimon [107] provides a survey of optimal control and differential game models in the pricing–production area. Eliashberg and Steinberg [73] study a differential game in which one firm is a production-smoother (having convex production and linear holding costs) and the other firm is an order-taker (holding no inventory and facing a linear production cost).

An area that involves production, inventories, and marketing is distribution channel competition which, in general, could involve both competition between rival channels and competition within a channel. Formally, the effects of the expenditures on distribution are similar to awareness advertising and results from advertising competition can be implemented. Jørgensen [134] and Eliashberg and Steinberg [72] deal with dynamic pricing, production, purchasing, and inventories in a noncooperative channel situation.

In some differential games in marketing, the outcomes are qualitatively similar to those known from monopoly. Mahajan et al. [167] state that a good part of our intuition concerning optimal marketing policies seems to carry over from monopoly to oligopoly. The reader should be aware that it may be a model's lack of product interdependence or strategic interdependence, or both, that imply that oligopolistic firms should act like monopolists. In general it does not seem advisable to believe that optimal marketing policies by and large carry over from monopoly to oligopoly.

The result on dominant diagonal matrices that was used in section 11.3 can be found in Takayama [227].

11.5 Exercises

1. Consider the advertising game in section 11.1 and assume that firms are identical in the sense that the parameters r, c, q, and S are the same for both firms. That is, we have a symmetric game. For the open-loop Nash equilibrium, solve the costate equations explicitly and study the monotonicity properties of the costate variables.
2. Consider the advertising game of section 11.1 and refer to section 8.2. Change the dynamics of the advertising game such that market share x evolves according to the stochastic differential equation

$$dx(t) = \left[u^1(t)\sqrt{1-x(t)} - u^2(t)\sqrt{x(t)}\right] dt + \sigma(x(t))\, d\omega(t),$$

where $\omega(t)$ represents a white noise stochastic process and the function $\sigma(x)$ satisfies $\sigma(x) \geq 0$ for $x \in [0, 1]$ and $\sigma(0) = \sigma(1) = 0$. The above stochastic differential equation intends to model a situation where market share is not completely determined by the firms' advertising efforts. Besides these efforts there are exogenous stochastic disturbances caused, for example, by unpredictable consumer behaviour. Is it correct to conclude that the Markovian Nash equilibrium of the deterministic model also constitutes a Markovian Nash equilibrium of the stochastic game?

3. Consider a differential duopoly game with dynamics

$$\dot{S}_i(t) = k[\log A_i(t) - \log A_{3-i}(t)],\ S_i(0) = S_0 > 0\,,\ i = 1, 2,$$

in which $S_i(t)$ denotes the sales rate of firm i and $A_i(t)$ is its advertising rate. Suppose that the total sales $S_1(t) + S_2(t)$ are constant, equal to $M = 2S_0$. The dynamics reflect a hypothesis that the greater a firm's advertising, the greater the flow of sales from its competitor. Letting q_i denote the unit margin of firm i, payoffs are given by

$$J^i = \int_0^T e^{-r_i t}[q_i S_i(t) - A_i(t)]\, dt.$$

Firm i determines its advertising rate over the horizon in order to maximize its payoff. Find a Nash equilibrium with open-loop advertising strategies and show that both equilibrium advertising rates are decreasing over time and are equal to 0 at $t = T$.

4. Consider the game in exercise 3 and replace the dynamics by

$$\dot{S}_1(t) = g(A_1(t), A_2(t)),$$

in which $S_1(t)$ is the sales rate of firm 1. Firm 2 gets the remaining sales, $M - S_1(t)$, where $M = 2S_0$. The function g satisfies

$$\frac{\partial g}{\partial A_1} > 0, \frac{\partial g}{\partial A_2} < 0, \frac{\partial^2 g}{\partial A_1^2} < 0, \frac{\partial^2 g}{\partial A_2^2} > 0, \frac{\partial^2 g}{\partial A_1 \partial A_2} \leq 0.$$

Extend the payoffs with salvage value functions $e^{-r_i T}\sigma_i S_i(T)$, where σ_i is the valuation of a unit of terminal sales of firm i. Find a Nash equilibrium in open-loop advertising strategies. Hint: a specific form of the right-hand side of the dynamics could be

$$g(A_1, A_2) = A_1^\alpha A_2^\beta\,,\ 0 < \alpha < 1\,,\ \beta < 0.$$

Consider also simplifications such as $r_i = 0$, $S_i = 0$, and $\partial^2 g/(\partial A_1 \partial A_2) = 0$.

5. Consider an N-firm oligopoly, let $S_i(t)$ denote the cumulative sales by time t of firm $i \in \{1, 2, \ldots, N\}$ and define $S(t) = \sum_{i=1}^{N} S_i(t)$. Let $A_i(t)$ denote firm i's advertising rate. The differential game has the diffusion dynamics

$$\dot{S}_i(t) = [a + b \log A_i(t) + dS(t)][M - S(t)], \quad S_i(0) = S_{i0} \geq 0,$$

which means that a firm can stimulate its sales through advertising (but subject to decreasing returns) and that demand learning effects (imitation) are industry-wide. (If these effects were firm-specific we would have S_i instead of S in the brackets on the right-hand side of the dynamics.) Payoffs are given by

$$J^i = \int_0^T e^{-r_i t}[(p_i - c_i)\dot{S}_i(t) - A_i(t)]\, dt,$$

in which prices and unit costs are constant. Determine a Nash equilibrium in open-loop advertising strategies and show that advertising rates are monotonically decreasing over time.

6. Consider the pricing game in section 11.2. Refer to the proof of theorem 3.2 and recall that the idea of the proof is to evaluate directly the difference between the objective functional when using an optimal control and the objective functional when using an arbitrary, feasible control. To establish that the first objective value is not smaller than the second, two conditions are used. First, the maximum condition is employed to produce one inequality. Second, concavity of the maximized Hamiltonian is invoked to produce a second inequality. Then the desired result follows.

There is an alternative approach in which we do not assume concavity of the maximized Hamiltonian. To simplify, consider the case $N = 2$ and assume that the unit cost of production is constant with respect to x_i. Denote this cost by c_i. Employ the dynamics in the form stated in (11.29) and introduce the plausible assumption that the functions $k^i(x_i)$ in (11.29) are concave. It suffices to do the proof for player 1. The Hamiltonian of player 1 becomes

$$\begin{aligned} H^1(x_1, x_2, p^1, \lambda_1^1, \lambda_2^1, t) = {} & (p^1 - c_1 + \lambda_1^1)k^1(x_1)q^1(p^1, p^2(t)) \\ & + \lambda_2^1 k^2(x_2) q^2(p^1, p^2(t)). \end{aligned}$$

Going through the steps of the proof of theorem 3.2 reveals that we have proved optimality if the following inequality is satisfied for all t:

$$H^1(x_1^*(t), x_2^*(t), p^{1*}(t), \lambda_1^1(t), \lambda_2^1(t), t) - H^1(x_1(t), x_2(t), p^1(t), \lambda_1^1(t), \lambda_2^1(t), t)$$
$$\geq [\dot{\lambda}_1^1(t) - r_1\lambda_1^1(t)][x_1(t) - x_1^*(t)] + [\dot{\lambda}_2^1(t) - r_1\lambda_2^1(t)][x_2(t) - x_2^*(t)].$$

Now prove that the inequality is satisfied. (If we replace on the left-hand side of this inequality $H^1(x_1^*(t), x_2^*(t), p^{1*}(t), \lambda_1^1(t), \lambda_2^1(t), t)$ by $H^{1*}(x_1^*(t), x_2^*(t), \lambda_1^1(t), \lambda_2^1(t), t)$ as well as $H^1(x_1(t), x_2(t), p^1(t), \lambda_1^1(t), \lambda_2^1(t), t)$ by $H^{1*}(x_1(t), x_2(t), \lambda_1^1(t), \lambda_2^1(t), t)$ we get the inequality that follows from concavity of the maximized Hamiltonian in the proof of theorem 3.2.)

12 Differential games in resources and environmental economics

In this chapter we present a number of models in resources and environmental economics in which economic agents (firms or countries) exploit natural resources or the environment in an intertemporal context, taking into account the strategic behaviour of other agents.

We begin with a simple model of exploitation of a common-property, nonrenewable resource such as an oil field. We compare the benchmark cooperative solution with a noncooperative open-loop Nash equilibrium and a Markov perfect Nash equilibrium. In the open-loop version of the game, we show that the nature of the solution depends on how we restrict the set of feasible strategies available to each agent. This is an important issue from the modelling point of view, especially in the context of common property resources, because this context dramatically highlights the interdependence among agents, not only in terms of payoffs but also in terms of what one agent can do given the actions of others.

After a thorough discussion, we consider some variations of the basic model of common-property nonrenewable resources: the so-called doomsday problem, and the problem where utility depends directly on both the stock of the resource and the flow of consumption.

Renewable resources such as fish stocks and forests are considered next, at first in the standard format of simultaneous choice of Markovian strategies. The game is then modified to allow for history-dependent strategies (such as trigger strategies), and for hierarchical moves, thus illustrating a Stackelberg leadership formulation of a fishery game. The final application in this chapter is a model of transboundary pollution, for which a Nash equilibrium consisting of linear Markovian strategies is computed and the equilibrium outcome is compared with the cooperative outcome.

12.1 Nonrenewable resources

In this section we present a simple model of common-property resources. Consider a natural resource, such as an oil field, that can be exploited simultaneously by N firms, or N countries. The equilibrium path of extraction depends on whether the players (firms or countries) cooperate or not. Let us begin by examining the benchmark case, where they cooperate.

12.1.1 The cooperative case

Let $x(t)$ and $c_i(t)$ denote respectively the stock of the resource and player i's rate of extraction at time t. We assume that $c_i(t) \geq 0$ and that, if $x(t) = 0$, then the only feasible rate of extraction is $c_i(t) = 0$. The transition equation is assumed to be

$$\dot{x}(t) = -\sum_{i=1}^{N} c_i(t).$$

Each player i has a utility function $u(c_i)$, defined for all $c_i > 0$, which is strictly concave and increasing (we allow the case where $u(0) = -\infty$). Utility is discounted at a constant rate $r > 0$, and player i's objective function is

$$\int_0^{\infty} e^{-rt} u(c_i(t))\, dt.$$

We restrict attention to the class of utility functions having a constant elasticity of marginal utility $\eta > 0$. If $\eta \neq 1$, then

$$u(c_i) = \frac{A c_i^{1-\eta}}{1-\eta} + B = v(c_i) + B,$$

and if $\eta = 1$, then

$$u(c_i) = A \ln c_i + B.$$

Here, A is positive and will be set at unity without loss of generality, whereas B is a constant which may be positive, negative, or zero. Since the time horizon is fixed, the constant B has no effect on the optimal path. Therefore, without loss of generality, one could set $B = 0$.[1]

Given that all players have the same utility function, it seems natural to assume that, when they cooperate, the objective is to maximize the sum of

[1] If the time horizon were free, B would play an important role; we do not deal with the free horizon problem in this section.

the integrals of discounted payoffs. Let $c(t)$ denote the rate of extraction of the representative player. The resulting optimal control problem consists in maximizing

$$\int_0^\infty e^{-rt} N u(c(t))\, dt \tag{12.1}$$

subject to

$$\dot{x}(t) = -Nc(t),$$
$$c(t) \geq 0,$$
$$x(0) = x_0 > 0,$$
$$\lim_{t\to\infty} x(t) \geq 0.$$

Notice that the above constraints imply that $x(t) \geq 0$ for all $t \geq 0$. Therefore it is not necessary to add a separate nonnegativity constraint for the state variable. The Hamiltonian function is

$$H(x, c, \psi) = Nu(c) - \psi Nc, \tag{12.2}$$

from which we obtain the conditions

$$u'(c(t)) - \psi(t) \leq 0, \quad c(t) \geq 0, \quad [u'(c(t)) - \psi(t)]c(t) = 0, \tag{12.3}$$

$$\dot{\psi}(t) = r\psi(t). \tag{12.4}$$

Since the Hamiltonian is concave in (c, x), any feasible path $\{c^*(\cdot), x^*(\cdot), \psi^*(\cdot)\}$ that satisfies conditions (12.3) and (12.4) is optimal if it also satisfies the transversality condition

$$\lim_{t\to\infty} e^{-rt}\psi^*(t)[\tilde{x}(t) - x^*(t)] \geq 0, \tag{12.5}$$

where $\tilde{x}(t)$ is any feasible path.[2]

It is easy to show that, along an optimal path, extraction is positive for all t and $\psi(t)$ is positive and rises at the rate of interest: $\psi(t) = \psi(0)e^{rt}$. Conditions (12.3) and (12.4) then imply that

$$\frac{\dot{c}(t)}{c(t)} = -\frac{r}{\eta} \tag{12.6}$$

Condition (12.6) says that marginal utility rises at the rate r. This is known as the Hotelling rule. This condition yields $c(t) = c(0)e^{-rt/\eta}$ and

[2]Readers are referred to the sufficient optimality conditions presented in chapter 3 and are urged to verify that, in all the examples given in the present chapter, the solutions satisfy these conditions.

we can solve for $c(0)$ by noting that an optimal path must exhaust the stock in the sense that $\lim_{t\to\infty} x(t) = 0$. (Clearly, any path that does not exhaust the stock will be dominated by a feasible path that exhausts the stock, giving rise to more consumption over some time interval.) Thus $x_0 = \int_0^\infty Nc(t)\,dt = \eta Nc(0)/r$. The optimal solution is therefore

$$c^*(t) = \frac{rx_0}{\eta N} e^{-rt/\eta} \tag{12.7}$$

and the resulting time path of the state variable is

$$x^*(t) = x_0 e^{-rt/\eta}. \tag{12.8}$$

Notice that the transversality condition (12.5) is satisfied.

From (12.7) and (12.8) it follows that $c^*(t) = rx^*(t)/(\eta N)$, which shows that the optimal control path is generated by the linear Markovian strategy $\phi(x) = rx/(\eta N)$. The value of the integral (12.1) is

$$V(x_0) = \frac{N}{1-\eta}\left(\frac{x_0}{N}\right)^{1-\eta}\left(\frac{r}{\eta}\right)^{-\eta} + \frac{NB}{r}$$

provided $\eta \neq 1$. If $\eta = 1$, then

$$V(x_0) = \left[-1 + \ln\left(\frac{rx_0}{N}\right)\right]\frac{N}{r} + \frac{NB}{r}.$$

Note that if $\eta \geq 1$ then $\lim_{x_0 \to 0} V(x_0) = -\infty$. The payoff for player i is

$$V_i(x_0) = V(x_0)/N. \tag{12.9}$$

12.1.2 The noncooperative case

We now turn to the noncooperative solution. An interesting question is whether the cooperative solution described above can be achieved as a Nash equilibrium of a noncooperative game. It turns out that the answer depends crucially on the values of η and N, and also on what set of strategies is available to the players. Let us begin the discussion by considering open-loop strategies.

In the open-loop case, each player must choose at the outset a time path of extraction. These choices are made simultaneously and noncooperatively.

If $\eta \geq 1$, then $\lim_{c\to 0} u(c) = -\infty$, and thus zero consumption over any nondegenerate time interval would give a payoff of minus infinity. In this case the cooperative solution is achievable as an open-loop Nash equilibrium of the noncooperative game, because all players will ensure that the resource does not become exhausted in finite time, and the sufficient conditions in the case of asymptotic exhaustion are the same as those of the cooperative solution. This claim is proved formally in theorem 12.1 below.

If $\eta < 1$, then $\lim_{c \to 0} u(c)$ is finite. Thus, as we will see later, under the common access regime there may exist an incentive for each player to exhaust the resource stock in finite time, as everyone tries to grab a bigger share of the total stock. Whether this occurs or not depends on how severely we restrict the choice set of each player. To analyse this problem, some formal definitions are required.

Consider player i's choice. For all $j \neq i$ let a time path $c_j(\cdot)$ be given such that

$$\int_0^\infty \sum_{j \neq i} c_j(t)\, dt < x_0.$$

We can now define what we mean by strictly feasible open-loop replies and weakly feasible open-loop replies, respectively.

Definition 12.1 Player i's open-loop strategy, the time path $c_i(\cdot)$, is a strictly feasible open-loop reply to the $n-1$ time paths $c_j(\cdot)$, $j \neq i$, if and only if

$$\int_0^\infty c_i(t)\, dt + \int_0^\infty \sum_{j \neq i} c_j(t)\, dt \leq x_0.$$

It is a weakly feasible open-loop reply to the $n-1$ time paths $c_j(\cdot), j \neq i$, if and only if

$$\int_0^{T_i} c_i(t)\, dt + \int_0^{T_i} \sum_{j \neq i} c_j(t)\, dt \leq x_0,$$

where T_i is the date beyond which player i's extraction is zero. Formally, one has $T_i = \inf\{t \mid c_i(s) = 0 \text{ for all } s \geq t\}$.

The difference between the two definitions lies in the fact that with a weakly feasible open-loop reply, player i is allowed to frustrate his opponents' plans. To see this, assume that there exists a value T_i such that player i's extraction path, given his opponents' extraction over $[0, T_i]$, implies exhaustion of the resource stock at T_i. According to the definition of a weakly open-loop reply, player i may choose this path if he sets $c_i(t) = 0$ for all $t > T_i$. If, for some $j \neq i$, player j's extraction rate $c_j(t)$ is strictly positive at some $t > T_i$, then player j's plan is frustrated by i because at that time t the resource is already exhausted. In the case of a game with strictly feasible open-loop replies such a situation cannot occur. Some may argue that requiring each player to choose only among strictly feasible open-loop replies effectively removes the essential feature of common property resources. We can now prove the result already mentioned, namely that with strictly feasible open-loop replies the cooperative outcome coincides with an open-loop Nash equilibrium.

Theorem 12.1 *If all players are restricted to strictly feasible open-loop replies then, regardless of the value of η, the cooperative extraction path coincides with the extraction path generated by an open-loop Nash equilibrium of the noncooperative game.*

Proof We show that the optimality conditions for both problems coincide. Consider player i's problem in the noncooperative game. For given time paths $c_j(\cdot)$, $j \neq i$, player i must choose a time path $c_i(\cdot)$ such that

$$\dot{x}(t) = -c_i(t) - \sum_{j \neq i} c_j(t) \tag{12.10}$$

and

$$\lim_{t \to \infty} x(t) \geq 0. \tag{12.11}$$

His objective is to maximize

$$\int_0^\infty e^{-rt} u(c_i(t))\, dt.$$

The Hamiltonian is

$$H_i(x, c_i, \psi, t) = u(c_i) - \psi_i \left[c_i + \sum_{j \neq i} c_j(t) \right],$$

from which we obtain the optimality conditions

$$u'(c_i(t)) - \psi_i(t) \leq 0, \quad c_i(t) \geq 0, \quad [u'(c_i(t)) - \psi_i(t)]c_i(t) = 0, \tag{12.12}$$

$$\dot{\psi}_i(t) = r\psi_i(t). \tag{12.13}$$

These conditions must hold along with (12.10), (12.11), and $x(0) = x_0$.

It is easy to check that, if $c_j(t) = rx_0(\eta N)^{-1}e^{-rt/\eta}$ holds for all $j \neq i$, then the time path $c_i^*(t) = rx_0(\eta N)^{-1}e^{-rt/\eta}$ together with $\psi_i(0) = u'(c_i^*(0))$ and $x^*(t) = x_0 e^{-rt/\eta}$ satisfies conditions (12.10)–(12.13), as well as the transversality condition

$$\lim_{t \to \infty} e^{-rt}\psi_i(t)[\tilde{x}(t) - x^*(t)] \geq 0,$$

where $\tilde{x}(t)$ can be any feasible path that satisfies (12.10) and (12.11). Because the Hamiltonian is concave in the control and the state variables, the above conditions are sufficient for player i's optimal control problem. Finally, the arguments apply for all players. ■

Notice that the above proof makes use of the requirement that players must choose strictly feasible open-loop replies. If weakly feasible open-loop replies are allowed then, as far as player i's choice of a best reply is concerned, condition (12.11) is not required, and must be replaced by

$$\lim_{t \to T_i} x(t) \geq 0, \tag{12.14}$$

or simply $x(T_i) \geq 0$ if T_i is finite. In other words, player i may choose an extraction path $c_i(\cdot)$ such that $c_i(s) = 0$ for all $s \geq T_i$, whereby the possibility that the other players become frustrated, i.e.,

$$\int_{T_i}^{\infty} \sum_{j \neq i} c_j(t)\, dt > x(T_i),$$

is not a matter of concern for player i.

We now show that if $0 < \eta < 1$ (so that $u(0)$ is not minus infinity) and if weakly feasible open-loop replies are allowed then the cooperative extraction path cannot be supported as an open-loop Nash equilibrium of the noncooperative game provided N is sufficiently large.

Theorem 12.2 *Let* $0 < \eta < 1$ *and assume that weakly feasible open-loop replies are permitted. Then each player has an incentive to deviate from the cooperative extraction path if and only if* $1 - \eta > 1/N$.

Proof We wish to show that, if player i's opponents all choose the time path $c_j(t) = r x_0 (\eta N)^{-1} e^{-rt/\eta}$, then player i's best reply is to exhaust the resource stock in finite time if and only if $1 - \eta > 1/N$.

Since $u(c) = v(c) + B$, where $v(0) = 0$, and since condition (12.11) has been replaced by the weaker condition (12.14), player i's optimal control problem consists in maximizing

$$\int_0^{T_i} e^{-rt} v(c_i(t))\, dt + \int_0^{\infty} e^{-rt} B\, dt \tag{12.15}$$

subject to the constraints

$$A(0) = 0, \tag{12.16}$$

$$\dot{A}(t) = c_i(t), \tag{12.17}$$

$$A(T_i) \leq x_0 - \int_0^{T_i} (N-1)c_j(t)\,dt, \tag{12.18}$$

and $c_i(t) \geq 0$. Here $A(t)$ stands for player i's accumulated extraction at time t. Note that while the time horizon is infinite, the time at which player i stops his extraction, T_i, is not fixed; it has to be chosen optimally by player i. We have chosen this formulation to emphasize the fact that for player i the constraint on the terminal value of the state variable depends on T_i: the greater is T_i, the smaller is his total accumulated extraction; see equation (12.18). To obtain the transversality conditions for this problem, we use Hestenes' theorem.[3]

We define

$$z(T_i) = \int_0^{T_i} (N-1)c_j(t)\,dt$$

so that condition (12.18) may be written as

$$h(A(T_i), T_i) := x_0 - z(T_i) - A(T_i) \geq 0.$$

Applying Hestenes' theorem one can show that there exists a multiplier $\mu \geq 0$ associated with this constraint such that

$$\mu[x_0 - z(T_i) - A(T_i)] = 0 \tag{12.19}$$

and such that the transversality conditions (12.22) and (12.23) recorded below are necessarily satisfied by an optimal path. The Hamiltonian function is $H(A, c_i, \xi_i) = v(c_i) + \xi_i c_i$, from which we obtain

$$v'(c_i(t)) + \xi_i(t) \leq 0, \quad c_i(t) \geq 0, \quad [v'(c_i(t)) + \xi_i(t)]c_i(t) = 0, \tag{12.20}$$

$$\dot{\xi}_i(t) = r\xi_i(t), \tag{12.21}$$

$$H(A(T_i), c_i(T_i), \xi_i(T_i)) - \mu z'(T_i) = 0, \tag{12.22}$$

$$\xi_i(T_i) - \mu h_A(A(T_i), T_i) = 0, \tag{12.23}$$

where h_A denotes the partial derivative of h with respect to its first argument. In condition (12.22), $H(A(T_i), c_i(T_i), \xi_i(T_i))$ represents what player i would gain (i.e., the increase in the value of the integral (12.15)) if T_i were to be increased marginally given the assumption that such an extension of time would not reduce his total accumulated extraction. However, that assumption is not valid: an extension of T_i would reduce his total accumulated extraction by the amount $z'(T_i)$. Therefore $H(A(T_i), c_i(T_i),$

[3]See section 12.5 for a reference to Hestenes' theorem.

$\xi_i(T_i)) - \mu z'(T_i) = 0$ is the requirement that a marginal increase in T_i adds zero net gain to player i. Conditions (12.19), (12.22), and (12.23) may be expressed as

$$H(A(T_i), c_i(T_i), \xi_i(T_i)) + \xi_i(T_i)(N-1)c_j(T_i) = 0, \tag{12.24}$$

$$\left[A(T_i) - x_0 - vintop_0^{T_i}(N-1)c_j(t)\,dt\right]\xi_i(T_i) = 0. \tag{12.25}$$

Let us find a candidate solution to this set of conditions. Assume that c_i $(t) > 0$ for all t in the interval $[0, T_i]$. Then (12.20) and (12.21) yield

$$c_i(t) = c_i(T_i)e^{r(T_i - t)/\eta} \qquad \text{for } t \in [0, T_i] \tag{12.26}$$

and $\xi_i(T_i) = -c_i(T_i)^{-\eta} < 0$. Condition (12.22) then becomes

$$\left[\frac{\eta c_i(T_i)}{1-\eta} - \frac{(N-1)x_0 r e^{-rT_i/\eta}}{\eta N}\right] c_i(T_i)^{-\eta} e^{-rT_i} = 0. \tag{12.27}$$

If T_i is finite this yields

$$c_i(T_i) = \frac{(1-\eta)(N-1)x_0 r e^{-rT_i/\eta}}{\eta^2 N}. \tag{12.28}$$

Use (12.28) and (12.26) to get

$$A(T_i) = \int_0^{T_i} c_i(t)\,dt = \frac{(1-\eta)(N-1)x_0[1 - e^{-rT_i/\eta}]}{\eta N}.$$

Substitute this into (12.25) to obtain

$$1 - e^{-rT_i/\eta} = \eta N/(N-1), \tag{12.29}$$

which has a positive and finite solution T_i if and only if $\eta N/(N-1) < 1$. This condition is equivalent to

$$1 - \eta > 1/N. \tag{12.30}$$

Finally let us compute the value of $V_i(x_0)$. From (12.15), (12.26), (12.28), and (12.29) we obtain

$$V_i(x_0) = \frac{B}{r} + \left(\frac{x_0}{N}\right)^{1-\eta}\left(\frac{r}{\eta}\right)^{-\eta}(N-1)^{-\eta}N(1-\eta)^{-\eta}\eta^{\eta}.$$

This value is greater than the payoff in the cooperative solution given by (12.9) if and only if

$$(N-1)^{-\eta}N(1-\eta)^{1-\eta}\eta^{\eta} > 1. \tag{12.31}$$

In lemma 12.1 in the appendix it is shown that (12.31) is implied by (12.30). ■

The intuition behind theorem 12.2 is as follows. Because $0 < \eta < 1$, zero consumption does not cause utility to fall to minus infinity. Therefore, it may pay for a player (and hence, by symmetry, for all players) to deviate from the cooperative extraction path and exhaust the stock at some finite date T_i. In doing so, the player would capture that part of the stock that his opponents plan to consume after T_i. This would improve his payoff provided that the gains he would get from this 'theft' exceed the loss caused by having nothing to consume after T_i when marginal utility of consumption would be high. Clearly the gains are greater, the greater the number of players that are 'robbed'. This is why the cooperative solution can be supported as an open-loop Nash equilibrium only if $N < 1/(1-\eta)$.[4]

In view of theorem 12.2 the following question arises: given that $0 < \eta < 1$ and $N > 1/(1-\eta)$, is there an open-loop Nash equilibrium that implies exhaustion of the stock at some finite time $T > 0$? The answer is no. Owing to limitation of space, we do not offer a formal argument here, but only sketch a proof. First, if a player i chooses a finite T_i, then player j's best reply must involve $T_j \leq T_i$. It follows that in any open-loop Nash equilibrium all players must plan to exhaust the stock at the same finite time, say T. Consider two cases: (a) the consumption path continuously falls to zero (i.e., $\lim_{t \to T} c(t) = 0$), and (b) there is a jump discontinuity at T. Clearly (a) violates the Hotelling rule because marginal utility has a constant elasticity η. As for (b), since each player's planning horizon (as distinct from planned exhaustion date) is infinite, and since each player i knows that his opponents stop extracting at time T, he can improve his payoff by rescheduling his consumption stream so that his consumption is always positive, without fear of being 'robbed' and without changing his total lifetime accumulated consumption.

We now turn our attention to Markov perfect Nash equilibria. We suppose that each player i chooses a rule $\phi_i(x, t)$ that specifies his extraction rate at time t as a function of the size of the stock at time t, $x(t)$, and of t itself:

$$c_i(t) = \phi_i(x(t), t). \tag{12.32}$$

The function ϕ_i is called player i's Markovian strategy. Let S_i denote the set of rules from which player i can choose this strategy. We restrict this set to contain only functions that have the property $\phi_i(0, t) = 0$ for all t. A Markov perfect Nash equilibrium is a strategy profile $\phi^* = (\phi_1^*,$

[4]Notice that in this section, as opposed to chapter 6, we do not allow players to use history-dependent strategies.

$\phi_2^*, \ldots, \phi_N^*)$ such that, for given $\phi_{-i}^* = (\phi_1^*, \ldots, \phi_{i-1}^*, \phi_{i+1}^*, \ldots, \phi_N^*)$ and for every $t_0 \geq 0$ and every $b \geq 0$, the strategy ϕ_i^* maximizes player i's payoff integral

$$\int_{t_0}^{\infty} e^{-rt} u(\phi_i(x(t), t))\, dt$$

subject to the constraints

$$x(t_0) = b,$$
$$\dot{x}(t) = -\phi_i(x(t), t) - \sum_{j \neq i} \phi_j^*(x(t), t)$$

over all strategies ϕ_i from the set S_i. The requirement that the integral be maximized for every $t_0 \geq 0$ and every $b \geq 0$ reflects the idea of subgame perfection.

Let us find a symmetric equilibrium for the case $u(c) = (1 - \eta)^{-1} c^{1-\eta}$, where $0 < \eta < 1$. Suppose that player i's opponents all use the same strategy, denoted by ϕ, and that this strategy is independent of t (i.e., it is a stationary Markovian strategy). To find player i's best reply, we seek a solution to the HJB equation

$$rV_i(x) = \max\{u(c_i) + V_i'(x)[-c_i - (N-1)\phi(x)] \mid c_i \geq 0\}. \tag{12.33}$$

Maximization with respect to c_i and the symmetry of the equilibrium yield the condition

$$c_i = \phi(x) = [V_i'(x)]^{-1/\eta}. \tag{12.34}$$

Substituting this into (12.33), we obtain

$$rV_i(x) = [V_i'(x)]^{(\eta-1)/\eta}\left(\frac{1}{1-\eta} - N\right). \tag{12.35}$$

Notice that, since $u(0) = 0$, the value function cannot be negative for $x \geq 0$. Let $D = (1-\eta)^{-1} - N$ and assume that $D > 0$ to ensure that $V_i(x) \geq 0$. Solve the differential equation (12.35) to get

$$V_i(x) = \left[\frac{x}{1-\eta}\left(\frac{\{r}{D}\right)^{\eta/(\eta-1)} + E\right]^{1-\eta}, \tag{12.36}$$

where E is a constant of integration. The requirement that $\phi(0) = 0$ implies that, if the initial stock is zero, consumption will be zero for ever and, since $0 < \eta < 1$, the value of the integral of the discounted utility stream is zero. It follows that we must set $E = 0$ to meet this requirement, and (12.36) becomes

$$V_i(x) = \left(\frac{r}{D}\right)^{-\eta}\left(\frac{x}{1-\eta}\right)^{1-\eta}. \tag{12.37}$$

From (12.34) and (12.37) we obtain the equilibrium strategy

$$\phi(x) = \frac{rx}{\eta N - (N-1)}. \tag{12.38}$$

Notice that with $\eta \in (0, 1)$ we have to assume $D > 0$ (which means that N is sufficiently small) to obtain a symmetric Markov perfect Nash equilibrium. If $D \leq 0$, then no such equilibrium exists. As theorem 12.2 and the ensuing discussion show, the condition $D > 0$ also ensures that the cooperative equilibrium is supported as an open-loop Nash equilibrium when players are allowed to choose weakly feasible open-loop strategies. Comparing the two equilibria, we see that the Markov perfect Nash equilibrium implies a faster extraction rate and a lower value $V_i(x)$ for all players. The intuition behind this is as follows. In the Markov perfect equilibrium each player knows that, if he tries to be more conservationist, this will encourage other firms to extract more. Therefore there is little incentive to be conservationist. By contrast, in the open-loop equilibrium, given the other players' extraction paths, player i's effort to be conservationist does not induce the others to extract more.

Finally, we consider the case $\eta > 1$ and $u(c) = (1-\eta)^{-1}c^{1-\eta}$. In this case the value function $V_i(x)$ is negative for all positive x. Then we have $rV_i(x) = -[V_i'(x)]^{(\eta-1)/\eta}G$, where $G = N + (\eta - 1)^{-1} > 0$. Hence

$$\left(\frac{r}{G}\right)^{\eta/(\eta-1)}[-V_i(x)]^{\eta/(\eta-1)}\,dx = dV_i(x).$$

Defining $z = -V_i(x) > 0$ we get

$$\left(\frac{r}{G}\right)^{\eta/(\eta-1)}dx = -z^{\eta/(1-\eta)}\,dz.$$

Integration yields

$$V_i(x) = -z = -\left[\frac{x}{\eta-1}\left(\frac{r}{G}\right)^{\eta/(\eta-1)} + E\right]^{1-\eta} < 0.$$

As before, we have to set $E = 0$ to ensure that $\phi(0) = 0$. Hence, we obtain again the equilibrium strategy (12.38).

12.2 Nonrenewable resources: some variations

In this section we consider some variations of the basic model presented in the previous section. First, we look at the case where each player can choose his termination date, that is, the date at which, from his point of view, the game ends. This is sometimes called the doomsday problem. The second variation concerns the case where the utility function depends on the stock as well as on the flow.

12.2.1 The doomsday problem

In the preceding section, we assumed that the time horizon is infinite so that, if the resource stock is exhausted at some finite time T, then after time T each player's utility flow is $u(0)$ (which may be positive or negative), and each player's total payoff is

$$\int_0^T e^{-rt}u(c_i(t))\,dt + \int_T^\infty e^{-rt}u(0)\,dt.$$

We now consider a different formulation. More specifically, we assume that each player i can choose a terminal date T_i such that his utility flow stops at T_i and anything that happens after T_i does not count as far as he is concerned. His total payoff is therefore

$$\int_0^{T_i} e^{-rt}u(c_i(t))\,dt.$$

This type of problem is sometimes referred to as the doomsday problem. This name suggests the gloomy picture of an individual or a community that has to decide on a time beyond which nothing matters any more. A slightly different interpretation is the case of a firm that dissolves itself when its business activities are declared completed. For example, consider a price-taking firm with the profit function $u(c_i) = pc_i - G(c_i) - F$, where c_i denotes the output of the firm, $p > 0$ is the market price of the product, $G(c_i)$ is the total variable cost, and F is the flow of fixed cost (i.e., that part of the cost which is independent of the output level c_i) that the firm must incur as long as it remains in business. An example of F is a municipal tax that a firm must pay independently of its output level.

Assume that there are N identical players (individuals or firms) extracting a common-property resource stock. The utility (or profit) function $u : [0, \infty) \mapsto \mathbb{R}$ is assumed to be continuously differentiable and strictly concave. The cooperative problem is to find a time path of extraction $c(\cdot)$ and a terminal date T so as to maximize the integral

$$\int_0^T e^{-rt}Nu(c(t))\,dt$$

subject to $c(t) \geq 0$ and

$$\dot{x}(t) = -Nc(t), \qquad x(0) = x_0, \qquad x(T) \geq 0. \tag{12.39}$$

Notice that it is possible to choose $T = 0$, in which case the payoff (the value of the integral) is zero. It follows that, for the doomsday problem, $V(x)$ is nonnegative. Thus, if the utility $u(c)$ is negative for all $c \geq 0$, then the optimal terminal time is $T = 0$. Therefore, we consider in what follows only the case where $u(c) > 0$ for at least some $c > 0$. On the other hand, if $u(0) > 0$ then it is never optimal to terminate the utility flow at some finite time T. For this reason, we assume that $u(0) < 0$. It follows that there exists a threshold consumption level $\bar{c}$ such that $u(\bar{c}) = 0$ and $u(c) < 0$ if $c < \bar{c}$.

The Hamiltonian is given by (12.2), from which we obtain the conditions (12.3), (12.4), and (12.39). Moreover, the transversality conditions are

$$e^{-rT}\psi(T)x(T) = 0$$

and

$$e^{-rT}H(x(T), c(T), \psi(T)) = e^{-rT}[Nu(c(T)) - \psi(T)Nc(T)] = 0. \tag{12.40}$$

Since $u'(c(t)) = \psi(t)$ if $c(t) > 0$, condition (12.40) may be written as

$$e^{-rT}[u(c(T)) - c(T)u'(c(T))] = 0.$$

This condition is satisfied if T is infinite or if the equation

$$u(c(T)) - c(T)u'(c(T)) = 0 \tag{12.41}$$

holds. Given our assumptions on the utility function, in the present model of optimal extraction of a finite stock of an exhaustible resource, it is not optimal to have $T = \infty$. This can be seen as follows. If T were infinite then there would exist a finite time t_1 beyond which $c(t) < \bar{c}$ and, thus,

$$\int_{t_1}^{\infty} e^{-rt}u(c(t))\,dt < 0$$

would hold. This, however, would imply that it would be better to choose $T \leq t_1$. We conclude that T is finite, and it follows that (12.41) holds.

Equation (12.41) has a unique solution $c(T) = c_M > 0$ if $u(\cdot)$ is strictly concave and satisfies $u(0) < 0$, $u(c) > 0$ for some $c > 0$, and $\lim_{c\to\infty} u'(c) = 0$. In fact, at c_M the average utility $u(c_M)/c_M$ equals the marginal utility $u'(c_M)$. (The reader should sketch the graph of u and determine the point c_M from the graph.) In the case of a price-taking firm, the economic interpretation is that c_M is that level of output which equates the average total cost, $[G(c) + F]/c$, to the marginal cost, $G'(c)$. In other words, it is the level of output that minimizes the cost per unit extracted. In particular, if $r = 0$, and p and F are time-invariant, then this solution tells us that the firm should extract at a constant rate $c(t) = c_M$ for all $t \leq T$, and that at T the resource is exhausted. (Satisfy yourself that in this case $T = x_0/c_M$.) Upon reflection, the result that the output of the mining firm, under the assumptions just stated, should be at the point of minimum average cost, rather than at the point where marginal cost equals price, is not surprising. To see this one should note the distinction between a resource-extracting firm and a 'normal' firm (i.e., the type of static firm considered in any first-year textbook in microeconomics). For a resource-extracting firm, one more unit of extraction (i.e., output) today means one less unit of extraction at some time in the future. This is not the case with the normal firm. The reader is invited to prove that if $r > 0$ (and p and F are constant) then the extraction $c(t)$ falls over time, approaching c_M as t approaches T.

In what follows, for concreteness, we take $u(c) = \ln c$. Then (12.41) implies $c_M = e$ (where by definition $\ln e = 1$). Now the Hotelling rule gives $c(t) = c_M e^{r(T-t)}$. Resource exhaustion implies

$$x(0) = x(0) - x(T) = N \int_0^T c(t)\, dt = \frac{Nc_M}{r}(e^{rT} - 1).$$

Solving for T gives

$$T = \frac{1}{r} \ln\left(\frac{eN + rx_0}{eN}\right).$$

The value of the payoff integral is

$$V(x_0) = \frac{N}{r} \ln\left(\frac{eN + rx_0}{eN}\right),$$

which is positive for all positive x_0. It is easy to verify that this value function satisfies the HJB equation $rV(x) = \max\{N \ln c - V'(x)Nc \mid c > 0\}$ and that the optimal Markovian (or feedback) control rule is $c = e + rx/N$ if $x > 0$ and $c = 0$ if $x = 0$.

12.2.2 Stock-dependent utility

In many resource extraction problems, the utility function of each player depends not only on his extraction rate, but also on the remaining stock. For example, the stock may be a good proxy for the recreational value of the resource (think of a sand beach), or it may have a direct effect on a firm's profit (for example, the cost of harvesting fish may depend on the stock of fish).

The following example is a special case of a class of problems involving the long-term decline in effectiveness of a pesticide. Let $x(t)$ denote the effectiveness of a pesticide and $a_i(t)$ the rate of application of the pesticide by farmer i. Insects tend to develop resistance to the pesticide over time. To capture this phenomenon, we postulate that $\dot{x}(t) = -b\sum_{i=1}^{N} a_i(t)$ if $x(t) > 0$ and $\dot{x}(t) = 0$ when $x(t) = 0$, where b is a positive parameter. We distinguish the nominal dose $a_i(t)$ from the effective dose $a_i(t)x(t)$ and we assume that farmer i's profit rate at time t is $\pi_i(t) = [a_i(t)x(t)]^\alpha$, where $0 < \alpha < 1$. The cooperative problem is to maximize

$$\int_0^\infty e^{-rt} N[a(t)x(t)]^\alpha \, dt$$

subject to $a(t) \geq 0$ and

$$\dot{x}(t) = -bNa(t), \qquad x(0) = x_0, \qquad \lim_{t\to\infty} x(t) \geq 0.$$

The HJB equation for this problem is

$$rV(x) = \max\{N(ax)^\alpha - V'(x)bNa \mid a \geq 0\}.$$

Clearly, since the integrand is zero if $x = 0$, we require $V(0) = 0$. The unique solution of the HJB equation with this property is

$$V(x) = \frac{x^{2\alpha}[N(1-\alpha)/r]^{1-\alpha}}{(2b)^\alpha}$$

and the optimal rate of application of the pesticide is

$$Na(t) = \frac{rx(t)}{2b(1-\alpha)}.$$

Thus $\dot{x}(t)/x(t) = -r/(2-2\alpha)$.

When the farmers act noncooperatively, the Markovian strategies $a_i(t) = \phi(x(t))$ with $\phi(x) = rx[2b(1-N\alpha)]^{-1}$ constitute a symmetric Markov perfect Nash equilibrium provided that $N < 1/\alpha$. In this case the value of the payoff integral for each firm i is

$$V_i(x) = \frac{x^{2\alpha}[(1-N\alpha)/r]^{1-\alpha}}{(2b)^\alpha}. \tag{12.42}$$

The reader should check that the above Markovian strategy satisfies the HJB equation

$$rV_i(x) = \max\left\{(a_i x)^\alpha + V_i'(x)\left[-a_i - \frac{(N-1)rx}{2b(1-N\alpha)}\right] \,\middle|\, a_i \geq 0\right\},$$

where $V_i(\cdot)$ is given by (12.42).

It is clear that the rate of decline in effectiveness of the pesticide in this Markov perfect Nash equilibrium is greater than the rate obtained in the cooperative solution. This result corresponds to similar results obtained in most models of noncooperative exploitation of a stock.

12.3 Renewable resources

We now provide some examples of differential games involving a common property renewable resource, such as a fish stock or a forest.

12.3.1 A fishery model

Let $x(t)$ denote the stock of fish at time t, and $c_i(t)$ country i's rate of harvest of this renewable resource. The state variable $x(t)$ evolves according to the differential equation

$$\dot{x}(t) = g(x(t)) - \sum_{i=1}^{N} c_i(t),$$

where $g(\cdot)$ is the natural growth function which is assumed to satisfy the properties

$$g(0) = 0, \quad g'(0) > 0, \quad g''(x) < 0, \quad g'(\tilde{x}) = 0 \text{ for some } \tilde{x} > 0.$$

This implies that $g(x)$ reaches a maximum at $\tilde{x} > 0$ and that there exists a stock level $x_M > 0$ such that $g(x)$ is negative for all $x > x_M$. We call $\tilde{x}$ the maximum sustainable yield stock level. The payoff to country i is

$$J_i = \int_0^\infty e^{-rt} u(c_i(t))\, dt,$$

where $u(\cdot)$ is a strictly concave and increasing function and $r > 0$ is the rate of discount.[5] In what follows, we assume that $g'(0) > r$. This assumption means that the consumers are not too impatient and, as will be seen shortly, it ensures that the cooperative solution implies convergence of the stock to a strictly positive steady state level.

Let us consider the benchmark case in which N symmetric players cooperate to maximize the sum of their payoffs,

$$\int_0^\infty e^{-rt} N u(c(t))\, dt,$$

[5] A more general version would allow a utility function $u(c, x)$, i.e., the state variable may affect utility.

subject to $c(t) \geq 0$ and

$$\dot{x}(t) = g(x(t)) - Nc(t), \qquad x(0) = x_0, \qquad \lim_{t\to\infty} x(t) \geq 0.$$

It is easy to show that the cooperative solution implies that the stock converges to a steady state level x^* that is the unique solution of the equation

$$g'(x^*) = r. \tag{12.43}$$

This condition has the intuitive interpretation that the biological rate of growth (which is a sort of interest rate) at the steady state is equal to the rate of time preference.

Now consider the case where the countries play a noncooperative game. What are the equilibria? We assume $u(c) = B + (1-\eta)^{-1}c^{1-\eta}$, where B is a constant (possibly equal to 0) and $\eta > 1$ so that $u(0) = -\infty$. With this utility function, no country has an incentive to drive the stock to zero in finite time, because to do so would result in a payoff integral of minus infinity. It follows from this fact and from the fact that the stock does not appear in the integral that the cooperative solution can be supported by a symmetric open-loop Nash equilibrium. (The reader is invited to verify this claim.) However, we have pointed out that these equilibria are in general not subgame perfect. Let us therefore find a Markov perfect Nash equilibrium for this game. To obtain an explicit solution, we assume that $\eta = 2$ and $g(x) = x - x^2$. Without loss of generality we set $B = 0$. Let us assume that player i believes that all other players use a stationary Markovian strategy $c_j = \phi(x)$. We now use the Hamiltonian approach to obtain a candidate for a Markov perfect equilibrium, and afterwards we verify that it is indeed such an equilibrium by using the HJB equation. The Hamiltonian for player i is

$$H(x, c_i, \psi) = -1/c_i + \psi[x - x^2 - c_i - M\phi(x)],$$

where $M = N - 1$. This yields the conditions

$$c_i(t)^{-2} = \psi(t), \tag{12.44}$$

$$\dot{\psi}(t) = \psi(t)[r - 1 + 2x(t) + M\phi'(x(t))], \tag{12.45}$$

$$\dot{x}(t) = x(t) - x(t)^2 - c_i(t) - M\phi(x(t)). \tag{12.46}$$

From (12.44) and (12.45) follows

$$\dot{c}_i(t) = -\frac{c_i(t)}{2}[r - 1 + 2x(t) + M\phi'(x(t))]. \tag{12.47}$$

Divide (12.47) by (12.46) to obtain

$$2\frac{dc_i(x)}{dx} = -\frac{c_i(x)[r - 1 + 2x + M\phi'(x)]}{x - x^2 - c_i(x) - M\phi(x)}. \tag{12.48}$$

In a symmetric equilibrium we have $c_i(x) = \phi(x)$ for all $i \in \{1, 2, \ldots, N\}$. Use this in (12.48) to get

$$\phi'(x) = \frac{1 - r - 2x}{2(x - x^2)/\phi(x) - M - 2}.$$

To solve this differential equation, we impose the boundary condition that $\phi(0) = 0$ (no harvest if there is no fish). Then we obtain the solution

$$\phi(x) = \frac{(1 + r)x}{N + 1}.$$

Substitute this into (12.46) to obtain the steady state stock

$$x^{**} = \frac{1 - rN}{N + 1}$$

provided $rN < 1$. This is smaller than the steady state stock in the cooperative solution, $x^* = (1 - r)/2$, given by (12.43). This is to be expected. Everyone tries to capture more, because each one knows that if he captures less, this will tend to increase x in the future, which will lead others to capture more because the harvesting strategy has the property that $\phi'(x) > 0$.

The reader should verify that the strategies described above satisfy the HJB equations

$$rV_i(x) = \max\{-1/c_i + V_i'(x)[x - x^2 - c_i - M\phi(x)] \mid c_i > 0\}$$

where $V_i(x) = -a/x + b$, $a = [(N + 1)/(r + 1)]^2$, and $b = -a/r$.

12.3.2 A fishery model with capacity constraints

We now consider a modified fishery model by assuming that the utility function is linear (that is, $u(c) = c$) and that no country can harvest more than a maximum amount c_m. This constraint may, for example, reflect the limited quantity of vessels. The natural growth function $g(x)$ is assumed to have all the properties listed in the fishery model of the preceding subsection. In particular, it is assumed that there exists a value $x^* > 0$ such that $g'(x^*) = r$.

Assume that $N = 2$ and that $2c_m > g(\tilde{x})$ (here $\tilde{x}$ is the maximum sustainable yield), so that if both countries harvest at the maximum rate then the resource stock will be exhausted in finite time.

Let us characterize the cooperative solution first. We seek to maximize

$$2\int_0^\infty e^{-rt}c(t)\,dt$$

subject to

$$\dot{x}(t) = g(x(t)) - 2c(t), \qquad \lim_{t\to\infty} x(t) \geq 0.$$

The current value Hamiltonian is $H(x, c, \psi) = 2c + \psi[g(x) - 2c]$, from which we obtain the conditions

$$c(t)\begin{cases} = 0 & \text{if } \psi(t) > 1, \\ = c_m & \text{if } \psi(t) < 1, \\ \in [0, c_m] & \text{if } \psi(t) = 1, \end{cases}$$

and

$$\dot{\psi}(t) = \psi(t)[r - g'(x(t))].$$

Since H is linear in the control variable, the optimal control has the so-called bang-bang property: $c(t) = \phi(x(t))$ with $\phi(x) = 0$ if $x < x^*$, $\phi(x) = c_m$ if $x > x^*$, and $\phi(x^*) = g(x^*)/2$, where x^* is determined by (12.43).

We now turn to noncooperative solutions. To simplify the problem we make the following assumption:

Assumption 12.1 $g(x_0) \leq c_m$.

The interpretation of this assumption is as follows. If one player, say player 1, chooses to harvest at the maximum rate c_m, then no matter what player 2 does, the stock of fish will keep on declining to zero in finite time. In other words, player 2 cannot ensure a steady state with a positive stock when player 1 is greedy.

Under assumption 12.1, the pair of strategies (ϕ_1, ϕ_2) given by $\phi_i(x) = c_m$ for $x > 0$ and $\phi_i(0) = 0$ seems to be a natural candidate for a Markovian Nash equilibrium. Notice that, if player 1 plays this strategy, then player 2 faces a very simple control problem and, given assumption 12.1, it is easy to show that player 2's optimal exploitation is to set $c_2(t) = c_m$ until the resource is exhausted. Note that this Markovian equilibrium is not a Markov perfect Nash equilibrium (see section 4.3 for the distinction between these two concepts). Recall that according to the definition of a Markov perfect Nash equilibrium it is required that, given player 1's strategy, player 2's reply is optimal for all $t_0 \geq 0$ and all initial conditions $\xi = x(t_0)$, even for those stock levels ξ that are not reachable. Let us replace this requirement with the weaker requirement that, given player 1's strategy and given x_0 (which together imply that $x(t)$ must lie in some interval Z for all $t \geq 0$), player 2's reply is optimal for all $t_0 \geq 0$ and all reachable initial conditions at t_0, that is, all $x(t_0) = \xi \in Z$. In our particular example, owing to assumption 12.1 and player 1's strategy of maximum exploitation, the set of all reachable initial conditions is $Z = [0, x_0]$. Within this set, player 2's reply is, in fact, optimal. (Note that this modification gives a requirement that is slightly weaker than that given in definition 4.5, because our set Z is dependent on player 1's given strategy, while in definition 4.5, the attainable set A is independent of the strategies.)

12.3.3 Trigger strategy equilibria

We now show how the cooperative exploitation path of the fishery model in the preceding subsection (with linear utility and capacity constraints) can be achieved by trigger strategies. (See chapter 6 for a general exposition of trigger strategies.)

Let us choose the cooperative exploitation path as the target path. Any departure from the target path is called a defection. A player who defects will be punished by the other player. Clearly if a defection can be detected immediately and there is no delay in punishment then there cannot be any gain from defecting because it occurs only at an instant and therefore does not contribute to the payoff, which is an integral of discounted consumption. In this subsection we therefore postulate that there is some delay between defection time and detection time. We assume that at any time t the players know only the history of the game up to time $t - \theta$, where θ denotes the delay. However, we do assume that the players know the initial stock x_0 at time 0.

Let $x^{\#}(\cdot)$ and $c_i^{\#}(\cdot)$ denote respectively the target path of the stock and the corresponding control path for player $i \in \{1, 2\}$. A player detects at time t that the other player defected at time $t - \theta$ if $\dot{x}(t - \theta)$ differs from $\dot{x}^{\#}(t - \theta)$. Once a player defects it is in his interest to extract at the maximum rate, c_m, because he knows that, in the punishment phase, both players will extract at the maximum rate. Without loss of generality, assume player 1 is considering a defection that is to take place at time τ. Then the stock will be exhausted at time $\tau + T(x(\tau))$, where $T(x(\tau))$ satisfies

$$x(\tau) = -\int_{\tau}^{\tau+T(x(\tau))} \dot{x}(t)\, dt,$$

with $\dot{x}(t) = g(x(t)) - c_m - c_2(t)$, $c_2(t) = c_2^{\#}(t)$ for all $t \in [\tau, \tau + \theta]$, and $c_2(t) = c_m$ for $t > \tau + \theta$. The payoff to the defector is

$$D(x(\tau)) = \left[1 - e^{-rT(x(\tau))}\right]c_m.$$

Let $V_i^{\#}(x)$ denote the payoff to player i if both players follow the target path and the initial stock is x. The pair of trigger strategies constitutes a Nash equilibrium if $D(x^{\#}(t)) \leq V_i^{\#}(x^{\#}(t))$ for all t. The reader is left to show that if θ is sufficiently small and c_m is sufficiently large then this condition holds.

12.3.4 Stackelberg leadership

In all the equilibria considered so far, the players choose their strategies $\phi_i(\cdot)$, $i = 1, 2$, simultaneously. We now modify the game so that one player, say player 1, can announce his strategy, say $c_1 = h_1(x)$, and commit himself to that strategy before the other player makes his choice of strategy. This kind of hierarchic game is called a Stackelberg game and the player that has the right to move first is called the Stackelberg leader, the other player being designated as the follower. Clearly the Stackelberg leader can always ensure himself a payoff at least as great as that which he would obtain in a Markov perfect Nash equilibrium of the game with simultaneous plays, because he can always choose $h_1(\cdot) = \phi_1(\cdot)$, where ϕ_1 $(\cdot)$ denotes his strategy in the Markov perfect equilibrium.

In this subsection we present a simple model of exploitation of a common-property renewable resource and compare the Markov perfect Nash equilibrium of the simultaneous-play game with the equilibrium of the hierarchic game.

We suppose that there are two players and adopt the Schaefer specification that player i's harvest rate $c_i(t)$ depends on his effort level $E_i(t)$ and on the stock $x(t)$ by the equation

$$c_i(t) = qx(t)E_i(t),$$

where $q > 0$ is the catchability coefficient, which we will set at unity for simplicity. The stock of fish grows according to Gompertz' law

$$\dot{x}(t) = x(t)[1 - \ln x(t)] - x(t)E_1(t) - x(t)E_2(t). \tag{12.49}$$

The two players sell all the fish they catch, and the price is uniquely determined by the inverse demand function $P = f(c_1 + c_2)$. Assume that there is a cost of γ_i per unit of effort exerted by player i. The payoff for player i is

$$J_i = \int_0^\infty e^{-r_i t}[f(x(t)E_1(t) + x(t)E(t))x(t)E_i(t) - \gamma_i E_i(t)]\,dt.$$

Note that we allow the two players to have different discount rates. Let us assume that the inverse demand function $f(\cdot)$ is

$$f(xE_1 + xE_2) = \frac{1}{xE_1 + xE_2}.$$

This assumption implies that the integrand is independent of the state variable. It is clear, therefore, that if player j uses a constant effort level, then for player i the problem is essentially reduced to a static optimization problem and his best reply is to apply constant effort as well. To see how, in principle, equilibria with nonconstant effort levels can be identified, it is convenient to transform the state variable by defining $y = \ln x$. Then equation (12.49) becomes

$$\dot{y}(t) = 1 - y(t) - E_1(t) - E_2(t).$$

We first consider the case where both players must choose their strategies simultaneously. Consider player 1's optimization problem. If he expects that player 2's strategy is $E_2 = \phi_2(y)$, then his Hamiltonian function is

$$H_1(y, E_1, \psi_1) = \left[(E_1 + \phi_2(y))^{-1} - \gamma_1\right]E_1 + \psi_1[1 - y - E_1 - \phi_2(y)],$$

from which we obtain the conditions

$$[E_1(t) + \phi_2(y(t))]^{-1} - \gamma_1 - E_1(t)[E_1(t) + \phi_2(y(t))]^{-2} = \psi_1(t), \tag{12.50}$$

$$\dot{\psi}_1(t) = \psi_1(t)\left[1 + r_1 + \phi_2'(y(t))\right] + E_1(t)[E_1(t) + \phi_2(y(t))]^{-2}\phi_2'(y(t)). \tag{12.51}$$

Suppose that player 1's optimal reply can be represented by $E_1(t) = \phi_1(y(t))$. Substitute this into (12.50) and differentiate with respect to time to get

$$\dot{\psi}_1(t) = -\frac{1 - y(t) - S(y(t))}{S(y(t))^2}\left[2\phi_1'(y(t)) + \phi_2'(y(t)) - \frac{2\phi_1(y(t))S'(y(t))}{S(y(t))}\right], \tag{12.52}$$

where $S(y) = \phi_1(y) + \phi_2(y)$. On the other hand, equation (12.51) can be written as

$$\dot{\psi}_1(t) = A_1(y(t))\left[1 + r_1 + \phi_2'(y(t))\right] + \phi_1(y(t))\phi_2'(y(t))/S(y(t))^2, \tag{12.53}$$

where $A_1(y) = 1/S(y) - \gamma_1 - \phi_1(y)/S(y)^2$. Subtracting (12.53) from (12.52) we obtain a differential equation that involves $\phi_1(y)$ and $\phi_2(y)$ and their derivatives. A similar differential equation can be derived by considering player 2's optimization problem. Since the state variable y can take on negative values (recall $y = tf =$ "*times*"*ln* x) while $\phi_i(y)$ must be nonnegative ($E_i \geq 0$), the set of admissible solutions to this pair of differential equations is likely to be quite small. One obvious solution (already mentioned) is that $\phi_i(y)$ is independent of y. Writing $\phi_i(y) = K_i$, we can solve for K_1 and K_2 to get

$$K_i = \gamma_j/[\gamma_1 + \gamma_2]^2.$$

It is easy to check that the constant equilibrium strategies $\phi_i(y) = K_i$ satisfy the HJB equations

$$r_i V_i(y) = \max\{E_i/(E_i + K_j) - \gamma_i E_i \\ + V_i'(y)[1 - y - E_i - K_j] \mid E_i \geq 0\}, \quad j \neq i,$$

where $V_i'(y) = 0$ and $V_i(y) = [K_i/(K_i + K_j) - \gamma_i K_i]/r_i$. In fact, if player i believes that player j's strategy is a constant effort level, say $E_j = z$, then his best reply is

$$E_i = (z/\gamma_i)^{1/2} - z \tag{12.54}$$

provided that $z \leq 1/\gamma_i$.

Let us turn to the hierarchic-play game and assume that player 1 is the first mover in the sense that he can announce his strategy and make a commitment to follow it before player 2 chooses his strategy. Player 1 knows that if he chooses a constant strategy $E_1 = z$ then player 2 will choose the best reply $E_2(z)$ given by (12.54). This gives player 1 the payoff

$$J_1(z) = \frac{z}{r_1}\left[\frac{1}{E_2(z) + z} - \gamma_1\right].$$

Player 1, being the leader, then chooses a number $z \geq 0$ that maximizes his payoff.

12.4 A transboundary pollution game

In this section we use the model of transboundary pollution that we formulated in chapter 5 (see examples 5.1–5.3), except that we now treat the two countries symmetrically. Recall that, in that model, there are two countries, country 1 and country 2. Country i produces a single consumption good, Y_i, with a given fixed endowment of factors of production. There is an emission-consumption tradeoff function

$$Y_i = G_i(E_i),$$

which indicates the amount of pollutants emitted when country i produces Y_i units of the consumption good. Emissions add to the stock of pollution S according to the equation

$$\dot{S}(t) = E_1(t) + E_2(t) - kS(t),$$

where $k > 0$ is the coefficient of natural purification.

The representative consumer in country i derives utility $U_i(Y_i)$ from consumption, but incurs a discomfort $D_i(S)$. Assume for simplicity that $U_i(G_i(E_i)) = aE_i - (1/2)E_i^2$ and $D_i(S) = (b/2)S^2$ where b is a positive constant. Country i seeks to maximize the payoff

$$\int_0^{\infty} e^{-rt}[aE_i(t) - (1/2)E_i(t)^2 - (b/2)S(t)^2]\,dt.$$

It is not difficult to show that if the two countries collude, then the stock of pollution converges to a steady state level S_c given by

$$S_c = \frac{2a(r+k)}{k(r+k)+4b}$$

and, owing to symmetry, both countries achieve the same payoff

$$V_1(S_0) = V_2(S_0) = \frac{M_c}{2} - \frac{1}{2}B_cS_0 - \frac{1}{4}A_cS_0^2,$$

where

$$A_c = \frac{1}{2}\left[-\left(k+\frac{r}{2}\right) + \sqrt{\left(k+\frac{r}{2}\right)^2 + 4b}\right] > 0,$$

$$B_c = \frac{2aA_c}{r+k+2A_c} > 0,$$

$$M_c = \frac{(B_c - a)^2}{r} > 0.$$

We now turn to the scenario where the two countries do not cooperate and look for Markov perfect Nash equilibria. Since the model has a linear quadratic structure, we can use the methods developed in section 7.1 to show that the strategies

$$E_1(S) = E_2(S) = a - B_m - A_mS$$

qualify as a symmetric Markov perfect Nash equilibrium if the constants A_m and B_m are defined as

$$A_m = \frac{1}{3}\left[-\left(k+\frac{r}{2}\right) + \sqrt{\left(k+\frac{r}{2}\right)^2 + 3b}\right] > 0,$$

$$B_m = \frac{2aA_m}{r+k+3A_m} > 0.$$

At this equilibrium, each country's payoff is

$$V_i(S_0) = M_m - B_mS_0 - (1/2)A_mS_0^2,$$

where

$$M_m = \frac{\{a^2 - 4aB_m + 3B_m^2}{2r}.$$

It is not difficult to see that the steady state pollution stock in this equilibrium is greater than that obtained under collusion. Both countries receive a lower payoff compared with the collusive outcome.

12.5 Further reading

Hestenes' theorem, which we applied in section 12.1.2, is stated and discussed in, for example, Léonard and Long [160, pp. 248–251].

The following is a short list of papers and books that are most closely related to the models presented in this chapter. The list is not meant to be exhaustive.

Ulph [235] gives a useful review of books on resources and environmental economics. Various versions of the open-loop model of common-property nonrenewable resources are considered and discussed in Khalatbari [152], Kemp and Long [150], Sinn [219], Bolle [9, 10], and McMillan and Sinn [174]. Mohr [180] gives a defence of the concept of open-loop Nash equilibria in some contexts of natural resource games. Some theorems on the relationship between open-loop Nash equilibria and Markov perfect Nash equilibria in a class of resource games are proven in Long and Shimomura [164] and in Long et al. [165].

For fishery economics, a standard reference is Clark [27]. Kaitala [137] gives a useful survey of applications of dynamic game theory to fishery management. Clemhout and Wan [30, 31] present several deterministic and stochastic models of renewable resources in which players use stationary Markovian strategies. Jørgensen and Sorger [135] find Markovian equilibria of a fishing game under nonclassical assumptions on production and preferences. Chiarella et al. [19] prove the efficiency of open-loop Nash equilibria for a class of fishery models. For games of entry deterrence in a fishery, see Salchenberger [210] and Crabbé and Long [36]. Kaitala [138] examines the nonuniqueness issue of Markovian equilibria in a fishery game. Dockner and Sorger [61] and Sorger [222] show the existence of multiple Markov perfect Nash equilibria (with discontinuous Markovian strategies) in fishery models.

Trigger strategies in fishery models are considered by Hämäläinen et al. [113], Kaitala and Pohjola [141], and Benhabib and Radner [7]. Dockner et al. [50] consider Stackelberg leadership in fishery. For cooperative games of fishery, see Munro [182, 183] and Kaitala and Munro [140]. Games involving the use of pesticides and antibiotics are discussed in detail in Cornes et al. [34] and Long and Shimomura [164]. Dockner and Long [56] compare various transboundary pollution equilibria.

For discrete-time models of fish wars, see Levhari and Mirman [161], Cave [17], Sundaram [226], and Dutta and Sundaram [68]. A continuous-time version of the model of Levhari and Mirman [161] can be found in Plourde and Yeung [191]. For a discrete-time model of a pollution game, see Dockner et al. [58].

12.6 Exercises

1. Let $x(t)$ be the stock at time t of a common-property resource that is exploited by N identical players. Assume that $\dot{x}(t) = -\sum_{i=1}^{N} c_i(t)$. Find a symmetric Markov perfect Nash equilibrium for the case where the utility function is $u_i(c_i) = \ln c_i$. Show that the equilibrium strategies are independent of N.
2. Repeat exercise 1 for the case $u(c) = \ln[w(c)]$, where $w(c)$ is an increasing function which is homogeneous of degree $\alpha > 0$.
3. Consider a model of resource extraction with $\dot{x}(t) = -\sum_{i=1}^{N} q_i(t)$ where the players are N identical oligopolists. Replace the utility $u(q_i)$ by the profit $\pi_i = P(q_i + \sum_{j \neq i} q_j) q_i$, where $P(\cdot)$ is the inverse demand function. For simplicity, take $P(Q) = Q^{-1/\eta}$ where $\eta > 0$ is the constant elasticity of demand. Find the cooperative solution and show that it is supported by an open-loop Nash equilibrium if the firms are restricted to use strictly feasible open-loop replies. What restrictions on N and η are necessary for the existence of a symmetric open-loop Nash equilibrium if the players are allowed to use weakly feasible open loop replies?
4. Show that the cooperative solution of the doomsday problem with $u(c) = \ln c$ cannot be supported as an open-loop Nash equilibrium if the players are allowed to use weakly feasible open-loop replies.
5. Find a symmetric open-loop Nash equilibrium for the pesticide model from section 12.2.2 when $N < 1/\alpha$ under the assumption that farmers can choose weakly feasible open-loop replies. Show that in this equilibrium the rate of decline of the effectiveness of the pesticide is greater than in the cooperative solution.
6. (A Gompertz fishing game.) Consider a fishery model with two players having the utility function $u(c_i) = \ln c_i$. Assume that $\dot{x}(t) = x(t)[1 - \ln x(t)] - c_1(t) - c_2(t)$. (The growth function $g(x) = x(1 - \ln x)$ is called the Gompertz growth function.) Define $\alpha_i(t) = c_i(t)/x(t)$ and $y(t) = \ln x(t)$. One may interpret $\alpha_i(t)$ as the fishing effort of player i at time t. Then the payoff for player i is

$$J_i = \int_0^\infty e^{-rt}[y(t) + \ln \alpha_i(t)]\, dt,$$

which is to be maximized by choosing the time path $\alpha_i(\cdot)$ subject to $\dot{y}(t) = 1 - y(t) - \alpha_1(t) - \alpha_2(t)$. Show that the pair of strategies $\alpha_i(t) = 1 + r$, $i = 1, 2$, constitutes a subgame perfect open-loop Nash equilibrium. Show that if both players cooperate then the payoff for each of them is

$$V_i^c(y_0) = \frac{y_0}{1+r} + \frac{1}{r}\left[\ln\left(\frac{1+r}{2}\right) - \frac{r}{1+r}\right]$$

and that, in the equilibrium given above, the corresponding payoff is lower.

7. In the model of exercise 6 examine the possibility of sustaining the cooperative solution by using trigger strategies.
8. Show that in the following fishing game, any symmetric open-loop Nash equilibrium leads to a steady state fish stock that is lower than the steady state fish stock in the corresponding cooperative game. The stock of fish grows according to the law $\dot{x}(t) = x(t)[1 - x(t)] - 2\sum_{i=1}^{2}[x(t)E_i(t)]^{1/2}$, where $E_i(t)$ denotes player i's effort level at time t. Player i takes the time path $E_j(\cdot)$, $j \neq i$, as given and seeks to maximize

$$\int_0^\infty e^{-rt}[PY_i(t) - wE_i(t)]\, dt,$$

where $r \in (0, 1)$ is the discount rate, $P > 0$ is the price of fish, $w > 0$ is the unit cost of effort, and $Y_i(t) = 2[x(t)E_i(t)]^{1/2}$ is player i's catch rate at time t. You may set $P = w = 1$ and assume $0 < r < 1/2$ for simplicity.

Appendix

Lemma 12.1 If $1 - \eta > 1/N$ then $(1 - \eta)^{1-\eta}\eta^\eta > (N - 1)^\eta N^{-1}$.

Proof Because $1 - \eta > 1/N$ there exists a real number $y \in (1, N)$ such that $1 - \eta = 1/y$. It is easy to verify that

$$(1 - \eta)^{1-\eta}\eta^\eta = \left(\frac{1}{y}\right)^{1/y}\left(\frac{y-1}{y}\right)\left(\frac{y-1}{y}\right)^{-1/y} = \frac{(y-1)^\eta}{y}.$$

Next, notice that from $N > y$ and $1 - \eta = 1/y$ it follows that

$$\frac{(y-1)^\eta}{y} > \frac{(N-1)^\eta}{N}$$

because the function $x \mapsto (x-1)^\eta x^{-1}$ is decreasing for all $x > 1/(1-\eta)$. ■

In fact, what we have shown is that the function $x \mapsto (1-\eta)^{1-\eta}\eta^\eta - (x-1)^\eta x^{-1}$, defined for $x \geq 1$ and $\eta \in (0,1)$, is non-negative and attains its minimum at $x = 1/(1-\eta)$.

Answers and hints for exercises

Chapter 2

1. Using the approach outlined in example 2.1 it is easy to construct the strategic form. This form exhibits two Nash equilibria but (A, b) is the only subgame perfect equilibrium. Backward induction yields the same result. See Selten [213].
2. Profit functions are given by

$$\pi^1(Q_1, Q_2) = [100 - 4(Q_1 + Q_2) + 3(Q_1 + Q_2)^2 - (Q_1 + Q_2)^3]Q_1 - 4Q_1,$$
$$\pi^2(Q_1, Q_2) = [100 - 4(Q_1 + Q_2) + 3(Q_1 + Q_2)^2 - (Q_1 + Q_2)^3]Q_2 - 2Q_2 - 0.1Q_2^2,$$

and in an equilibrium with positive outputs we get

$$\frac{\partial \pi^1(Q_1, Q_2)}{\partial Q_1} = 96 - 4(Q_1 + Q_2) + 3(Q_1 + Q_2)^2 + 6Q_1(Q_1 + Q_2) - (Q_1 + Q_2)^3 - 3Q_1(Q_1 + Q_2)^2 - 4Q_1 = 0,$$
$$\frac{\partial \pi^2(Q_1, Q_2)}{\partial Q_2} = 98 - 4Q_1 - 8.2Q_2 + 6Q_2(Q_1 + Q_2) + 3(1 - Q_2)(Q_1 + Q_2)^2 - (Q_1 + Q_2)^3 = 0,$$

which yields the Cournot–Nash equilibrium $Q_1 = 2.028$, $Q_2 = 2.081$. See Friedman [100].
3. Let $i \in \{1, 2\}$ and denote the first-period actions of the two players by $a_i \in \mathbb{R}$, $i \in \{1, 2\}$, and the second-period actions by $b_i \in \mathbb{R}$, $i \in \{1, 2\}$. Denote the payoffs by $J^i(a, b)$, where $a = (a_1, a_2)$ and $b = (b_1, b_2)$. An open-loop Nash equilibrium is a quadruple $(a^{\mathrm{OL}}, b^{\mathrm{OL}})$ which satisfies for all $i, j \in \{1, 2\}$ with $i \neq j$

$$a_i^{\mathrm{OL}} = \operatorname{argmax}\left\{J^i((a_i, a_j^{\mathrm{OL}}), b^{\mathrm{OL}}) \mid a_i \in A_i\right\},$$
$$b_i^{\mathrm{OL}} = \operatorname{argmax}\left\{J^i(a^{\mathrm{OL}}, (b_i, b_j^{\mathrm{OL}})) \mid b_i \in B_i\right\}.$$

Since action sets A_i and B_i are open, an open-loop Nash equilibrium satisfies the first order condition

$$\frac{\partial J^i(a, b)}{\partial a_i} = \frac{\partial J^i(a, b)}{\partial b_i} = 0, \quad i = 1, 2. \tag{0.1}$$

Next, suppose that the players let their second-period actions depend on the first period ones according to functions $b_i^{AH}(\cdot)$, $i = 1, 2$. The argument of these functions is a and the superscript AH signifies that now second-period actions are contingent on first-period actions. Requiring subgame perfectness means that the second-period actions must be a Nash equilibrium in the period 2 game for any choice of period 1 actions (that is, for any history). Formally we write this requirement as

$$b_i^{AH}(a) = \operatorname{argmax}\left\{J^i(a, (b_i, b_j^{AH}(a))) \mid b_i \in B_i\right\}, \quad i = 1, 2, \quad i \neq j. \tag{0.2}$$

As to the period-1 actions we note – in contrast to the open-loop case – that when determining these actions the players need to take into account that the period-2 actions will depend on the period-1 actions according to the strategies in (0.2). The first-order condition for an optimal a_i then becomes

$$\frac{\partial J^i(a, b^{AH}(a))}{\partial a_i} + \frac{\partial J^i(a, b^{AH}(a))}{\partial b_j}\frac{\partial b_j^{AH}(a)}{\partial a_i} = 0, \quad i = 1, 2, \quad i \neq j, \tag{0.3}$$

provided the functions defined in (0.2) are differentiable. Comparing (0.1) and (0.3) shows that, in the action-history dependent equilibrium, the first-order condition is extended with an additional term which accounts for the influence of a_i on b_j. After having studied section 4.2, the reader will note the similarity of (0.1) and (0.3) and the maximum principle equilibrium conditions in a differential game. See Fudenberg and Tirole [105].

4. Table 2.6 shows that the entrant's best reply to 'collude' is 'enter' and his best reply to 'fight' is 'stay out'. The incumbent's best reply to 'enter' is 'collude' and his best replies to 'stay out' are 'collude' and 'fight'. The conclusion follows.
5. Backward induction tells us to start looking at the game at the time when the follower makes his decision. Whatever action the leader might take, the follower maximizes his payoff by choosing his action according to the best reply (2.5). Working backward, it only remains to determine the leader's optimal choice. He chooses his Stackelberg output (equal to 1/2) since this gives him more profit than any other feasible output, including his Cournot output (equal to 1/3).
6. The game has $N + 1$ players. There is one chain store, having branches in N cities, and one potential entrant in each of the N cities. There are many Nash equilibria: every terminal history having in all stages the outcome of either 'stay out' or 'enter, collude' is the outcome of a Nash equilibrium. Thus, the identical repetition of the Nash equilibrium in the one-stage game is a Nash equilibrium in the N-stage game. (Recall that ('stay out', 'fight') and ('enter', 'collude') were the Nash equilibria in the one-entrant game, see example 2.6.) There is, however, only one subgame perfect Nash equilibrium. In this equilibrium each entrant chooses 'enter' and the chain store always plays 'collude'. (Recall that ('enter', 'collude') was the subgame perfect Nash equilibrium in the one-entrant game of example 2.6.) The argument is simple and follows by

backward induction: in the last city, city N, it is clear that entrant N should enter and the chain store must choose 'collude'. This is true no matter how the history has been in the preceding $N-1$ cities. In city $N-1$, the chain store must do the same since it cannot influence city N through its decision in city $N-1$. Working backward the result is that in a subgame perfect outcome the chain store colludes with all entrants.

The result is known as the 'chain store paradox' and was studied in Selten [215]. The source of the paradox is that the subgame perfect outcome seems counterintuitive, from the point of view that the chain store would like to build a reputation of being tough, by playing 'fight' against a certain number of the first entrants. The last few entrants will probably not be deterred from entering the market although they know that the chain store has a reputation of being tough. With only a few cities left, it may not be worthwhile for the entrant to play 'fight': rather it could collude to obtain some last-minute profits. This intuitively plausible behaviour is ruled out in the original setting of the game but can be rationalized if one modifies the game. The modification turns the game into one of incomplete information, assuming that the chain store's type ('strong' or 'weak') is unknown to the entrants. See also Friedman [99] and Osborne and Rubinstein [188].

Chapter 3

1. The adjoint equation is $\dot{\lambda}(t) = \lambda(t)^2 M^2[1 - x(t)]$ and the unique control value which maximizes the Hamiltonian is equal to $u(t) = \lambda(t)M[1 - x(t)]$. Differentiating the latter equation with respect to t and using the adjoint equation and the state dynamics yields $\dot{u}(t) = 0$. Substituting $u(t) = \bar{u}$ into the state dynamics and into the objective functional gives $J(u(\cdot)) = 1 - e^{-\bar{u}MT} - \bar{u}^2 T/2$. Differentiation with respect to $\bar{u}$ yields the result.
2. One can verify the conditions of theorem 3.1 or theorem 3.2. The optimal value function is $V(x, t) = A(t)x$, where $A(\cdot)$ is the unique solution of the Riccati differential equation $\dot{A}(t) = rA(t) - [p - A(t)]^2/(2\gamma)$ with the boundary condition $A(T) = q$. The adjoint variable is $\lambda(t) = A(t)$. An optimal Markovian strategy is $u(t) = \phi(x(t), t)$ with $\phi(x, t) = x[p - A(t)]/\gamma$. If $q > p$ then it is optimal not to extract anything during the final interval $[\tau, T]$, where $\tau = \max\{0, T - (1/r)\ln(q/p)\}$. This can be shown as before with the following differential equation for $A(\cdot)$:
$$\dot{A}(t) = \begin{cases} rA(t) - [p - A(t)]^2/(2\gamma) & \text{if } A(t) \le p, \\ rA(t) & \text{if } A(t) \ge p. \end{cases}$$
3. $B(t) = \left[1 - (1+r)e^{-r(T-t)}\right]/r$, $A(t) = (1/4)\int_t^T e^{-r(s-t)}(1+s)B(s)^2\,ds$, $u(t) = (1+t)B(t)/2$.
4. To prove the first statement you can verify the conditions of theorem 3.1 with a value function of the form $V(x_1, x_2, t) = A(t) + B(t)x_1 + Cx_1^2 + D(t)x_2$. This yields $B(t) = [7e^{3(t-T)} + 3e^{t-T}]/2 - 1$, $C = -1$, $D(t) = 1 - e^{t-T}$, and $A(t) = (1/2)\int_t^T e^{-(s-t)}[B(s) + D(s)]^2\,ds$. The open-loop representation is

$$u(t) = \frac{1}{15}e^{-2t}\left[-30x_{10} + 5e^{-T} + 21e^{-3T}\right] + \frac{1}{6}e^{t-T} + \frac{21}{10}e^{3(t-T)}.$$

Both ψ_1 and ψ_2 generate this path.

5. Verify the conditions of theorem 3.3 and theorem 3.2. To this end, first compute the state trajectory corresponding to $u(\cdot)$. Then show that the adjoint variable defined by

$$\lambda(t) = \begin{cases} e^{(r+1)(t-\tau)} & \text{if } t \in [0, \tau), \\ \left[(r+1)e^{(r+2)(t-\tau)} + 1\right]/(r+2) & \text{if } t \in [\tau, \sigma), \\ 2/(r+1) & \text{if } t \in [\sigma, \infty) \end{cases}$$

satisfies the adjoint equation. Finally, verify the maximum condition. Note that every feasible state trajectory remains bounded and that $\lambda(t)$ is also bounded. Because $r > 0$ the transversality condition follows from lemma 3.1(iii).

6. Verify the conditions of theorem 3.6. The adjoint variable is given by $\lambda(t) = u(t)$.

7. $x(t) = 1/(4\alpha^2) + [x_0 - 1/(4\alpha^2)]e^t$. If $\alpha > 1/(2\sqrt{x_0})$ and $\tilde{x}(\cdot)$ is any bounded state trajectory then the transversality condition (3.31) does not hold with the value function $V(\cdot;\alpha)$ because $\liminf_{t\to\infty} e^{-t}[V(\tilde{x}(t), t; \alpha) - V(x(t), t; \alpha)] = -\alpha[x_0 - 1/(4\alpha^2)] < 0$. If $\alpha = 1/(2\sqrt{x_0})$ then $x(\cdot)$ is bounded and lemma 3.1(i) can be applied to verify the transversality condition. The same argument can also be applied with $V(x, t) = \sqrt{x}$.

Chapter 4

1.

$$u^1(t) = \begin{cases} 1 & \text{if } t \in [0, T-1), \\ 0 & \text{if } t \in [T-1, T-1/2), \\ 1 & \text{if } t \in [T-1/2, T], \end{cases} \qquad u^2(t) = \begin{cases} 1 & \text{if } t \in [0, T-1/2), \\ 0 & \text{if } t \in [T-1/2, T]. \end{cases}$$

This can most easily be shown by applying theorem 4.2 with the adjoint variables defined by $\lambda^1(t) = T - t$ and

$$\lambda^2(t) = \begin{cases} 2(T-t) - 1/2 & \text{if } t \in [0, T-1), \\ T - t + 1/2 & \text{if } t \in [T-1, T-1/2), \\ 2(T-t) & \text{if } t \in [T-1/2, T]. \end{cases}$$

2. To prove the first part apply theorem 4.1 with $V^i(x, t) = -A(t)x$. To prove the second part, apply theorem 4.2 with $\lambda^i(t) = -B(t)$.

3. In a symmetric Markov perfect Nash equilibrium each player uses the strategy $u^i(t) = \phi(x(t), t)$ with $\phi(x, t) = x[p - A(t)]/\gamma$, where the function $A(\cdot)$ is the unique solution of the differential equation $\dot{A}(t) = rA(t) - [p - A(t)][p - (2N-1)A(t)]/(2\gamma)$ and $A(T) = q$. This can be shown by verifying the conditions of theorem 4.4 with the optimal value functions $V^i(x, t) = A(t)x$ for all $i \in \{1, 2, \ldots, N\}$.

4. Set $V^i = V$ and $\phi^i = \phi$ for all $i \in \{1, 2, \ldots, N\}$. Then equation (4.11) is

$$rV(x, t) - V_t(x, t) = V_x(x, t)K(x) + (2 - N)/[4V_x(x, t)]$$

and equation (4.14) is

$$\frac{r-(1-N)\phi_x(x,t)-K'(x)}{2\sqrt{\phi(x,t)}}+\frac{\phi_t(x,t)+\phi_x(x,t)[K(x)-N\phi(x,t)]}{4\phi(x,t)^{3/2}}=0.$$

The latter equation can be simplified as follows

$$2\phi(x,t)[r-K'(x)]+\phi_t(x,t)+\phi_x(x,t)[K(x)-(2-N)\phi(x,t)]=0.$$

5. Any pair (ϕ^1, ϕ^2) where $\phi^1 : [0,\infty) \mapsto \{0, \alpha/r, 1\}$ is a piecewise constant function and $\phi^2(t) = \phi^1(t)$ holds for all t constitutes a subgame perfect Markovian Nash equilibrium. This can be proved by applying theorem 4.2 with $\lambda^i(t) = 1/r$.
6. Apply theorem 4.2 with $\lambda^i(t) = (1 + e^{-x_0})/r$. To prove that the equilibrium is not subgame perfect, consider the subgame $\Gamma(0, T)$ where T is sufficiently large. Show that in this subgame each player is better off to choose $u^i(t) = 0$ for all $t \geq T$ than to choose the equilibrium path $u^i(t) = 1$.

Chapter 5

1. First, consider the perfect competition scenario. The inverse demand function is $p = C_i^{-b}$, where $0 < b < 1$. There is no 'choke price' and, as p tends to infinity, C_i tends to zero. Since $p(t) = p_0^* e^{rt}$ and $C_H(t) = C_F(t) = [p(t)]^{-1/b}$, the market clearing condition is

$$S_0 = 2\int_0^\infty C_F(t)\,dt = \frac{2b[p_0^*]^{-1/b}}{r}.$$

Thus, given S_0, we can compute the competitive equilibrium price $p_0^* = [rS_0/(2b)]^{-b}$. Now assume the home country's consumers form a coalition to restrict oil imports. The coalition chooses the total cumulative demand Z_0. Then $S_0 - Z_0$ is supplied to consumers in country F. The market-clearing condition is

$$S_0 - Z_0 = \int_0^\infty C_F(t)\,dt = \frac{b[p_0^{**}]^{-1/b}}{r}.$$

Solving for p_0^{**} we get

$$p_0^{**} = f(Z_0) := \left(\frac{rS_0 - rZ_0}{b}\right)^{-b}.$$

Because $v'(0) = \infty$, the coalition will plan to have positive consumption at every instant t. The integral of discounted utility for H is

$$\int_0^\infty e^{-rt}[v(C_H(t)) + y - p(t)C_H(t)]\,dt = \int_0^\infty e^{-rt}[v(C_H(t)) + y]\,dt - f(Z_0)Z_0.$$

With $Z(t)$ as the state variable and $\pi(t)$ as the costate variable, we obtain the transversality condition at the initial time

$$\pi(0) = f(Z_0) + Z_0 f'(Z_0).$$

But $\pi(0) = C_H(0)^{-b} = [(r/b)Z_0]^{-b}$. Using this to substitute for $\pi(0)$ in the above transversality condition, we obtain

$$\left(\frac{rZ_0}{b}\right)^{-b} - \left(\frac{rS_0 - rZ_0}{b}\right)^{-b}\left(1 + \frac{brZ_0}{rS_0 - rZ_0}\right) = 0.$$

This equation determines the optimal Z_0. Let $G(Z_0)$ denote the left-hand side of the above equation. Since $G(0) = \infty$, $G(S_0/2) < 0$, and $G'(Z) < 0$ for all Z, there exists a unique Z_0 that satisfies the above equation. This value satisfies $0 < Z_0 < S_0/2$.

2. (For more details, see Benchekroun et al. [6].)
(a) Let $r > 0$ be the rate of discount. The costate variable for firm i is λ_i. In an open-loop Nash equilibrium $\lambda_2(t) = 0$ identically and the Cournot equilibrium outputs at t, given $\lambda_1(t)$ and $K(t)$, are

$$Q_1(t) = \frac{1}{3b}[a - 2c(K(t)) + 2\lambda_1(t)] \quad \text{and} \quad Q_2(t) = \frac{1}{3b}[a + c(K(t)) - \lambda_1(t)].$$

Thus we obtain the pair of equations

$$\dot{\lambda}_1(t) = (r + \delta)\lambda_1(t) + \left[\frac{a - 2c(K(t)) + 2\lambda_1(t)}{3b}\right]c'(K(t))$$

and

$$\dot{K}(t) = \frac{a - 2c(K(t)) + 2\lambda_1(t)}{3b} - \delta K(t).$$

Assumimg $a > 2\bar{c}$ and

$$3\delta b > \gamma\left[\frac{a}{\bar{c}} + \frac{2\delta}{r + \delta}\right]$$

the system has the positive steady-state solution

$$K_\infty = \frac{(r + \delta)(a - 2\bar{c})}{3b\delta(r + \delta) - 2\gamma(r + 2\delta)} \quad \text{and} \quad \lambda_{1\infty} = \frac{\delta\gamma K_\infty}{r + \delta}.$$

The Jacobian of the system has two real roots of opposite sign. Take the negative root to ensure convergence.
(b) The pair of strategies

$$Q_1(K) = X^* + Y^*K, \qquad Q_2(K) = e^* + f^*K$$

constitutes a Markov perfect Nash equilibrium. Here, under the assumptions made in part (a) above regarding restrictions on parameter values,

$$f^* = \frac{-(r+2\delta)+\Delta_1}{4} < 0,$$
$$e^* = \frac{4af^* + (r+\delta)(a+\bar{c})}{8bf^* + 3b(r+\delta)},$$
$$Y^* = \frac{r+2\delta-\Delta_1}{2} > 0,$$
$$X^* = \frac{2a-\bar{c}}{b[3+8f/(r+\delta)]},$$

and

$$\Delta_1 = \sqrt{(r+2\delta)^2 - (r+2\delta)(8\gamma/2b)} > 0.$$

(c) To derive a Stackelberg equilibrium we restrict attention to linear affine strategies. Firm 2 announces its strategy $Q_2(t) = e + fK(t)$. Firm 1, knowing e and f, chooses its best reply $Q_1(t) = X(e,f) + Y(e,f)K(t)$. Using the HJB equation for firm 1, one can show that

$$Y(e,f) = \frac{1}{2b}[b(r+2\delta) - \Delta]$$

and

$$X(e,f) = \frac{1}{2b}\left[\frac{(a-be-\bar{c})(r+\delta)}{r+\delta-Y(e,f)}\right],$$

where[1]

$$\Delta = \sqrt{[b(r+2\delta)]^2 - 2b(r+2\delta)(\gamma - bf)}.$$

Knowing $Y(e,f)$ and $X(e,f)$, firm 2 then chooses e and f to maximize its payoff

$$\int_0^\infty e^{-rt}\{a - b[e + fK(t) + X(e,f) + Y(e,f)K(t)]\}[e + fK(t)]\,dt$$

subject to

$$\dot{K}(t) = X(e,f) + Y(e,f)K(t) - \delta K(t)$$

and $K(0) = K_0$.

For given (e,f), we can solve for $K(t)$ from the differential equation, and the result is substituted into the objective function. The optimal values of e and f can finally be found.

3. The follower's Hamiltonian is

$$H^F(k,c,\lambda,t) = \frac{c^{1-\sigma}-1}{1-\sigma} + \ln g(t) + \lambda\{[1-\theta(t)]Ak^\sigma - c\},$$

from which we get

[1] We assume $2(\gamma - bf) < b(r+2\delta)$ so that Δ is real.

$$c(t)^{-\sigma} = \lambda(t),$$
$$\dot{k}(t) = [1 - \theta(t)]Ak(t)^{\sigma} - c(t),$$
$$\dot{\lambda}(t) = \rho\lambda(t) - \sigma\lambda(t)Ak(t)^{\sigma-1}[1 - \theta(t)].$$

These three equations yield a general solution with the property that $\lambda(t)^{1/\sigma}k(t) = (\sigma/\rho) + Qe^{\rho t/\sigma}$, where Q is a constant. To prove this, write $h(t) = \lambda(t)^{1/\sigma}k(t)$, so that $\dot{h}(t)/h(t) = (\rho/\sigma) - [1/h(t)]$. When $Q = 0$, then $c(t) = \rho k(t)/\sigma$ and $-\rho k(t)/\sigma \leq \dot{k}(t) \leq Ak(t)^{\sigma} - \rho k(t)/\sigma$, and this implies that $k(t)$ is positive and bounded from above, because $\sigma < 1$ and $k(0) > 0$. Also, when $Q = 0$,

$$\lim_{t\to\infty} k(t)\lambda(t)e^{-\rho t} = \lim_{t\to\infty} k(t)^{1-\sigma}h(t)^{\sigma}e^{-\rho t} = \lim_{t\to\infty} k(t)^{1-\sigma}(\sigma/\rho)^{\sigma}e^{-\rho t} = 0.$$

So, when $Q = 0$, the follower's transversality condition is satisfied, and the follower's consumption rule is $c(t) = \rho k(t)/\sigma$, independent of the future tax rate. The leader's problem can be solved as a simple optimal control problem.

4. Use the HJB equations.
5. Check that the HJB equations are satisfied.
6. Solve the differential equation for $E(t)$ and substitute it into the integrand of the objective function, then integrate, and differentiate the resulting expression with respect to b to obtain the first order condition for the optimal b. Check the second order condition.
7. The Hamiltonian for the follower is

$$H^F(A, C, L, \lambda, t) = \ln(C - L) + \lambda\{R(t)[1 - \theta_K(t)]A + W(t)[1 - \theta_L(t)]L - C\},$$

from which we get, for an interior solution,

$$\frac{1}{C(t) - L(t)} = \lambda(t) = \lambda(t)W(t)[1 - \theta_L(t)]$$

and

$$\dot{\lambda}(t) = \rho\lambda(t) - R(t)\lambda(t)[1 - \theta_K(t)].$$

Using these conditions, we obtain

$$\dot{A}(t) = R(t)A(t)[1 - \theta_K(t)] - \frac{1}{\lambda(t)}.$$

Therefore

$$\frac{1}{A(t)\lambda(t)}\frac{d[A(t)\lambda(t)]}{dt} = \rho - \frac{1}{A(t)\lambda(t)}.$$

This equation and the transversality condition $\lim_{t\to\infty} A(t)\lambda(t)e^{-\rho t} = 0$ imply $A(t)\lambda(t) = 1/\rho$. It follows that $C(t) - L(t) = \rho A(t)$.
Since $W(t)[1 - \theta_L(t)] = 1$ and $W(t)$ is equal to the marginal product of labour at time t, we have $L(t)/K(t) = \{(1 - \beta)[1 - \theta_L(t)]\}^{1/\beta}$. Therefore,

$$R(t)K(t) = \beta\{(1 - \beta)[1 - \theta_L(t)]\}^{(1-\beta)/\beta}K(t),$$
$$W(t)L(t) = (1 - \beta)\{(1 - \beta)[1 - \theta_L(t)]\}^{(1-\beta)/\beta}K(t).$$

The government's problem is to maximize

$$\int_0^\infty e^{-\rho t}[\ln(\rho A(t)) + \ln G]\, dt$$

subject to

$$\begin{aligned}
\dot{A}(t) &= [1-\theta_K(t)]\beta\{(1-\beta)[1-\theta_L(t)]\}^{(1-\beta)/\beta}A(t) - \rho A(t),\\
\dot{B}(t) &= G - \theta_K(t)\beta\{(1-\beta)[1-\theta_L(t)]\}^{(1-\beta)/\beta}A(t)\\
&\quad - \theta_L(t)(1-\beta)\{(1-\beta)[1-\theta_L(t)]\}^{(1-\beta)/\beta}[A(t)-B(t)]\\
&\quad + \beta\{(1-\beta)[1-\theta_L(t)]\}^{(1-\beta)/\beta}B(t).
\end{aligned}$$

Note that G is assumed to be exogenous and constant.
The optimal paths $\theta_L(\cdot)$ and $\theta_K(\cdot)$ are

$$\theta_L(t) = 0 \text{ for all } t,$$

$$\theta_K(t) = \begin{cases} 1 & \text{if } t < T,\\ 0 & \text{if } t \geq T,\end{cases}$$

where T is given by

$$\frac{\gamma A_0}{\rho+\gamma}\left[1 - e^{-(\rho+\gamma)T}\right] = B_0 + \frac{G}{\gamma}$$

and $\gamma = \beta(1-\beta)^{(1-\beta)/\beta}$. See Xie [245] for details.

Chapter 6

1. For $f(x) = x^\alpha$ it holds that

$$\bar{x} = \left(\frac{r+\gamma}{\alpha}\right)^{1/(\alpha-1)} \quad\text{and}\quad \tau(x_0) = \frac{1}{\gamma(1-\alpha)}\ln\left(\frac{1-\gamma x_0^{1-\alpha}}{1-\gamma\bar{x}^{1-\alpha}}\right).$$

   To verify that $\tilde{x}(\cdot)$ as specified in the exercise is indeed the Pareto optimal state trajectory, substitute into $\dot{\tilde{x}}(t) = \tilde{x}(t)^\alpha - \gamma\tilde{x}(t)$. Condition (6.15) reduces in this case to $\alpha(1-\alpha)\gamma^2 > 0$, which is obviously true.
2. Part (i) follows in the same way as in the proof of theorem 6.3(i). Note that (6.16) holds. The proof of (ii) follows from (6.16) in an analogous way to that in which theorem 6.3(ii) follows from (6.8).
3. (i) Nonanticipating and not regular, (ii) nonanticipating and regular, (iii) anticipating and regular.
4. (i) To find the Pareto optimal solution, note that $\tilde{J} = J^1 + J^2$ can be written as

$$\begin{aligned}
\tilde{J} &= \int_0^\infty e^{-rt}\left[x(t) - x(t)^2/2 + u^1(t) - [u^1(t)]^2/2 + u^2(t)\right] dt\\
&= \int_0^\infty e^{-rt}\left[x(t) - x(t)^2/2 + u^1(t) - [u^1(t)]^2/2 + \dot{x}(t)\right] dt\\
&= \int_0^\infty e^{-rt}\left[x(t) - x(t)^2/2 + u^1(t) - [u^1(t)]^2/2 + rx(t)\right] dt - x_0.
\end{aligned}$$

Since $u^1(t)$ does not appear in the system dynamics it is optimal to choose $u^1(t)$ such that the integrand is maximized at each point in time. This yields $u^1(t) = \tilde{\phi}^1(t) = 1$. Furthermore, since the integrand is a strictly concave function of $x(t)$ which is maximized at the unique value $x(t) = 1 + r$ it is optimal to choose $u^2(\cdot)$ such that the state trajectory stays as close as possible to the constant level $1 + r$. Because of the control constraint for u^2 this yields $u^2(t) = \tilde{\phi}^2(x(t))$ with

$$\tilde{\phi}^2(x) = \begin{cases} 0 & \text{if } x \leq 1 + r, \\ -1 & \text{if } x > 1 + r. \end{cases}$$

(ii) Since player 1 cannot influence the state equation he must choose his control $u^1(t)$ such that the utility function F^1 is maximized at any $t \in [0, \infty)$. This yields $u^1(t) = \theta^1(t) = 0$. Player 2's best response to this strategy can be found in the same way as the Pareto optimal solution. It is given by $u^2(t) = \theta^2(x(t))$ with

$$\theta^2(x) = \begin{cases} 0 & \text{if } x \leq r, \\ -1 & \text{if } x > r. \end{cases}$$

(iii) If $x_0 = 1 + r$ then it follows that the Pareto optimal solution is constant with $\tilde{x}(t) = 1 + r$, $\tilde{u}^1(t) = 1$, and $\tilde{u}^2(t) = 0$ for all $t \in [0, \infty)$. By evaluating the objective functionals one obtains

$$J^1(\tilde{\phi}^1; \tilde{\phi}^2, \tilde{u}(\cdot), t) = 1 + \frac{1}{2r} \quad \text{and} \quad J^2(\tilde{\phi}^2; \tilde{\phi}^1, \tilde{u}(\cdot), t) = \frac{1}{r} - \frac{(1+r)^2}{2r}.$$

To verify (6.17) one has to calculate $V^i(1 + r, t; \theta)$, $i = 1, 2$. For the initial state $x_0 = 1 + r$ the Markov perfect Nash equilibrium generates the control paths $u^1(t) = 0$ for all $t \in [0, \infty)$ and

$$u^2(t) = \begin{cases} -1 & \text{for } t \in [0, 1), \\ 0 & \text{for } t \in [1, \infty). \end{cases}$$

The corresponding state trajectory is given by

$$x(t) = \begin{cases} 1 + r - t & \text{for } t \in [0, 1), \\ r & \text{for } t \in [1, \infty). \end{cases}$$

Evaluating the objective functionals along this solution yields

$$V^1(\tilde{x}(t), t; \theta) = 1 + \frac{1}{r} - \frac{1}{r^2}(1 - e^{-r})$$

and

$$V^2(\tilde{x}(t), t; \theta) = \frac{1}{r^2} - \frac{(1+r)^2}{2r} - \frac{1}{r^3}(1 - e^{-r}).$$

Condition (6.17) holds for $i = 2$ for all $r > 0$ but it holds for $i = 1$ if and only if $e^{-r} < 1 - r/2$.

(iv) If $x_0 \leq r$, the Pareto optimal solution is given by $\tilde{x}(t) = x_0$, $\tilde{u}^1(t) = 1$, and $\tilde{u}^2(t) = 0$ for all $t \in [0, \infty)$. This yields

$$J^1(\tilde{\phi}^1;\tilde{\phi}^2,\tilde{u}(\cdot),t)=x_0/r-1/(2r)\ \text{ and }\ J^2(\tilde{\phi}^2;\tilde{\phi}^1,\tilde{u}(\cdot),t)=1/r-x_0^2/(2r).$$

The Markov perfect Nash equilibrium generates constant paths as well with $x(t)=x_0$, $u^1(t)=0$, and $u^2(t)=0$ for all $t\in[0,\infty)$, which implies that

$$V^1(\tilde{x}(t),t;\theta)=x_0/r\ \text{ and }\ V^2(\tilde{x}(t),t;\theta)=-x_0^2/(2r).$$

Condition (6.17) for $i=2$ is satisfied for $\epsilon<1/r$ but it cannot be satisfied for any nonnegative ϵ if $i=1$.

Chapter 7

1. The open-loop equilibrium strategies of the knowledge accumulation game are

$$u^i(t)=\left(W^i-\frac{1}{r+\alpha}\right)e^{(r+\alpha)(t-T)}+\frac{1}{r+\alpha}.$$

It is easily seen that $\tilde{u}^i(t)=\lim_{T\to\infty}u^i(t)=1/(r+\alpha)$ is an equilibrium strategy for the infinite horizon game.

2. The existence of a Markov perfect equilibrium with constant strategies can be established by assuming a value function of the form $V^i(x)=\lambda^i x$, where λ^1 and λ^2 are constants. It can now be established that the optimal prices that correspond to this equilibrium are characterized by

$$u^i\left(1+\frac{1}{\eta^i}\right)=c^i+\lambda^i\left(1+\frac{\alpha^j_{u^i}(u^j,u^i)}{\alpha^i_{u^i}(u^i,u^j)}\right),$$

where $\eta^i=u^i\alpha^i_{u^i}(u^i,u^j)/\alpha^i(u^i,u^j)$ is the price elasticity of demand. Since $\lambda^i>0$ and the effect of a price change of firm i on its own demand is higher than on the demand of the rival, we see immediately that the dynamic prices are higher than the static ones.

3. To derive the Markov perfect Nash equilibrium by means of the maximum principle one uses the costate equation

$$\dot{\lambda}^i(t)=\left[r+\alpha^i(u^i(t),u^j(t))+\alpha^j(u^j(t),u^i(t))\right]\lambda^i(t)-[u^i(t)-c^i]\alpha^i(u^i(t),u^j(t))$$

and the maximum condition

$$u^i(t)\left[1+\frac{1}{\eta^i(t)}\right]=c^i+\lambda^i(t)\left[1+\frac{\alpha^j_{u^i}(u^j(t),u^i(t))}{\alpha^i_{u^i}(u^i(t),u^j(t))}\right],$$

where, as in the previous exercise, $\eta^i(t)$ is the price elasticity.

4. If we make use of the following state variable transformation

$$z(t)=y(t)^\alpha$$

we get by means of differentiation with respect to time t

$$\frac{\dot{z}(t)}{z(t)}=\alpha\frac{\dot{y}(t)}{y(t)},$$

which demonstrates that the game can be transformed into a linear state game.

5. The open-loop equilibrium is characterized by $\dot{\lambda}^1(t)=r\lambda^1(t)-\pi^1$, $\dot{\lambda}^2(t)=r\lambda^2(t)+\pi^2$, $\lambda^1(t)=2[u^1(t)]^2$, and $\lambda^2(t)=-2[u^2(t)]^2$. If we differentiate

the last two equations with respect to time and substitute from the first two equations we get

$$\dot{u}^1(t) = \frac{1}{4u^1(t)}\{2r[u^1(t)]^2 - \pi^1\},$$
$$\dot{u}^2(t) = \frac{1}{4u^2(t)}\{2r[u^2(t)]^2 - \pi^2\}.$$

A phase diagram analysis can easily be conducted.

Chapter 8

1. Verify the conditions of theorem 8.1 with the value function

$$V(h,x) = \begin{cases} 0 & \text{if } h = 1 \text{ and } x < 1 - 1/r, \\ \dfrac{(r-1)^{2+r}x^{-(1+r)}}{r^{2+r}(1+r)(2+r)} + \dfrac{x}{2+r} + \dfrac{1-r}{r(1+r)} & \text{if } h = 1 \text{ and } x \geq 1 - 1/r, \\ 1/r & \text{if } h = 2. \end{cases}$$

2. For the stationary Markovian Nash equilibrium, verify the conditions of theorem 8.2 with

$$V^1(1,x) = -(498 - 520x + 169x^2)/676,$$
$$V^1(2,x) = -(3388 - 3640x + 1183x^2)/2704,$$

$V^2(1,x) = V^1(2,-x)$, and $V^2(2,x) = V^1(1,-x)$. For the piecewise open-loop Nash equilibrium, verify the conditions of theorem 8.3 with

$$V^1(1,x,y,t) = -(498 - 520x + 169x^2)/676,$$
$$\begin{aligned} V^1(2,x,y,t) = -[&4(1841 - 2730x + 1183x^2) \\ &+ 8e^{-t/2}(300 - 260x + 195y - 169xy) \\ &+ e^{-t}(400 + 520y + 169y^2)]/8112, \end{aligned}$$

$V^2(1,x,y,t) = V^1(2,-x,-y,t)$, and $V^2(2,x,y,t) = V^1(1,-x,-y,t)$. The state trajectories are given by

$$x_{1y}(s) = \left(y - \frac{20}{13}\right)e^{-s/2} + \frac{20}{13},$$
$$x_{2y}(s) = \left(y + \frac{20}{13}\right)e^{-s/2} - \frac{20}{13}$$

in both the stationary Markovian Nash equilibrium and the piecewise open-loop Nash equilibrium.

3.

$$A = \left[\frac{r - \alpha\beta + \alpha(1-\alpha)\gamma^2/2}{1-\alpha} \right]^{\alpha-1}.$$

To prove that $x(t)$ as defined in the exercise solves the state equation, use Itô's lemma with $G(x, t) = \ln x$. Finally,

$$E_{u(\cdot)} V(x(t), t) = Ax(0)^\alpha e^{\frac{\alpha t}{1-\alpha}[\beta - r - (1-\alpha)\gamma^2/2]}$$

so that $e^{-rt} E_{u(\cdot)} V(x(t), t)$ converges to 0 for $t \to \infty$ because of the parameter restrictions.

4. The parameter restrictions are $0 < \alpha < 1/N$, $\max\{0, \alpha\beta - \alpha(1-\alpha)\gamma^2/2\} < r$. The solution is described by

$$A = \left[\frac{r - \alpha\beta + \alpha(1-\alpha)\gamma^2/2}{1-\alpha N} \right]^{\alpha-1},$$

$$u^i = \phi(x, t) = \frac{r - \alpha\beta + \alpha(1-\alpha)\gamma^2/2}{1-\alpha N} x,$$

$$x(t) = x(0) \exp\left\{ \left[\frac{\beta - rN}{1-\alpha N} - \frac{(1-\alpha^2 N)\gamma^2}{2(1-\alpha N)} \right] t + \gamma w(t) \right\},$$

$$E_{u(\cdot)} V(x(t), t) = Ax(0)^\alpha e^{\frac{\alpha t}{1-\alpha N}[\beta - rN - (1-\alpha)\gamma^2/2]}.$$

5. Note that the optimal value functions for the Markovian Nash equilibrium derived in exercise 4.2. are linear with respect to the state. Consequently, they also solve the HJB equations of the stochastic problem.
6. The equilibrium strategies are given by $u^i(t) = \phi(x(t), t)$, with $\phi(x, t) = 2A(t)x$ and

$$A(t) = \frac{(r - \alpha^2 + C)\left[(r - \alpha^2 - C)e^{C(t-T)} - r + \alpha^2 + C\right]}{4(2N-1)\left[(r - \alpha^2 - C)e^{C(t-T)} - r + \alpha^2 - C\right]}.$$

Here, $C = \sqrt{(r-\alpha^2)^2 + 8N - 4}$. To prove this, it is necessary to verify the conditions of theorem 8.5 with $V^i(x, t) = A(t)x^2$. The equilibrium state process is given by

$$x(t) = x(0)e^{2N\int_0^t A(s)\,ds - (\alpha^2/2)t + \alpha w(t)}.$$

Chapter 9

1. Dynamic programming may be used to derive a Markov perfect Nash equilibrium. Use the quadratic value function

$$V^i(K^i, K^j) = \alpha + \beta_i K^i + \gamma K^j + (\delta/2)(K^i)^2 + (\epsilon/2)(K^j)^2 + \sigma K^i K^j$$

and show that the constants α, β_i, γ, δ, σ, and ϵ satisfy equations similar to those stated in example 9.1. In particular, the investment strategies become $\phi^i(K^i, K^j) = (1/k)(\beta_i + \delta K^i + \sigma K^j)$, with $\sigma < 0$. The steady state of this equilibrium corresponds to a static conjectural variations equilibrium with a negative conjecture (because $\sigma < 0$). This implies that the steady state level in the Markov perfect Nash equilibrium is higher than in the open-loop game, which corresponds to a conjecture of zero.

2. The open-loop equilibrium is characterized by the following equations

$$\dot{K}^1(t) = I^1(t) - \delta^1 K^1(t),$$
$$\dot{K}^2(t) = I^2(t) - \delta^2 K^2(t),$$
$$\dot{\lambda}_1(t) = (r + \delta^1)\lambda_1(t) - \pi^1_{K^1}(K^1(t), K^2(t)),$$
$$\dot{\lambda}_2(t) = (r + \delta^2)\lambda_2(t) - \pi^2_{K^2}(K^2(t), K^1(t)),$$

with the maximum conditions $\lambda_i(t) = kI^i(t)$. Hence we can derive a differential equation system in the state and control variables or the state and costate variables.

3. The differential game is given by the objective functionals

$$\int_0^\infty e^{-rt}\{[1 - c^i - Q^i(t) - Q^j(t)]Q^i(t) - (1/2)P(t)^2\}\,dt$$

and the state equation

$$\dot{P}(t) = Q^1(t) + Q^2(t) - \delta P(t).$$

This game, however, has the same structure as the knowledge accumulation game discussed in section 9.5. Hence the open-loop and the Markov perfect Nash equilibrium can be derived following the calculations presented in that section.

4. Follow section 9.5.
5. Let i = 2 be the private firm. Its output is given by $Q^2=(1-c^2-Q^1)/2$. Given that firm 1 maximizes social welfare, output levels are higher than in the static Cournot model.

Chapter 10

1. Inserting the conjectured value function from (10.10) and its derivative $V'(p) = Kp - E$ into the HJB equation for case (B) yields

$$r(Kp^2/2 - Ep + G) = (p - c)[p - c - s(Kp - E)] - [p - c - s(Kp - E)]^2/2 + (Kp - E)s\{a - p - 2[p - c - s(Kp - E)]\}.$$

Collect the coefficients of p^2, p^1, and p^0, respectively, and equate the three resulting expressions to zero (since the HJB equation must be satisfied for all p). From the coefficients of p^2 we obtain a quadratic equation in K, namely, $3s^2K^2 + 1 - K(r + 6s) = 0$. Further, from the coefficients of p^1 we obtain $E = (-asK + c - 2scK)/(r - 3s^2K + 3s)$ and from those of p^0 we get

$G = [c^2 + 3s^2E^2 - 2sE(2c + a)]/(2r)$. For more details, see Fershtman and Kamien [89].

2. When the game is played over a fixed and finite horizon, the HJB equations in (10.7) must be modified as follows. The value functions must depend on state p as well as time t since a stationarity assumption will not work when the horizon is finite. Moreover, on the left-hand side of (10.7) we must subtract the term $V_t^i(p, t)$ (the partial derivative of the value function with respect to time, cf. equation (4.10)). When outputs are positive, we conjecture quadratic value functions

$$V^i(p, t) = \frac{1}{2}K_i(t)2p^2 - E_i(t)p + G_i(t).$$

Differentiate these value functions partially with respect to p and t and insert the resultant derivatives into the HJB equations. This yields a system of six differential equations for the coefficients $K_i(t)$, $E_i(t)$, $G_i(t)$. It can be proved (see Fershtman and Kamien [90]) that equilibrium output strategies must be symmetric, that is, we can delete the subscript i. We end up with a system of three differential equations for the coefficient functions $K(\cdot)$, $E(\cdot)$, and $G(\cdot)$ which must be solved subject to the terminal conditions $K(T) = E(T) = G(T) = 0$. The latter follow from condition (iii) of theorem 4.1. A complete solution of the exercise, including a treatment of the case where output is zero, can be found in Fershtman and Kamien [90].

3. To identify a Markovian Nash equilibrium in the R&D game we have the N HJB equations

$$\begin{aligned}-V_t^i(y, t) = \max\Big\{&y\Big[\lambda P_I u_i + \lambda P_F \sum_{j\neq i}\phi_j(y, t) - e^{-rt}u_i^2/2\Big]\\ &- V_y^i(y, t)\lambda y\Big[u_i + \sum_{j\neq i}\phi_j(y, t)\Big]\,\Big|\, u_i \geq 0\Big\}\end{aligned}$$

and we confine our interest to interior solutions (as in the maximum principle approach in section 10.2). Candidate strategies are given by $\phi_i(y, t) = \lambda e^{rt}[P_I - V_y^i(y, t)]$ (see also (10.22)). Substituting these strategies into the curly brackets on the right-hand side of the HJB equations yields a system of N ordinary differential equations for the value functions. We look for a symmetric Markovian Nash equilibrium (which is plausible since the game is symmetric) and conjecture a value function of the form $V^i(y, t) = [P_I + b(t)]y$, where the function $b(\cdot)$ must satisfy the differential equation stated in (10.26) and the terminal condition $b(T) = -P_I$. The Markovian Nash equilibrium strategies are those reported at the end of section 10.2. For a complete solution of this exercise, see Reinganum [198].

4. For $P_F = 0$ the R&D strategy becomes

$$\phi(t) = \frac{2\lambda P_I(N-1)e^{rt}}{2N - 1 - \exp[P_I\lambda^2(N-1)(e^{rt} - e^{rT})/r]}.$$

It is not difficult to prove that this strategy involves greater effort for all t than the strategy where $P_F > 0$. (For details, see Reinganum [198, proposition 3].)

This result is intuitive: a firm uses greater effort if it can be assured of collecting the whole reward (provided that it is successful). An increase in the number of firms results in an increase in the equilibrium R&D effort. Thus, an increased number of rival firms tends to hasten innovations. To see this, differentiate partially the equilibrium strategy of section 10.2 with respect to N and the conclusion readily follows.

5. Changing the dynamics to

$$\dot{y}(t) = -\alpha\lambda y(t)N\sum_{i=1}^{N} u_i(t)$$

leads to the candidate strategies $\phi_i(y, t) = \lambda e^{rt}\left[P_I - \alpha\mu_i(t)y^{(\alpha-1)/\alpha}\right]$, which for $\alpha = 1$ becomes (10.22). Supposing strategies of the same structure as in section 10.2 leads to

$$\mu(t) = \frac{P_I + b(t)}{\alpha} y(t)^{(1-\alpha)/\alpha},$$

which for $\alpha = 1$ becomes (10.25). Proceeding from here in a similar way as in section 10.2 one sees that the function $b(\cdot)$ must satisfy the differential equation

$$\dot{b}(t) = \lambda^2 e^{rt}\Big\{ - b(t)^2[\alpha(N-1) + 1/2] + b(t)(N-1)[\alpha P_F - (2\alpha - 1)P_I] + (N-1)(\alpha - 1)P_I(P_F - P_I)\Big\},$$

which for $\alpha = 1$ becomes (10.26). The remaining details can be found in Dockner et al. [51]. Notice that in the game with the modified dynamics, Markovian strategies are no longer degenerate: they depend on state y as well as time t.

6. The cooperative R&D strategy is given by

$$\phi^{\text{coop}}(t) = \frac{(1+\gamma)\lambda P_I e^{rt}}{1 - P_I\lambda^2(1+\gamma)^2(e^{rt} - e^{rT})/r},$$

which is obtained by maximizing the two firms' joint profits

$$J = 2\int_0^T y(t)[\lambda P_I u(t) - (e^{-rt}/2)u(t)^2]\,dt.$$

This profit functional is obtained from (10.19) by putting $N = 2$ and $P_F = 0$ and by introducing the state transformation $y(t) = e^{-\lambda[z_1(t)+z_2(t)]}$. In the noncooperative case we obtain from section 10.2, for $N = 2$, the strategy $\phi(t) = 2\lambda P_I e^{rt}/[3 - e^{m(t)}]$, where $m(t) = \lambda^2 P_I(e^{rt} - e^{rT})/r$. Comparing $\phi^{\text{coop}}(t)$ and $\phi(t)$ shows that for $\gamma = 0$ the former is greater than the latter for all $t < T$. This means that, on average, noncooperative duopolists make the innovation sooner than a joint venture in which the firms are unable to share know-how. Details for this exercise can be found in Reinganum [197, 200].

Chapter 11

1. For the symmetric game, the solutions of the adjoint equations in (11.5) are symmetric and we obtain $\dot{\eta}(t) = r\eta(t) - q + \eta(t)^2/c$. The solution of this equation with terminal condition $\eta(T) = S$ is

$$\eta(t) = \frac{c[\mu - r + \alpha(\mu + r)e^{\mu t}]}{2(1 - \alpha e^{\mu t})},$$

in which $\mu = \sqrt{r^2 + 4q/c}$ and

$$\alpha = e^{-\mu T} \frac{S - (c/2)(\mu - r)}{S + (c/2)(\mu + r)}.$$

This shows that $\eta(t)$ is decreasing over time if $S < (c/2)(\mu - r)$, and otherwise increasing.

2. The HJB equations for a stochastic differential game are given by (8.31). We see that these equations differ from those for the deterministic game, stated in (11.10), only by the terms $(1/2)\sigma^2(x)V^i_{xx}(x, t)$, which, however, vanish in the model at hand where the value functions are linear in state x. Hence the value functions given by (11.12) also solve the HJB equations for the stochastic game and the Markovian Nash equilibrium strategies identified in the deterministic game yield an equilibrium in the stochastic game, too. We conclude that the strategies given by (11.16) are robust to exogenous random disturbances. The reader should note that this is not a general result. It pertains to the specific model only. In many other cases one will see quite dramatic differences between the outcome of a stochastic differential game and that of its deterministic counterpart.

3. Note that feasible advertising strategies take values in $(0, \infty)$, that is, zero advertising is excluded. Current-value Hamiltonians are given by

$$H^i(S, A_i, \lambda_i, t) = q_i S_i - A_i + \lambda_i k \log(A_i/A_j(t)), \; i = 1, 2, \; i \neq j,$$

where $S_1 = S$ and $S_2 = M - S$. The equilibrium conditions are $\dot{\lambda}_1(t) = r_1\lambda_1(t) - q_1$, $\dot{\lambda}_2(t) = r_2\lambda_2(t) + q_2$, $\lambda_1(T) = \lambda_2(T) = 0$, $A_1(t) = k\lambda_1(t)$, and $A_2(t) = -k\lambda_2(t)$, which shows that advertising rates are zero at the terminal instant of time. It is easy to solve the costate equations, and inserting these solutions into the candidate advertising rates yields

$$A_i(t) = \frac{kq_i}{r_i}\left[1 - e^{r_i(T-t)}\right].$$

These advertising rates are exponentially decreasing over time. If discount rates are zero it is straightforward to show that advertising rates are linearly decreasing over time. For a proof that the above conditions are actually sufficient for the equilibrium property as well as further details, see Jørgensen [131].

4. The costate equations are the same as in exercise 3, except that the transversality condition must be replaced by $\lambda_1(T) = \sigma_1$ and $\lambda_2(T) = -\sigma_2$. The costate equations are explicitly solvable and we obtain

$$\lambda_1(t) = q_1/r_1 + [\sigma_1 - q_1/r_1]e^{-r_1(T-t)}$$
$$\lambda_2(t) = -q_2/r_2 - [\sigma_2 - q_2/r_2]e^{-r_2(T-t)},$$

which shows that the signs of the costate variables are determined by the relationship between σ_i, r_i, and q_i. Candidate advertising strategies are (implicitly) given by the equations

$$\lambda_i(t)\frac{\partial g(A_1(t), A_2(t))}{\partial A_i} = 1,$$

which have a unique solution due to the assumptions on the function g. An analysis of the solution can proceed by phase plane analysis, see Feichtinger and Dockner [83].

5. Define the transformed state variable $S(t) = \sum_{i=1}^{N} S_i(t)$ and note that

$$\dot{S}(t) = \sum_{i=1}^{N}[a + b\log A_i(t) + dS(t)][M - S(t)]$$
$$= [M - S(t)][Na + b\sum_{i=1}^{N}\log A_i(t) + dNS(t)].$$

The Hamiltonian of firm $i \in \{1, 2, \ldots, N\}$ is given by

$$H_i(S, A_i, \lambda_i, t) = (p_i - c_i)[a + b\log A_i + dS](M - S)$$
$$+\lambda_i(M - S)\Big[Na + b\log A_i + \textstyle\sum_{j\neq i}\log A_j(t) + dNS\Big] - A_i,$$

where λ_i is the costate variable of firm i associated with the state S. If the equilibrium advertising rate of firm i is positive, a necessary condition for maximization of the Hamiltonian is $b[M - S(t)][p_i - c_i + \lambda_i(t)] = A_i(t)$. The costate satisfies

$$\dot{\lambda}_i(t) = \lambda_i(t)\left[r_i + Na + b\sum_{j=1}^{N}\log A_j(t) + 2NdS(t) - dNM\right]$$
$$+ (p_i - c_i)[a + b\log A_i(t) + 2dS(t) - dM].$$

Differentiate in the Hamiltonian maximization condition to obtain, for instance, that $\dot{A}_i(t) < 0$. Thus, equilibrium advertising rates are decreasing over time. For other results, see Dockner and Jørgensen [55].

6. Start by using the costate equations in (11.24) and (11.25) and the result follows from straightforward manipulations.

Chapter 12

1. Try a value function of the form $V_i(x) = A\ln x + B$, and use the HJB equations.
2. Try a value function of the form $V_i(x) = (\alpha/r)\ln x + B$, and use the HJB equations.
3. (a) Note that cooperation among the oligopolists means trying to achieve the outcome under a monopoly. The monopolist maximizes the integral of the

discounted stream of profit flows $P(Q(t))Q(t)$ under the resource constraint. Let $r > 0$ be the discount rate. Such a problem makes sense only if $\eta > 1$ because if $\eta < 1$ then $P(Q)Q$ will be decreasing in Q, and if $\eta = 1$ then $P(Q)Q$ is a constant. Let us assume $\eta > 1$. Then the monopolist's optimal time path of extraction is $Q_M(t) = \eta r x(0) e^{-\eta r t}$. This path satisfies

$$\left(1 - \frac{1}{\eta}\right) Q_M(t)^{-1/\eta} = \psi_M(0) e^{rt}$$

where $\psi_M(0) = [1 - (1/\eta)](\eta r x_0)^{-1/\eta}$. On the other hand, under oligopoly, the oligopolist i's extraction path satisfies

$$\frac{-q_i(t)}{\eta}[Q_{-i}(t) + q_i(t)]^{-1-(1/\eta)} + [Q_{-i}(t) + q_i(t)]^{-1/\eta} = \psi_i(0) e^{rt},$$

where $Q_{-i}(t) = \sum_{j \neq i} q_j(t)$. If $N > 1/\eta$, then we can set $q_i(t) = (1/N) Q_M(t)$ for all i and

$$\psi_i(0) = \psi_M(0) \frac{N\eta - 1}{N(\eta - 1)} > 0.$$

This shows that the open-loop Nash equilibrium of the oligopolistic game (when only strictly feasible replies are allowed) yields the same path of industry output as under monopoly. Note that the second order condition for maximizing the Hamiltonian is satisfied if $2N - 1 > 1/\eta$.

(b) Consider the case where firms are allowed to choose strategies from a larger set, namely the set of weakly feasible open-loop replies. In this case, if a firm j plans to have a positive output for all $t \leq T$, then firm i is allowed to find a best reply that implies exhaustion at some $T' \leq T$. (The following argument is based on Bolle [10].) A symmetric equilibrium where the resource is exhausted at some finite T exists only if $dJ_i(T)/dT \geq 0$, where

$$J_i(T) = \max_{q_i(\cdot)} \int_0^T e^{-rt} q_i(t) P(Q_{-i}(t) + q_i(t))\, dt$$

subject to

$$\dot{A}(t) = q_i(t), \quad A(0) = 0,$$

$$A(T) \leq x_0 - \int_0^T Q_{-i}(t)\, dt.$$

Now

$$dJ_i(T)/dT = e^{-rT}\{q_i(T) P(Q(T)) - [P(Q(T)) + q_i(T) P'(Q(T))] Q(T)\},$$

where $Q(T) = Q_{-i}(T) + q_i(T)$. Using symmetry, $Q_{-i}(T) = (N-1) q_i(T)$, we get

$$dJ_i(T)/dT = e^{-rT} P(Q_{-i}(T) + q_i(T)) q_i(T)[(1/\eta) - N + 1],$$

which is nonnegative only if $N-1 \leq 1/\eta$. This inequality is also necessary for the existence of a symmetric open-loop Nash equilibrium with asymptotic exhaustion, if players are allowed to use weakly feasible open-loop replies.

4. Let $c^*(\cdot)$ denote the cooperative control path of the representative player, and let $T^* = (1/r)\ln((eN + rx_0)/eN)$. Show that, if players are noncooperative, then $H_i(T^*) = \ln(c_i(T^*)) - [1/c_i(T^*)][c_i(T^*) + (N-1)c^*(T^*)]$ is different from zero if $c_i(T^*) = c^*(T^*)$.
5. Assume $\alpha \leq 1/2$ to ensure concavity of the Hamiltonian of each player. Without loss of generality, set $b = 1$. In a symmetric open-loop Nash equilibrium where the effectiveness declines to zero as time tends to infinity, the equilibrium control path is

$$a_i(t) = \frac{rx_0}{1+(1-2\alpha)N} e^{-rNt/[1+(1-2\alpha)N]}$$

and $\dot{x}(t)/x(t) = -rN/[1+(1-2\alpha)N]$, implying a faster rate of depletion than that under cooperation, where $\dot{x}(t)/x(t) = -r/[2-2\alpha]$.

6. The cooperative solution is $\alpha_i(t) = (1+r)/2$ for $i = 1, 2$ and for all t. The open-loop Nash equilibrium for this game is $\alpha_i(t) = 1 + r$ independently of the initial state. It follows that this equilibrium is subgame perfect. The value functions

$$V_i^D(y) = \frac{y}{1+r} + \frac{1}{r}\left[\ln(1+r) + \frac{1}{1+r} - 2\right]$$

satisfy the HJB equations. Here, the superscript D denotes non-cooperation.

7. If h is the time interval that must elapse between the time of cheating and the time of detection, and if player 2 wants to cheat at time 0 while player 1 remains cooperative (i.e., $\alpha_1 = (1+r)/2$) until detection time then, to maximize his payoff, player 2 must solve the optimal control problem

$$\max_{E_2} \int_0^h e^{-rt}[y(t) + \ln \alpha_2(t)]\, dt + e^{-rh} V_2^D(y(h))$$

subject to

$$\dot{y}(t) = 1 - y(t) - \frac{1+r}{2} - \alpha_2(t)$$

and $y(0) = y_0$. Show that this control problem yields $\alpha_2(t) = 1 + r$ for all $t \in [0, h]$. (A phase diagram might help, bearing in mind the transversality condition at h.) Next, compute player 2's payoff given that he cheats at time 0. Show that this payoff is smaller than $V_i^D(y_0)$ if h and r are small.

8. For the cooperative solution, let $E(t) = E_1(t) = E_2(t)$ because of symmetry. Define $z(t) = E(t)/x(t)$. Show that when effort levels are positive, the following equations hold:

$$-\frac{z(t)^{1/2}}{2}\frac{\dot{z}(t)}{z(t)} = [1 - z(t)^{1/2}][r - 1 + 2x(t)] - 2z(t)$$

$$\frac{\dot{x}(t)}{x(t)} = 1 - x(t) - 4z(t)^{1/2}$$

Thus in the steady state

$$\left[1 - \frac{1 - x(t)}{4}\right][r - 1 + 2x(t)] = 2\left[\frac{1 - x(t)}{4}\right]^2.$$

This yields the root

$$x^* = \frac{1}{3}\left(-6 - r + \sqrt{57 - 6r + r^2}\right) > 0.$$

For the open-loop Nash equilibrium, we obtain in the steady state

$$\left[1 - \frac{1 - x(t)}{4}\right]\left[r - 1 + 2x(t) + \frac{1 - x(t)}{4}\right] = \left[\frac{1 - x(t)}{4}\right]^2$$

and this yields

$$x^{**} = \frac{1}{3}\left(-5 - r + \sqrt{40 - 8r + r^2}\right) > 0.$$

It is easy to see that $x^* > x^{**}$.

Bibliography

[1] Arnold, L. (1974): *Stochastic Differential Equations: Theory and Applications*, John Wiley & Sons, New York

[2] Barro, R. J., and X. Sala-i-Martin (1995): *Economic Growth*, McGraw-Hill, New York

[3] Başar, T., and A. Haurie (1984): Feedback equilibria in differential games with structural and modal uncertainties, in Cruz, J. B. (ed.), *Advances in Large Scale Systems*, JIT Press, New York, 163–201

[4] Başar, T., and G. J. Olsder (1995): *Dynamic Noncooperative Game Theory*, 2nd edn., Academic Press, London

[5] Benchekroun, H., and N. V. Long (1998): Efficiency inducing taxation for polluting oligopolists, *Journal of Public Economics* 70, 325–342

[6] Benchekroun, H., N. V. Long, and H. Tian (1999): Learning-by-doing and strategic trade policy, *Review of International Economics*, 7, 493–508

[7] Benhabib, J., and R. Radner (1992): The joint exploitation of a productive asset: a game-theoretic approach, *Economic Theory* 2, 155–190

[8] Binmore, K. (1992): *Fun and Games*, D. C. Heath, Lexington

[9] Bolle, F. (1980): The efficient use of a non-renewable common-pool resource is possible but unlikely, *Journal of Economics* 40, 391–397

[10] Bolle, F. (1986): On the oligopolistic extraction of non-renewable common-pool resources, *Economica* 53, 519–527

[11] Breton, M., and A. Haurie (1989): Two-layer piecewise deterministic games, *Proceedings of the 28th IEEE Conference on Decision and Control*, IEEE Publications, New York, 198–200

[12] Brock, W. A., and A. G. Malliaris (1989): *Differential Equations, Stability and Chaos in Dynamic Economics*, North-Holland, Amsterdam

[13] Buiter, W. H. (1983): Optimal and time-consistent policies in continuous-time rational expectations models, Technical Working Paper No. 29, NBER, Cambridge

[14] Calvo, G. A. (1978): On the time consistency of optimal policy in a monetary economy, *Econometrica* 46, 1411–1428

[15] Carlson, D. A., and A. Haurie (1991): *Infinite Horizon Optimal Control: Theory and Applications*, 2nd edn, Springer-Verlag, Berlin

[16] Case, J. (1979): *Economics and the Competitive Process*, New York University Press, New York

[17] Cave, J. (1987): Long-term competition in a dynamic game: the cold fish war, *RAND Journal of Economics* 18, 596–610

[18] Chiang, A. C. (1992): *Elements of Dynamic Optimization*, McGraw-Hill, New York

[19] Chiarella, C., M. C. Kemp, N. V. Long, and K. Okuguchi (1984): On the economics of international fisheries, *International Economic Review* 25, 85–92

[20] Chintagunta, P. K. (1993): Investigating the sensitivity of equilibrium profits to advertising dynamics and competitive effects, *Management Science* 39, 1146–1162

[21] Chintagunta, P. K., and D. Jain (1992): A dynamic model of channel member strategies for marketing expenditures, *Marketing Science* 11, 168–188

[22] Chintagunta, P. K., and D. Jain (1995): Empirical analysis of a dynamic duopoly model of competition, *Journal of Economics and Management Strategy* 4, 109–131

[23] Chintagunta, P. K., and V. R. Rao (1996): Pricing strategies in a dynamic duopoly: a differential game model, *Management Science* 42, 1501–1514

[24] Chintagunta, P. K., V. R. Rao, and N. J. Vilcassim (1993): Equilibrium pricing and advertising strategies for nondurable experience products in a dynamic duopoly, *Managerial and Decision Economics* 14, 221–234

[25] Chintagunta, P. K., and N. J. Vilcassim (1992): An empirical investigation of advertising strategies in a dynamic duopoly, *Management Science* 38, 1230–1244

[26] Chintagunta, P. K., and N. J. Vilcassim (1994): Marketing investment decisions in a dynamic duopoly: a model and empirical analysis, *International Journal of Research in Marketing* 11, 287–306

[27] Clark, C. W. (1990): *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*, 2nd edn, Wiley, New York

[28] Clarke, F. H. (1983): *Optimization and Nonsmooth Analysis*, Wiley, New York

[29] Clarke, F. H., and R. B. Vinter (1983): Local optimality conditions and Lipschitzian solutions to the Hamilton–Jacobi equation, *SIAM Journal of Control and Optimization* 21, 856–870

[30] Clemhout, S., and H. Y. Wan, Jr. (1985): Dynamic common-property resources and environmental problems, *Journal of Optimization Theory and Applications* 46, 471–481

[31] Clemhout, S., and H. Y. Wan, Jr. (1991): Environmental problem as a common-property resource game in Ehtamo, H., and R. P. Hämäläinen (eds.), *Dynamic Games in Economic Analysis*, Springer-Verlag, New York, 132–154

[32] Clemhout, S., and H. Y. Wan, Jr. (1994): Differential games–economic applications, in Aumann, R. J., and S. Hart (eds.), *Handbook of Game Theory with Economic Applications II*, Elsevier, Amsterdam, 801–825

[33] Cohen, D., and P. Michel (1988): How should control theory be used to calculate a time-consistent government policy?, *Review of Economic Studies* 55, 263–274

[34] Cornes, R., N. V. Long, and K. Shimomura (1995): Drugs and pests: intertemporal productive externalities, Working Paper, Australian National University, Canberra

[35] Cournot, A. (1838): *Recherches sur les principes mathématiques de la théorie des richesses*, L. Hachette, Paris

[36] Crabbé, P., and N. V. Long (1993): Entry deterrence and overexploitation of the fishery, *Journal of Economic Dynamics and Control* 17, 679–704

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

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

[39] Davis, M. H. A., M. A. H. Dempster, S. P. Sethi, and D. Vermes (1987): Optimal capacity expansion under uncertainty, *Advances in Applied Probability* 19, 156–176

[40] Debreu, G. (1952): A social equilibrium existence theorem, *Proceedings of the National Academy of Sciences* 38, 886–893

[41] Dixit, A., and B. Nalebuff (1991): *Thinking Strategically*, W. W. Norton, New York

[42] Dockner, E. J. (1985): Local stability analysis in optimal control problems with two state variables, in Feichtinger, G. (ed.), *Optimal Control Theory and Economic Analysis 2*, North-Holland, Amsterdam, 89–103

[43] Dockner, E. J. (1985): Optimal pricing in a dynamic duopoly game model, *Zeitschrift fur Operations Research* 29, 1–16

[44] Dockner, E. J. (1988): On the relation between dynamic oligopolistic competition and long-run competitive equilibrium, *European Journal of Political Economy* 4, 47–64

[45] Dockner, E. J. (1990): Strategic trade policy with uncertain entry, Working Paper, Department of Business Administration, University of Vienna

[46] Dockner, E. J. (1992): A dynamic theory of conjectural variations, *Journal of Industrial Economics* 40, 377–395

[47] Dockner, E. J., and G. Feichtinger (1986): Dynamic advertising and pricing in an oligopoly: a Nash equilibrium approach, in Başar, T. (ed.), *Dynamic Games and Applications in Economics*, Springer-Verlag, Berlin, 238–251

[48] Dockner, E. J., and G. Feichtinger (1991): On the optimality of limit cycles in dynamic economic systems, *Journal of Economics* 53, 31–50

[49] Dockner, E. J., G. Feichtinger, and S. Jørgensen (1985): Tractable classes of nonzero-sum open-loop Nash differential games: theory and examples, *Journal of Optimization Theory and Applications* 45, 179–197

[50] Dockner, E. J., G. Feichtinger, and A. Mehlmann (1989): Noncooperative solutions for a differential game model of fishery, *Journal of Economic Dynamics and Control* 13, 1–20

[51] Dockner, E. J., G. Feichtinger, and A. Mehlmann (1993): Dynamic R&D competition with memory, *Journal of Evolutionary Economics* 3, 145–152

[52] Dockner, E. J., and A. Gaunersdorfer (1996): Strategic new product pricing when demand obeys saturation effects, *European Journal of Operational Research* 90, 589–598

[53] Dockner, E. J., J. M. Hartwick, and G. Sorger (1995): Dynamic duopolistic competition with sticky prices: a note, Working Paper, Department of Economics, University of Vienna

[54] Dockner, E. J., and S. Jørgensen (1988): Optimal pricing strategies for new products in dynamic oligopolies, *Marketing Science* 7, 315–334

[55] Dockner, E. J., and S. Jørgensen (1992): New product advertising in dynamic oligopolies, *Zeitschrift för Operations Research* 36, 459–473

[56] Dockner, E. J., and N. V. Long (1993): International pollution control: cooperative versus noncooperative strategies, *Journal of Environmental Economics and Management* 24, 13–29

[57] Dockner, E. J., N. V. Long, and G. Sorger (1995): Multiperiod rent seeking investment: the winner-takes-all case, unpublished manuscript

[58] Dockner, E. J., N. V. Long, and G. Sorger (1996): Analysis of Nash equilibria in a class of capital accumulation games, *Journal of Economic Dynamics and Control* 20, 1209–1235

[59] Dockner, E. J., and R. Neck (1987): Zur Effizienz dezentralisierter Stabilisierungspolitik, in Isermann, H., et al. (eds.), *Operations Research Proceedings 1986*, Springer-Verlag, Berlin, 528–536

[60] Dockner, E. J., and R. Neck (1989): Time-consistency, subgame-perfectness, solution concepts and information patterns in dynamic models of stabilization policies, unpublished manuscript

[61] Dockner, E. J., and G. Sorger (1996): Existence and properties of equilibria for a dynamic game on productive assets, *Journal of Economic Theory* 71, 209–227

[62] Dockner, E. J., and H. Takahashi (1988): Further turnpike properties of capital accumulation games, *Economics Letters* 28, 321–325

[63] Dockner, E. J., and H. Takahashi (1990): On the saddle-point stability for a class of dynamic games, *Journal of Optimization Theory and Applications* 67, 247–258

[64] Dockner, E. J., and H. Takahashi (1993): Turnpike properties and comparative dynamics of general capital accumulation games, in Boldrin, M., et al. (eds.), *General Equilibrium, Growth and Trade II: The Legacy of L. McKenzie*, Academic Press, San Diego, 352–366

[65] Dolan, R. J., A. P. Jeuland, and E. Muller (1986): Models of new product diffusion: extension to competition against existing and potential firms over time, in Mahajan, V., and Y. Wind (eds.), *Innovation Diffusion Models of New Product Acceptance*, Ballinger, Cambridge, MA, 117–150

[66] Dowrick, S. (1986): Von Stackelberg and Cournot duopoly: choosing rules, *RAND Journal of Economics* 17, 251–260

[67] Driskill, R. A., and S. McCafferty (1989): Dynamic duopoly with adjustment costs: a differential game approach, *Journal of Economic Theory* 49, 324–338

[68] Dutta, P. K., and R. Sundaram (1993): How different can strategic models be?, *Journal of Economic Theory* 60, 42–61

[69] Ehtamo, H., and R. P. Hämäläinen (1993): A cooperative incentive equilibrium for a resource management problem, *Journal of Economic Dynamics and Control* 17, 659–678

[70] Ehtamo, H., J. Ruusunen, V. Kaitala, and R. P. Hämäläinen (1988): Solution for a dynamic bargaining problem with an application to resource management, *Journal of Optimization Theory and Applications* 59, 391–406

[71] Eliashberg, J., and A. P. Jeuland (1986): The impact of competitive entry in a developing market upon dynamic pricing strategies, *Marketing Science* 5, 20–36

[72] Eliashberg, J., and R. Steinberg (1987): Marketing-production decisions in an industrial channel of distribution, *Management Science* 33, 981–1000

[73] Eliashberg, J., and R. Steinberg (1991): Competitive strategies for two firms with asymmetric production cost structures, *Management Science* 37, 1452–1473

[74] Eliashberg, J., and R. Steinberg (1993): Marketing-production joint decision-making, in Eliashberg, J., and G. L. Lilien (eds.), *Handbooks of Operations Research and Management Science: Marketing*, North-Holland, Amsterdam, 827–880

[75] Engwerda, J. C. (1998): On the open-loop Nash equilibrium in LQ-games, *Journal of Economic Dynamics and Control* 22, 729–762

[76] Engwerda, J. C. (1998): On the scalar feedback Nash equilibria in the infinite horizon LQ-game, Eighth International Symposium on Dynamic Games and Applications, Maastricht

[77] Erickson, G. M. (1985): A model of advertising competition, *Journal of Marketing Research* 22, 297–304

[78] Erickson, G. M. (1991): *Dynamic Models of Advertising Competition*, Kluwer, Boston

[79] Erickson, G. M. (1992): Empirical analysis of closed-loop duopoly advertising strategies, *Management Science* 38, 1732–1749

[80] Erickson, G. M. (1995): Differential game models of advertising competition, *European Journal of Operational Research* 83, 431–438

[81] Feichtinger, G. (1982): Optimal policies for two firms in a noncooperative research project, in G. Feichtinger (ed.), *Optimal Control Theory and Economic Analysis*, North-Holland, Amsterdam, 373–397

[82] Feichtinger, G. (1983): A differential games solution to a model of competition between a thief and the police, *Management Science* 29, 686–699

[83] Feichtinger, G., and E. J. Dockner (1984): A note to Jørgensen's logarithmic advertising game, *Zeitschrift für Operations Research* 28, 133–153

[84] Feichtinger, G., and E. J. Dockner (1985): Optimal pricing in a duopoly: a noncooperative differential games solution, *Journal of Optimization Theory and Applications* 45, 199–218

[85] Feichtinger, G., and R. F. Hartl (1986): *Optimale Kontrolle Ökonomischer Prozesse*, de Gruyter, Berlin

[86] Feichtinger, G., R. F. Hartl, and S. P. Sethi (1994): Dynamic optimal control models in advertising: recent developments, *Management Science* 40, 195–226

[87] Fershtman, C. (1984): Goodwill and market shares in oligopoly, *Economica* 51, 271–281

[88] Fershtman, C. (1987): Identification of classes of differential games for which the open-loop is a degenerate feedback Nash equilibrium, *Journal of Optimization Theory and Applications* 55, 217–231

[89] Fershtman, C., and M. I. Kamien (1987): Dynamic duopolistic competition with sticky prices, *Econometrica* 55, 1151–1164

[90] Fershtman, C., and M. I. Kamien (1990): Turnpike properties in a finite-horizon differential game: dynamic duopoly with sticky prices, *International Economic Review* 31, 49–60

[91] Fershtman, C., V. Mahajan, and E. Muller (1990): Market share pioneering advantage: a theoretical approach, *Management Science* 36, 900–918

[92] Fershtman, C., and E. Muller (1984): Capital accumulation games of infinite duration, *Journal of Economic Theory* 33, 322–339

[93] Fershtman, C., and E. Muller (1986): Turnpike properties of capital accumulation games, *Journal of Economic Theory* 38, 167–177

[94] Fershtman, C., and S. Nitzan (1991): Dynamic voluntary provision of public goods, *European Economic Review* 35, 1057–1067

[95] Fisher, S. (1980): Dynamic inconsistency, cooperation and the benevolent dissembling government, *Journal of Economic Dynamics and Control* 2, 93–107

[96] Fleming, W. H., and R. W. Rishel (1975): *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York

[97] Fornell, C., W. T. Robinson, and B. Wernerfelt (1985): Consumption experience and sales promotion expenditure, *Management Science* 31, 1084–1105

[98] Friedman, A. (1971): *Differential Games*, Wiley, New York

[99] Friedman, J. W. (1986): Non-cooperative equilibria in time-dependent supergames, *Econometrica* 42, 221–237

[100] Friedman, J. W. (1991): *Game Theory with Applications to Economics*, Oxford University Press, New York

[101] Friedman, J. W. (ed.) (1994): *Problems of Coordination in Economic Activity*, Kluwer, Boston

[102] Fruchter, G., and S. Kalish (1997): Closed-loop advertising strategies in a duopoly, *Management Science* 43, 54–63

[103] Fudenberg, D., and J. Tirole (1983): Capital as a commitment: strategic investment to deter mobility, *Journal of Economic Theory* 31, 227–250

[104] Fudenberg, D., and J. Tirole (1986): *Dynamic Models of Oligopoly*, Harwood, London

[105] Fudenberg, D., and J. Tirole (1991): *Game Theory*, MIT Press, Cambridge, MA

[106] Gaimon, C. (1989): Dynamic game results of the acquisition of new technology, *Operations Research* 37, 410–425

[107] Gaimon, C. (1994): The price-production problem: an operations and marketing interface, Mimeo, School of Management, Georgia Institute of Technology, Atlanta

[108] Gardner, R. (1995): *Games for Business and Economics*, Wiley, New York

[109] Gaugusch, J. (1984): The noncooperative solution of a differential game: advertising versus pricing, *Optimal Control Applications and Methods* 5, 353–360

[110] Gihman, I., and A. V. Skorohod (1972): *Stochastic Differential Equations*, Springer-Verlag, New York

[111] Gilboa, I., and D. Schmeidler (1994): Infinite histories and steady orbits in repeated games, *Games and Economic Behavior* 6, 370–399

[112] Hamada, K. (1967): On the optimal transfer and income redistribution in a growing economy, *Review of Economic Studies* 34, 295–299

[113] Hämäläinen, R. P., A. Haurie, and V. Kaitala (1985): Equilibria and threats in a fishery management game, *Optimal Control Applications and Methods* 6, 315–333

[114] Hamilton, J. H., and S. M. Slutsky (1990): Endogenous timing in duopoly games: Stackelberg or Cournot equilibria, *Games and Economic Behavior* 2, 29–46

[115] Hanssens, D. M., L. J. Parsons, and R. L. Schultz (1990): *Market Response Models: Econometric and Time Series Analysis*, Kluwer, Boston

[116] Harris, C., and J. Vickers (1995): Innovation and natural resources: a dynamic game with uncertainty, *RAND Journal of Economics* 26, 418–430

[117] Harsanyi, J. C. (1967/68): Games with incomplete information played by 'Bayesian players' (Parts I, II, and III), *Management Science* 14, 159–182, 320–334, and 486–502

[118] Harsanyi, J. C., and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, MA

[119] Hartl, R. F. (1986): Arrow-type sufficient optimality conditions for nondifferentiable optimal control problems with state constraints, *Applied Mathematics and Optimization* 14, 229–247

[120] Hartl, R. F., Sethi, S. P., and R. Vickson (1995): A survey of the maximum principle for optimal control problems with state constraints, *SIAM Review* 37, 181–218

[121] Haurie, A. (1991): Piecewise deterministic and piecewise diffusion differential games with modal uncertainties, in Ricci, G. (ed.), *Decision Processes in Economics*, Springer, Berlin, 107–120

[122] Haurie, A. (1993): From repeated to differential games: how time and uncertainty pervade the theory of games, in Binmore, K., A. Kirman, and P. Tani (eds.), *Frontiers of Game Theory*, MIT Press, Cambridge, MA, 165–194

[123] Haurie, A., J. B. Krawczyk, and M. Roche (1993): Monitoring cooperative equilibria in a stochastic differential game, *Journal of Optimization Theory and Applications* 81, 73–95

[124] Haurie, A., and M. Pohjola (1987): Efficient equilibria in a differential game of capitalism, *Journal of Economic Dynamics and Control* 11, 65–78

[125] Haurie, A., and B. Tolwinski (1984): Acceptable equilibria in dynamic bargaining games, *Large Scale Systems* 6, 73–89

[126] Haurie, A., and B. Tolwinski (1985): Definition and properties of cooperative equilibria in a two-player game of infinite duration, *Journal of Optimization Theory and Applications* 46, 525–533

[127] Hoel, M. (1978): Distribution and growth as a differential game between workers and capitalists, *International Economic Review* 19, 335–350

[128] Horsky, D., and K. Mate (1988): Dynamic advertising strategies of competing durable goods producers, *Marketing Science* 7, 356–367

[129] Hughes Hallett, A. J. (1984): Non-cooperative strategies for dynamic policy games and the problem of time inconsistency, *Oxford Economic Papers* 36, 381–399

[130] Isaacs, R. (1965): *Differential Games*, Wiley, New York

[131] Jørgensen, S. (1982): A differential games solution to a logarithmic advertising model, *Journal of the Operational Research Society* 33, 425–432

[132] Jørgensen, S. (1982): A survey of some differential games in advertising, *Journal of Economic Dynamics and Control* 4, 341–369

[133] Jørgensen, S. (1986): Optimal dynamic pricing in an oligopolistic market: a survey, in Başar, T. (ed.), *Dynamic Games and Applications in Economics*, Springer-Verlag, Berlin, 179–237

[134] Jørgensen, S. (1986): Optimal production, purchasing and pricing: a differential game approach, *European Journal of Operational Research* 24, 64–76

[135] Jørgensen, S., and G. Sorger (1990): Feedback Nash equilibria in a problem of optimal fishery management, *Journal of Optimization Theory and Applications* 64, 293–310

[136] Jørgensen, S., and G. Zaccour (1999): Equilibrium pricing and advertising strategies in a marketing channel, *Journal of Optimization Theory and Applications* 102, 111–125

[137] Kaitala, V. (1986): Game theory models of fishery management: a survey, in Başar, T. (ed.), *Dynamic Games and Applications in Economics*, Springer-Verlag, New York, 252–266

[138] Kaitala, V. (1989): Non-uniqueness of no-memory feedback equilibria in a fishery resource game, *Autonomica* 25, 587–592

[139] Kaitala, V. (1993): Equilibria in a stochastic resource management game under imperfect information, *European Journal of Operational Research* 71, 439–453

[140] Kaitala, V., and G. Munro (1993): The economic management of high seas fishery resources: some game theoretic aspects, Working Paper No. 93-41, University of British Columbia, Vancouver

[141] Kaitala, V., and M. Pohjola (1988): Optimal recovery of a shared resource stock: a differential game model with efficient memory equilibria, *Natural Resource Modelling* 3, 91–119

[142] Kaitala, V., and M. Pohjola (1990): Economic development and agreeable redistribution in capitalism: efficient game equilibria in a two-class neoclassical growth model, *International Economic Review* 31, 421–438

[143] Kalish, S. (1988): Pricing new products from birth to decline: an expository review, in Devinney, T. M. (ed.), *Issues in Pricing*, Lexington Books, Lexington, 119–144

[144] Kamien, M. I., and N. L. Schwartz (1971): Limit pricing and uncertain entry, *Econometrica* 39, 441–454

[145] Kamien, M. I., and N. L. Schwartz (1971): Optimal maintenance and sale age for a machine subject to failure, *Management Science* 17, 427–449

[146] Kamien, M. I., and N. L. Schwartz (1972): Timing of innovations under uncertainty, *Econometrica* 40, 43–60

[147] Kamien, M. I., and N. L. Schwartz (1982): *Market Structure and Innovation*, Cambridge University Press, Cambridge

[148] Kamien, M. I., and N. L. Schwartz (1991): *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, 2nd edn., North-Holland, Amsterdam

[149] Kemp, M. C., and N. V. Long (1980): Optimal tariffs and exhaustible resources, in Kemp, M. C., and N. V. Long (eds.), *Exhaustible Resources, Optimality, and Trade*, North-Holland, Amsterdam, 187–196

[150] Kemp, M. C., and N. V. Long (1980): Resource extraction under conditions of common access in, Kemp, M. C., and N. V. Long (eds.), *Exhaustible Resources, Optimality, and Trade*, North-Holland, Amsterdam, 127–136

[151] Kemp, M. C., N. V. Long, and K. Shimomura (1993): Cyclical and noncyclical redistributive taxation, *International Economic Review* 34, 415–430

[152] Khalatbari, F. (1977): Market imperfections and the optimum rate of depletion of natural resources, *Economica* 44, 409–414

[153] Krasovskii, N. N., and A. I. Subbotin (1988): *Game-Theoretical Control Problems*, Springer-Verlag, Berlin

[154] Kreps, D. M. (1990): *Game Theory and Economic Modelling*, Oxford University Press, Oxford

[155] Kydland, F., and E. Prescott (1977): Rules rather than discretion: the inconsistency of optimal plans, *Journal of Political Economy* 85, 473–492

[156] Lancaster, K. (1973): The dynamic inefficiency of capitalism, *Journal of Political Economy* 81, 1092–1109

[157] Leitmann, G. (1974): *Cooperative and Non-cooperative Many Players Differential Games*, Springer-Verlag, New York

[158] Leitmann, G., and W. E. Schmitendorf (1978): Profit maximization through advertising: nonzero-sum differential game approach, *IEEE Transactions on Automatic Control* 23, 645–650

[159] Leitmann, G., and H. Stalford (1971): A sufficiency theorem for optimal control, *Journal of Optimization Theory and Applications* 8, 169–174

[160] Léonard, D., and N. V. Long (1992): *Optimal Control Theory and Static Optimization in Economics*, Cambridge University Press, Cambridge

[161] Levhari, D., and L. J. Mirman (1980): The great fish war: an example using a dynamic Cournot-Nash solution, *Bell Journal of Economics* 11, 322–334

[162] Lockwood, B. (1996): Uniqueness of Markov-perfect equilibrium in infinite-time affine quadratic differential games, *Journal of Economic Dynamics and Control* 20, 751–765

[163] Long, N. V. (1992): Pollution control: a differential game approach, *Annals of Operations Research* 37, 283–296

[164] Long, N. V., and K. Shimomura (1998): Some results on the Markov equilibria of a class of homogeneous differential games, *Journal of Economic Behavior and Organization* 33, 557–566

[165] Long, N. V., K. Shimomura, and H. Takahashi (1997): Comparing open-loop and Markov perfect equilibria in a class of differential games, Working Paper No. 97-22, CIRANO, Montreal

[166] Luce, R. D., and H. Raiffa (1957): *Games and Decisions*, John Wiley, New York

[167] Mahajan, V., E. Muller, and F. M. Bass (1993): New-product diffusion models, in Eliashberg, J., and G. L. Lilien (eds.), *Handbooks in Operations Research and Management Science: Marketing*, North-Holland, Amsterdam, 349–408

[168] Malliaris, A. G., and W. A. Brock (1982): *Stochastic Methods in Economics and Finance*, North-Holland, Amsterdam

[169] Maskin, E., and D. Newbery (1978): Rational expectations with market power: the paradox of the disadvantageous tariff on oil, Economic Research Paper No. 129, University of Warwick

[170] Maynard Smith, J. (1982): *Evolution and the Theory of Games*, Cambridge University Press, Cambridge

[171] McKenzie, L. W. (1986): Optimal economic growth, turnpike theorems and comparative dynamics, in Arrow, K. J., and M. D. Intriligator (eds.), *Handbook of Mathematical Economics III*, North-Holland, Amsterdam, 1281–1355

[172] McMillan, J. (1986): *Game Theory and International Economics*, Harwood, London

[173] McMillan, J. (1992): *Games, Strategies, and Managers*, Oxford University Press, Oxford

[174] McMillan, J., and H.-W. Sinn (1984): Oligopolistic extraction of a common property resource: dynamic equilibria, in Kemp, M. C., and N. V. Long (eds.), *Essays in the Economics of Exhaustible Resources*, North-Holland, Amsterdam, 199–214

[175] Mehlmann, A. (1985): *Differentialspiele – Die Analyse dynamischer Konfliktsituationen*, Hain Scientific, Königstein

[176] Mehlmann, A. (1988): *Applied Differential Games*, Plenum Press, New York

[177] Mehlmann, A. (1994): On differential games with Markov perfect triggering, *Journal of Optimization Theory and Applications* 80, 273–288

[178] Mesak, H. I., and A. F. Darrat (1993): A competitive advertising model: some theoretical and empirical results, *Journal of the Operational Research Society* 44, 491–502

[179] Miller, M., and M. Salmon (1985): Dynamic games and the time inconsistency of optimal policy in open economies, *Economic Journal* 95S, 124–137

[180] Mohr, E. (1988): Appropriation of common access natural resources through exploration: the relevance of the open-loop concept, *International Economic Review* 29, 307–319

[181] Moorthy, K. S. (1993): Competitive marketing strategies: game-theoretic models, in Eliashberg, J., and G. L. Lilien (eds.), *Handbooks in Operations Research and Management Science: Marketing*, North-Holland, Amsterdam, 143–190

[182] Munro, G. (1979): The optimal management of transboundary renewable resources, *Canadian Journal of Economics* 12, 355–376

[183] Munro, G. (1990): The optimal management of transboundary fisheries: game theoretic considerations, *Natural Resource Modelling* 4, 403–426

[184] Myerson, R. B. (1991): *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, MA

[185] Nash, J. (1950): Equilibrium points in N-person games, *Proceedings of the National Academy of Sciences of the USA* 36, 48–49

[186] Neck, R., and E. J. Dockner (1995): Commitment and coordination in a dynamic game model of international economic policy making, *Open Economies Review* 6, 5–28

[187] Nerlove, M., and K. J. Arrow (1962): Optimal advertising policy under dynamic conditions, *Economica* 29, 129–142

[188] Osborne, M. J., and A. Rubinstein (1994): *A Course in Game Theory*, MIT Press, Cambridge, MA

[189] Papavassilopoulos, G. P., J. V. Medanic, and J. B. Cruz (1979): On the existence of Nash strategies and solutions to coupled Riccati equations in linear-quadratic games, *Journal of Optimization Theory and Applications* 28, 49–76

[190] Petit, M. L. (1990): *Control Theory and Dynamic Games in Economic Policy Analysis*, Cambridge University Press, Cambridge

[191] Plourde, C., and D. Yeung (1989): Harvesting of a transboundary replenishable fish stock: a non-cooperative game solution, *Marine Resource Economics* 6, 57–71

[192] Pohjola, M. (1986): Applications of dynamic game theory to macroeconomics, in Başar, T. (ed.), *Dynamic Games and Applications in Economics*, Springer-Verlag, Berlin, 103–133

[193] Ramsey, F. P. (1928): A mathematical theory of saving, *Economic Journal* 38, 543–559

[194] Rao, R. C. (1988): Strategic pricing of durables under competition, in Devinney, T. M. (ed.), *Issues in Pricing*, Lexington Books, Lexington, 197–215

[195] Rao, R. C. (1990): The impact of competition on strategic marketing decisions, in Day, G., B. Weitz, and R. Wensley (eds.), *The Interface of Marketing and Strategy*, JAI Press, Greenwich, CT, 101–152

[196] Rasmusen, E. (1989): *Games and Information*, Basil Blackwell, New York

[197] Reinganum, J. F. (1981): Dynamic games of innovation, *Journal of Economic Theory* 25, 21–41

[198] Reinganum, J. F. (1982): A dynamic game of R&D: patent protection and competitive behavior, *Econometrica* 50, 671–688

[199] Reinganum, J. F. (1983): A class of differential games for which the open-loop and closed-loop equilibria coincide, *Journal of Optimization Theory and Applications* 36, 253–262

[200] Reinganum, J. F. (1984): Practical implications of game theoretic models of R&D, *American Economic Review* 74, 61–66

[201] Reinganum, J. F. (1989): The timing of innovation: research, development and diffusion, in Schmalensee, R., and R. D. Willig (eds.), *Handbook of Industrial Organization, I*, Elsevier, Amsterdam, 849–908

[202] Reinganum, J. F., and N. L. Stokey (1985): Oligopoly extraction of a common property resource: the importance of the period of commitment in dynamic games, *International Economic Review* 26, 161–173

[203] Reynolds, S. S. (1987): Capacity investment, preemption and commitment in an infinite horizon model, *International Economic Review* 28, 69–88

[204] Reynolds, S. S. (1991): Dynamic oligopoly with capacity adjustment costs, *Journal of Economic Dynamics and Control* 15, 491–514

[205] Rishel, R. (1975): Control of systems with jump Markov disturbances, *IEEE Transactions on Automatic Control* 20, 241–244

[206] Rishel, R. (1975): Dynamic programming and minimum principles for systems with jump Markov disturbances, *SIAM Journal of Control* 13, 338–371

[207] Rishel, R. (1990): Controlled continuous-time Markov processes, in Heyman, D. P., and M. J. Sobel (eds.), *Handbooks in Operations Research and Management Science*, II, North-Holland, Amsterdam, chapter 9

[208] Roos, C. F. (1925): A mathematical theory of competition, *American Journal of Mathematics* 46, 163–175

[209] Roos, C. F. (1927): A dynamic theory of economics, *Journal of Political Economy* 35, 632–656

[210] Salchenberger, L. M. (1989): Sole owner harvesting policies under threat of entries: a two-stage linear game, *Journal of Environmental Economics and Management* 16, 121–133

[211] Schelling, T. (1960): *The Strategy of Conflict*, Harvard University Press, Cambridge, MA

[212] Seierstad, A., and K. Sydsæter (1986): *Optimal Control Theory with Economic Applications*, North-Holland, Amsterdam

[213] Selten, R. (1965): Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit, *Zeitschrift für die Gesamte Staatswissenschaft* 121, 401–424 and 667–689

[214] Selten, R. (1975): Reexamination of the perfectness concept for equilibrium points in extensive games, *International Journal of Game Theory* 4, 25–55

[215] Selten, R. (1978): The chain store paradox, *Theory and Decision* 9, 127–159

[216] Sethi, S. P., and G. L. Thompson (1981): *Optimal Control Theory: Applications to Management Science*, M. Nijhoff, Boston

[217] Shimomura, K. (1991): The feedback equilibrium of a differential game of capitalism, *Journal of Economic Dynamics and Control* 15, 317–335

[218] Simaan, M., and T. Takayama (1978): Game theory applied to dynamic duopoly problems with production constraints, *Automatica* 14, 161–166

[219] Sinn, H.-W. (1984): Common property resources, storage facilities, and ownership structure: a Cournot model of the oil market, *Economica* 51, 235–252

[220] Sorger, G. (1989): Competitive dynamic advertising: a modification of the Case game, *Journal of Economic Dynamics and Control* 13, 55–80

[221] Sorger, G. (1997): Efficient income redistribution in a growing economy, *Central European Journal for Operations Research and Economics* 5, 41–50

[222] Sorger, G. (1998): Markov-perfect Nash equilibria in a class of resource games, *Economic Theory* 11, 79–100

[223] Spence, A. M. (1977): Entry, capacity, investment, and oligopolistic pricing, *Bell Journal of Economics* 8, 534–544

[224] Spence, A. M. (1979): Investment strategy and growth in a new market, *Bell Journal of Economics* 10, 1–19

[225] Stern, L. E. (1984): Criteria of optimality in the infinite-time optimal control problem, *Journal of Optimization Theory and Applications* 44, 497–508

[226] Sundaram, R. (1989): Perfect equilibrium in a class of symmetric dynamic games, *Journal of Economic Theory* 47, 153–177

[227] Takayama, A. (1985): *Mathematical Economics*, Dryden Press, Hinsdale

[228] Teng, J.-T., and G. L. Thompson (1983): Oligopoly models for optimal advertising when production costs obey a learning curve, *Management Science* 29, 1087–1101

[229] Thépot, J. (1983): Marketing and investment policies of duopolists in a growing industry, *Journal of Economic Dynamics and Control* 5, 387–404

[230] Thompson, G., and J.-T. Teng (1984): Optimal pricing and advertising policies for new product oligopoly models, *Marketing Science* 3, 148–168

[231] Tolwinski, B., A. Haurie, and G. Leitmann (1986): Cooperative equilibria in differential games, *Journal of Mathematical Analysis and Applications* 119, 182–202

[232] Tsutsui, S. (1996): Capacity constraints and voluntary output cutback in dynamic Cournot competition, *Journal of Economic Dynamics and Control* 20, 1683–1708

[233] Tsutsui, S., and K. Mino (1990): Nonlinear strategies in dynamic duopolistic competition with sticky prices, *Journal of Economic Theory* 52, 136–161

[234] Turnovsky, S. J., T. Başar, and V. d'Orey (1988): Dynamic strategic monetary policies and coordination in interdependent economies, *American Economic Review* 78, 1443–1445

[235] Ulph, A. (1989): A review of books on resource and environmental economics, *Bulletin of Economic Research* 41, 219–228

[236] Van Damme, E. (1991): *Stability and Perfection of Nash Equilibria*, Springer-Verlag, Berlin

[237] Van Hilten, O., P. Kort, and P. Van Loon (1993): *Dynamic Policies of the Firm: an Optimal Control Approach*, Springer-Verlag, Berlin

[238] Vermes, D. (1985): Optimal control of piecewise-deterministic Markov processes, *Stochastics* 14, 165–208

[239] Von Neumann, J., and O. Morgenstern (1944): *The Theory of Games and Economic Behavior*, Princeton University Press, Princeton

[240] Wernerfelt, B. (1984): Consumers with differing reaction speeds, scale advantages and industry structure, *European Economic Review* 24, 257–270

[241] Wernerfelt, B. (1985): The dynamics of prices and market shares over the product's life cycle, *Management Science* 31, 928–939

[242] Wernerfelt, B. (1986): A special case of dynamic pricing policy, *Management Science* 32, 1562–1566

[243] Wernerfelt, B. (1991): Brand loyalty and market equilibrium, *Marketing Science* 10, 229–245

[244] Wirl, F. (1996): Dynamic voluntary provision of public goods: extension for nonlinear strategies, *European Journal of Political Economy* 12, 555–560

[245] Xie, D. (1997): On time inconsistency: a technical issue in Stackelberg differential games, *Journal of Economic Theory* 76, 412–430

[246] Xie, J., and M. Sirbu (1995): Price competition and compatibility in the presence of positive demand externalities, *Management Science* 41, 909–926

Index