

A decorative border with intricate floral and scrollwork patterns in a dark brown color, framing the central text.

**Dynamic  
noncooperative  
game theory**

*Gaming - CCVI*

**Tamer Basar, Geert Jan  
Olsder**

P ; ^ Dynamic Noncooperative Game Theory 8^ Second Edition \*£ Tamer Bařar Geert Jan Olsder I In Applied Mathematics 23

p Dynamic Noncooperative Game Theory ^5 &\*. Second Edition X3k

SIAM's Classics in Applied Mathematics series consists of books that were previously allowed to go out of print. These books are republished by SIAM as a professional service because they continue to be important resources for mathematical scientists. Editor-in-Chief Robert E. O'Malley, Jr., University of Washington Editorial Board Richard A. Brualdi, University of Wisconsin-Madison Herbert B. Keller, California Institute of Technology Andrzej Z. Manitius, George Mason University Ingram Olkin, Stanford University Stanley Richardson, University of Edinburgh Ferdinand Verhulst, Mathematisch Instituut, University of Utrecht Classics in Applied Mathematics C. C. Lin and L. A. Segel, Mathematics Applied to Deterministic Problems in the Natural Sciences Johan G. F. Belinfante and Bernard Kolman, A Survey of Lie Groups and Lie Algebras with Applications and Computational Methods James M. Ortega, Numerical Analysis: A Second Course Anthony V. Fiacco and Garth P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques F. H. Clarke, Optimisation and Nonsmooth Analysis George F. Carrier and Carl E. Pearson, Ordinary Differential Equations Leo Breiman, Probability R. Bellman and G. M. Wing, An Introduction to Invariant Imbedding Abraham Berman and Robert J. Plemmons, Nonnegative Matrices in the Mathematical Sciences Olvi L. Mangasarian, Nonlinear Programming \*Carl Friedrich Gauss, Theory of the Combination of Observations Least Subject to Errors: Part One, Part Two, Supplement. Translated by G. W. Stewart Richard Bellman, Introduction to Matrix Analysis U. M. Ascher, R. M. M. Mattheij, and R. D. Russell, Numerical Solution of Boundary Value Problems for Ordinary Differential Equations K. E. Brenan, S. L. Campbell, and L. R. Petzold, Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations Charles L. Lawson and Richard J. Hanson, Solving Least Squares Problems J. E. Dennis, Jr. and Robert B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations Richard E. Barlow and Frank Proschan, Mathematical Theory of Reliability Cornelius Lanczos, Linear Differential Operators Richard Bellman, Introduction to Matrix Analysis, Second Edition Beresford N. Parlett, The

Symmetric Eigenvalue Problem \*First time in print. u

Classics in Applied Mathematics (continued) Richard Haberman, Mathematical Models: Mechanical Vibrations, Population Dynamics, and Traffic Flow Peter W. M. John, Statistical Design and Analysis of Experiments Tamer Başar and Geert Jan Olsder, Dynamic Noncooperative Game Theory, Second Edition Emanuel Parzen, Stochastic Processes Petar Kokotovic, Hassan K. Khalil, and John O'Reilly, Singular Perturbation Methods in Control: Analysis and Design Jean Dickinson Gibbons, Ingram Olkin, and Milton Sobel, Selecting and Ordering Populations: A Neiv Statistical Methodology James A. Murdock, Perturbations: Theory and Methods Ivar Ekeland and Roger Teman, Convex Analysis and Variational Problems Ivar Stakgold, Boundary Value Problems of Mathematical Physics, Volumes I and II J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables David Kinderlehrer and Guido Stampacchia, An Introduction to Variational Inequalities and Tkeir Applications F. Natterer, The Mathematics of Computerized Tomography Avinash C. Kak and Malcolm Slaney, Principles of Computerized Tomographic Imaging R. Wong, Asymptotic Approximations of Integrals O. Axelsson and V. A. Barker, Finite Element Solution of Boundary Value Problems: Theory and Computation David R. Brillinger, Time Series: Data Analysis and Theory Joel N. Franklin, Methods of Mathematical Economics: Linear and Nonlinear Programming, Fixed-Point Theorems Philip Hartman, Ordinary Differential Equations, Second Edition Michael D. Intriligator, Mathematical Optimization and Economic Theory Philippe G. Ciarlet, The Finite Element Method for Elliptic Problems Jane K. Cullum and Ralph A. Willoughby, Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Vol. I: Theory M. Vidyasagar, Nonlinear Systems Analysis, Second Edition Robert Mattheij and Jaap Molenaar, Ordinary Differential Equations in Theory and Practice Shanti S. Gupta and S. Panchapakesan, Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations Eugene L. Allgower and Kurt Georg, Introduction to Numerical Continuation Methods Heinz-Otto Kreiss and Jens Lorenz, Initiaf-Boundary Value Problems and the Nawer-Stofces Equations J. L. Hodges, Jr. and E. L. Lehmann, Basic Concepts of Probability and Statistics, Second Edition in

To Tangiil, Gozen, and Elif (T.B.) and Elke, Rena, Theda, and Kike (G.J.O.)

Dynamic Noncooperative Game Theory <sup>^</sup><sup>^</sup> Second Edition <sup>^-</sup><sup>^</sup> Tamer Basar  
University of Illinois at Urbana-Champaign Urbana, Illinois Geert Jan Olsder  
Delft University of Technology Delft, The Netherlands BiHJTL Society for  
Industrial and Applied Mathematics Philadelphia

Copyright ©1999 by the Society for Industrial and Applied Mathematics. This  
SIAM edition is an unabridged, revised republication of the work first  
published by Academic Press, New York, in 1982 and revised in 1995. 10  
9876543 All rights reserved. Printed in the United States of America. No part  
of this book may be reproduced, stored, or transmitted in any manner without  
the written permission of the publisher. For information, write to the Society  
for Industrial and Applied Mathematics, 3600 University City Science Center,  
Philadelphia, PA 19104-2688. Library of Congress Cataloging-in-Publication  
Data Basar, Tamer. Dynamic noncooperative game theory / Tamer Basar, Geert  
Jan Olsder. •• 2nd ed. p. cm. - (Classics in applied mathematics ; 23) Revised  
from the 2nd ed. published in 1995 by Academic Press, New York Includes  
bibliographical references and index. ISBN 0-89871-429-X (softcover) 1.  
Differential games. 2. Noncooperative games (Mathematics) I. Olsder, Geert  
Jan. II. Title. III. Series. QA272.B37 1998 519.3-dc21 98-46719 51HJTL is a  
registered trademark.

Contents Preface to the Classics Edition xi Preface to the Second Edition xiii 1  
Introduction and Motivation 1 1.1 Preliminary Remarks 1 1.2 Preview on  
Noncooperative Games 3 1.3 Outline of the Book 12 1.4 Conventions,  
Notation and Terminology 13 Part I 2 Noncooperative Finite Games: Two-  
Person Zero-Sum 17 2.1 Introduction 17 2.2 Matrix Games 18 2.3 Computation  
of Mixed Equilibrium Strategies 29 2.4 Extensive Forms: Single-Act Games  
36 2.5 Extensive Games: Multi-Act Games 45 2.6 Zero-Sum Games with  
Chance Moves 57 2.7 Two Extensions 60 2.7.1 Games with repeated decisions  
61 2.7.2 Extensive forms with cycles 63 2.8 Action-Dependent Information  
Sets 65 2.8.1 Duels 66 2.8.2 A searchlight game 66 2.9 Problems 70 2.10  
Notes 75 3 Noncooperative Finite Games: TV-Person Nonzero-Sum 77 3.1  
Introduction 77 3.2 Bimatrix Games 78 3.3 TV-Person Games in Normal Form  
88 3.4 Computation of Mixed-Strategy Nash Equilibria in Bimatrix Games 95  
3.5 Nash Equilibria of  $N$ -Person Games in Extensive Form 97 vu

VIII CONTENTS 3.5.1 Single-act games: Pure-strategy Nash equilibria 100

3.5.2 Single-act games: Nash equilibria in behavioral and mixed strategies	116
3.5.3 Multi-act games: Pure-strategy Nash equilibria	118
3.5.4 Multi-act games: Behavioral and mixed equilibrium strategies	126
3.5.5 Other refinements on Nash equilibria	128
3.6 The Stackelberg Equilibrium Solution	131
3.7 Nonzero-Sum Games with Chance Moves	148
3.8 Problems	153
3.9 Notes	159
4 Static Noncooperative Infinite Games	161
4.1 Introduction	161
4.2 e Equilibrium Solutions	162
4.3 Continuous-Kernel Games: Reaction Curves, and Existence and Uniqueness of Nash and Saddle-Point Equilibria	168
4.4 Stackelberg Solution of Continuous-Kernel Games	179
4.5 Consistent Conjectural Variations Equilibrium	186
4.6 Quadratic Games with Applications in Microeconomics	190
4.7 Braess Paradox	203
4.8 Problems	205
4.9 Notes	210
Part II 5 General Formulation of Infinite Dynamic Games	215
5.1 Introduction	215
5.2 Discrete-Time Infinite Dynamic Games	216
5.3 Continuous-Time Infinite Dynamic Games	224
5.4 Mixed and Behavioral Strategies in Infinite Dynamic Games . .	230
5.5 Tools for One-Person Optimization	233
5.5.1 Dynamic programming for discrete-time systems	233
5.5.2 Dynamic programming for continuous-time systems . . .	236
5.5.3 The minimum principle	241
5.6 Representations of Strategies Along Trajectories, and Time Consistency of Optimal Policies	247
5.7 Viscosity Solutions	255
5.8 Problems	260
5.9 Notes	262
6 Nash and Saddle-Point Equilibria of Infinite Dynamic Games	265
6.1 Introduction	265
6.2 Open-Loop and Feedback Nash and Saddle-Point Equilibria for Dynamic Games in Discrete Time	266
6.2.1 Open-loop Nash equilibria	267
6.2.2 Closed-loop no-memory and feedback Nash equilibria . . .	276
6.2.3 Linear-quadratic games with an infinite number of stages	288
6.3 Informational Properties of Nash Equilibria in Discrete-Time Dynamic Games	292
6.3.1 A three-person dynamic game illustrating informational nonuniqueness	292
6.3.2 General results on informationally nonunique equilibrium solutions	296
6.4 Stochastic Nonzero-Sum Games with Deterministic Information Patterns	303
6.5 Open-Loop and Feedback Nash and Saddle-Point Equilibria of Differential Games	310
6.5.1 Open-loop Nash equilibria	310
6.5.2 Closed-loop no-memory and feedback Nash equilibria . . .	320
6.5.3 Linear-quadratic differential games on an infinite time horizon	333
6.6 Applications in Robust Controller Designs: $H^\infty$ -Optimal Control	342
6.7 Stochastic Differential Games with Deterministic Information Patterns	350
6.8 Problems	355
6.9 Notes	361
7 Stackelberg Equilibria of Infinite Dynamic Games	365
7.1 Introduction	365
7.2 Open-Loop	

Stackelberg Solution of Two-Person Dynamic Games in Discrete Time 366 7.3  
 Feedback Stackelberg Solution Under CLPS Information Pattern 373 7.4  
 (Global) Stackelberg Solution Under CLPS Information Pattern 376 7.4.1 An  
 illustrative example (Example 7.1) 376 7.4.2 A second example (Example  
 7.2): Follower acts twice in the game 382 7.4.3 Linear Stackelberg solution of  
 linear-quadratic dynamic games 385 7.4.4 Incentives (deterministic) 392 7.5  
 Stochastic Dynamic Games with Deterministic Information Patterns 396 7.5.1  
 (Global) Stackelberg solution 396 7.5.2 Feedback Stackelberg solution 402  
 7.5.3 Stochastic incentive problems 403 7.6 Stackelberg Solution of  
 Differential Games 407 7.6.1 The open-loop information structure 407 7.6.2  
 The CLPS information pattern 412 7.7 Problems 418 7.8 Notes 421

X CONTENTS 8 Pursuit-Evasion Games 423 8.1 Introduction 423 8.2  
 Necessary and Sufficient Conditions for Saddle-Point Equilibria . 424 8.2.1  
 The Isaacs equation 425 8.2.2 Upper and lower values, and viscosity solutions  
 432 8.3 Capturability 434 8.4 Singular Surfaces 442 8.5 Solution of a Pursuit-  
 Evasion Game: The Lady in the Lake . . . 448 8.6 An Application in Maritime  
 Collision Avoidance 451 8.7 Role Determination and an Application in  
 Aeronautics 456 8.8 Problems 464 8.9 Notes 467 Appendix A Mathematical  
 Review 471 A.1 Sets 471 A.2 Normed Linear (Vector) Spaces 472 A.3  
 Matrices 473 A.4 Convex Sets and Functional 473 A.5 Optimization of  
 Functional 474 Appendix B Some Notions of Probability Theory 477 B.1  
 Ingredients of Probability Theory 477 B.2 Random Vectors 478 B.3 Integrals  
 and Expectation 480 B.4 Norms and the Cauchy-Schwarz Inequality 481  
 Appendix C Fixed Point Theorems 483 Bibliography 485 List of Corollaries,  
 Definitions, Examples, Lemmas, Propositions, Remarks and Theorems 507  
 Index 515

Preface to the Classics Edition This is the revised second edition of our 1982  
 book with the same title, which presents a rather comprehensive treatment of  
 static and dynamic noncooperative game theory, with emphasis placed (as in  
 the first and second editions) on the interplay between dynamic information  
 patterns and the structural properties of several different types of equilibria.  
 Whereas the second edition (1995) was a major revision with respect to the  
 original edition, this Classics edition only contains some moderate changes  
 with respect to the second one. There has been a number of reasons for the  
 preparation of this edition: • The second edition was sold out surprisingly fast.

• After some fifty years from its creation, the field of game theory is still very alive and active, as also reinforced by the selection of three game theorists (John Harsanyi, John Nash and Reinhard Selten) to share the 1994 Nobel prize in economics. Quite a few books on game theory have been published during the last ten years or so (though most of them essentially deal with static games only and are at the undergraduate level). • The recent interest in such fields as biological games, mathematical finance and robust control gives a new impetus to noncooperative game theory. • The topic of dynamic games has found its way into the curricula of many universities, sometimes as a natural supplement to a graduate level course on optimal control theory, which is actively taught in many engineering, applied mathematics and economics graduate programs. • At the level of coverage of this book, dynamic game theory is well established by now and has reached a level of maturity, which makes the book a timely addition to SIAM's prestigious Classics in Applied Mathematics series. For a brief description of and the level of the contents of the book, the reader is referred to the Preface of the second edition, which follows. It suffices to mention here the major changes made in this Classics edition with respect to the second edition. They are: XI

XII PREFACE TO THE CLASSICS EDITION • Inclusion of the "Braess paradox" in Chapter 4. • Inclusion of new material on the relationship between the existence of solutions to Riccati equations on the one hand and the existence of Nash equilibrium solutions to linear-quadratic differential games on the other, in Chapter 6. In the same chapter some new results have been included also on infinite-horizon differential games. We hope that this revised edition may find its way as its predecessors did. Tamer Başar Urbana, June 1998 Geert Jan Olsder Delft, June 1998

Preface to the Second Edition This is the second edition of our 1982 book with the same title, which presents an extensive and updated treatment of static and dynamic noncooperative game theory, with emphasis placed again (as in the first edition) on the interplay between dynamic information patterns and the structural properties of several different types of equilibria. There were essentially two reasons for producing this revised edition. One was the favorable reception of the first edition and the other one was the need to include new theoretical developments. Yet another reason was that the topic of dynamic games has found its way into the curricula of many universities. This

new edition contains some substantial changes and additions (and also a few deletions), but the flavor and the theme of the original text remain intact. The first part of the book, part I, which comprises Chapters 2 to 4, covers the material that is generally taught in an advanced undergraduate or first-year graduate course on noncooperative game theory. The coverage includes static finite and infinite games of both the zero-sum and nonzero-sum type and in the latter case both Nash and Stackelberg solution concepts are discussed. Furthermore, this part includes an extensive treatment of the class of dynamic games in which the strategy spaces are finite—the so-called multi-act games. Through an extensive tree formulation, the impact of information patterns on the existence, uniqueness and the nature of different types of noncooperative equilibria of multi-act games is thoroughly investigated. Most of the important concepts of static and dynamic game theory are introduced in these three chapters, and they are supplemented by several illustrative examples. Exposition of the material is quite novel, emphasizing concepts and techniques of multi-person decision making rather than mathematical details. However, mathematical proofs of most of the results are also provided, but without hindering the flow of main ideas. The major changes in this part over the first edition are the inclusion of additional material on: randomized strategies, finite games with repeated decisions and action-dependent information sets (Chapter 2); various refinements on the Nash equilibrium concept, such as trembling-hand, proper and perfect equilibria (Chapter 3); and in the context of static infinite games, stability of Nash equilibria and its relation with numerical schemes, consistent conjectural xm

XIV PREFACE TO THE SECOND EDITION variations equilibrium, and some new theorems on existence of Nash equilibria (Chapter 4). Some specific types of zero-sum games on the square (Chapter 4) have been left out. The second part of the book, Part II, which includes Chapters 5 to 8, extends the theory of the first part to infinite dynamic games in both discrete and continuous time (the so-known differential games). Here the emphasis is again on the close interrelation between information patterns and noncooperative equilibria of such multi-person dynamic decision problems. We present a unified treatment of the existing, but scattered, results in the literature on infinite dynamic games as well as some new results on the topic. The treatment is confined to deterministic games and to stochastic games under perfect state information, mainly because inclusion of a complete investigation on stochastic



dynamic games under imperfect state information would require presentation of some new techniques and thereby a volume much larger than the present one. Again, some of the major changes in this part over the first edition are the inclusion of new material on: time consistency (Chapters 5-7); viscosity solutions of the Hamilton-Jacobi-Bellman-Isaacs equation (Chapters 5 and 8); affine-quadratic dynamic games and results on infinite-horizon games in discrete and continuous time (Chapters 5 and 6); applications in robust ( $H^\infty$ ) controller designs (Chapter 8); incentive theory and relationship with Stackelberg solutions (Chapter 7); and Stackelberg equilibrium in the continuous time (Chapter 7). The material on the dolichobrachistochrone which was in Chapter 8 of the first edition, has been left out. Furthermore, the three appendices (A-C) are expanded versions of the earlier ones, and present the necessary background material on vector spaces, matrix algebra, optimization, probability and stochastic processes and fixed point theorems. Each chapter (with the exception of the first) is supplemented with a problem section. Each problem section contains standard exercises on the contents of the chapter, which a reader who has carefully followed the text should have no difficulty in solving, as well as exercises which can be solved by making use of the techniques developed in the text, but which require some elaborate thinking on the part of the reader. Following the problem section in each chapter (except the first) is a notes section, which is devoted to historical remarks and sources for further reading on the topics covered in that particular chapter. A one-semester course on noncooperative game theory, taught using this book, would involve mainly Part I and also an appropriate blend of some of the topics covered in Part II, this latter choice depending to a great extent on the taste of the instructor and the background of the students. Such a course would be suitable for advanced undergraduate or first-year graduate students in engineering, economics, mathematics, operations research and business administration. In order to follow the main flow of Chapters 2 and 3 the student need not have a strong mathematical background—apart from some elementary analysis—but he should be able to think in mathematical terms. However, proofs of some of the theorems in these two chapters, as well as the contents of Chapter 4, require some basic knowledge of real analysis and probability, which is summarized in the three appendices that are included towards the end of the

PREFACE TO THE SECOND EDITION xv book. Part II of the book, on the

other hand, is intended more for the researcher in the field, to provide him with the state of art in infinite dynamic game theory. However, selected topics from this part could also be used as a part of a graduate course on dynamic optimization, optimal control theory or mathematical economics. This edition has been prepared by means of  $\text{W}\text{T}\text{\$}$ i. Toward that end the text of the original edition was scanned; the formulas and figures were added separately. Despite the flexibility of  $\text{IATgX}$  and an intensive use of electronic mail, the preparation of this new edition has taken us much longer than anticipated, and we are grateful to the publisher for being flexible with respect to deadlines. The preparation would have taken even longer if our secretaries, Francie Bridges and Tatiana Tijanova, respectively, with their expertise of  $\text{JMgX}$ , had not been around. In the first stage Francie was responsible for the scanning and the retyping of the formulas, and Tatiana for the labeling and the figures, but in later stages they helped wherever necessary, and their assistance in this matter is gratefully acknowledged. The first author would also like to acknowledge his association with the Center for Advanced Study at the University of Illinois during the Fall 1993 semester, which provided him with released time (from teaching) which he spent partly in the writing of the new material in this second edition. The second author would like to thank INRIA Sophia-Antipolis in the person of its director, Pierre Bernhard, for allowing him to work on this revision while he spent a sabbatical there. We both would also like to take this opportunity to thank many of our colleagues and students for their input and suggestions for this second edition. Particular recognition goes to Niek Tholen, of Delft University of Technology, who read through most of the manuscript, caught many misprints and made penetrating comments and suggestions. We hope that this revised edition may find its way as its predecessor did. Tamer Başar Urbana, March 1994 Geert Jan Olsder Delft, March 1994

This page intentionally left blank

Chapter 1 Introduction and Motivation 1.1 Preliminary Remarks This book is concerned with dynamic noncooperative game theory. In a nutshell, game theory involves multi-person decision making; it is dynamic if the order in which the decisions are made is important, and it is noncooperative if each person involved pursues his or her own interests which are partly conflicting with others'. A considerable part of everything that has been written down, whether it is history, literature or a novel, has as its central theme a conflict

situation—a collision of interests. Even though the notion of "conflict" is as old as mankind, the scientific approach has started relatively recently, in the years around 1930, with, as a result, a still growing stream of scientific publications. We also see that more and more scientific disciplines devote time and attention to the analysis of conflicting situations. These disciplines include (applied) mathematics, economics, engineering, aeronautics, sociology, politics and mathematical finance. It is relatively easy to delineate the main ingredients of a conflict situation: an individual has to make a decision and each possible decision leads to a different outcome or result, which are valued differently by that individual. This individual may not be the only one who decides about a particular outcome; a series of decisions by several individuals may be necessary. If all these individuals value the possible outcomes differently, the germs for a conflict situation are there. The individuals involved, also called players or decision makers, or simply persons, do not always have complete control over the outcome. Sometimes there are uncertainties which influence the outcome in an unpredictable way. Under 1 Without any preference to sexes, a decision maker, in this book, is most times referred to as a "he". It could equally well be a "she". 1

T. BAŞAR AND G. J. OLSDER Table 1.1: The place of dynamic game theory.

One player	Many players	Static	Dynamic	Mathematical programming	Optimal control theory
		(Static) game theory	Dynamic (and/or differential) game theory		
such circumstances, the outcome is (partly) based on data not yet known and not determined by the other players' decisions. Sometimes it is said that such data is under the control of "nature", or "God", and that every outcome is caused by the joint or individual actions of human beings and nature. The established names of "game theory" (developed from approximately 1930) and "theory of differential games" (developed from approximately 1950, parallel to that of optimal control theory) are somewhat unfortunate. "Game theory", especially, appears to be directly related to parlor games; of course it is, but the notion that it is only related to such games is far too restrictive. The term "differential game" became a generally accepted name for games where differential equations play an important role. Nowadays the term "differential game" is also being used for other classes of games for which the more general term "dynamic game" would be more appropriate. The applications of "game theory" and the "theory of differential games" mainly deal with economic and political conflict situations, worst-case designs and also modeling of war					

games. However, it is not only the applications in these fields that are important; equally important is the development of suitable concepts to describe and understand conflict situations. It turns out, for instance, that the role of information—what one player knows relative to others—is very crucial in such problems. Scientifically, dynamic game theory can be viewed as a child of the parents game theory and optimal control theory.<sup>2</sup> Its character, however, is much more versatile than that of its parents, since it involves a dynamic decision process evolving in (discrete or continuous) time, with more than one decision maker, each with his own cost function and possibly having access to different information. This view is the starting point behind the formulation of "games in extensive form", which started in the 1930s through the pioneering work of Von Neumann, which culminated in his book with Morgenstern (Von Neumann and Morgenstern, 1947), and then made mathematically precise by Kuhn.<sup>2</sup> In almost all analogies there is a deficiency; a deficiency in the present analogy is that the child is as old as one of his parents—optimal control theory. For the relationship between these theories and the theory of mathematical programming, see Table 1.1.

INTRODUCTION AND MOTIVATION 3 (1953), all within the framework of "finite" games. The general idea in this formulation is that a game evolves according to a road or tree structure, where at every crossing or branching a decision has to be made as how to proceed. In spite of this original set-up, the evolution of game theory has followed a rather different path. Most research in this field has been, and is being, concentrated on the normal or strategic form of a game. In this form all possible sequences of decisions of each player are set out against each other. For a two-player game this results in a matrix structure. In such a formulation dynamic aspects of a game are completely suppressed, and this is the reason why game theory is classified as basically "static" in Table 1.1. In this framework emphasis has been more on (mathematical) existence questions, rather than on the development of algorithms to obtain solutions. Independently, control theory gradually evolved from Second World War servomechanisms, where questions of solution techniques and stability were studied. Then followed Bellman's "dynamic programming" (Bellman, 1957) and Pontryagin's "maximum principle" (Pontryagin et al., 1962), which spurred the interest in a new field called optimal control theory. Here the concern has been on obtaining optimal (i.e., minimizing or maximizing) solutions and developing numerical algorithms for

one-person single-objective dynamic decision problems. The merging of the two fields, game theory and optimal control theory, which leads to even more concepts and to actual computation schemes, has achieved a level of maturity, which the reader will hopefully agree with after he/she goes through this book. At this point, at the very beginning of the book, where many concepts have yet to be introduced, it is rather difficult to describe how dynamic game theory evolved in time and what the contributions of relevant references are. We therefore defer such a description until later, to the "notes" section of each chapter (except the present), where relevant historical remarks are included.

1.2 Preview on Noncooperative Games A clear distinction exists between two-player (or, equivalently, two-person) zero-sum games and the others. In a zero-sum game, as the name implies, the sum of the cost functions of the players is identically zero. Mathematically speaking, if  $u_i$  and  $U_i$  denote, respectively, the decision variable and the cost function of the  $i$ th player (to be written  $P_i$ ), then  $\sum_{i=1}^2 U_i(u_i) = 0$  in a zero-sum game. If this sum is, instead, equal to a nonzero constant (independent of the decision variables), then we talk about a "constant-sum" game which can, however, easily be transformed to a zero-sum game through a simple translation without altering the essential features of the game. Therefore, constant-sum games can be treated within the framework of zero-sum games, without any loss of generality, which we shall choose to do in this book. A salient feature of two-person zero-sum games that distinguishes them from other types of games is that they do not allow for any cooperation between the players, since, in a two-person zero-sum game, what one player gains incurs a

4 T. BAŞAR AND G. J. OLSDER Figure 1.1: The rope-pulling game. loss to the other player. However, in other games, such as two-player nonzero-sum games (wherein the quantity  $\sum_{i=1}^2 U_i(u_i, v_i)$  is not a constant) or three- or more-player games, the cooperation between two or more players may lead to their mutual advantage. Example 1.1 A point object (with mass one) can move in a plane which is endowed with the standard  $(x_1, x_2)$ -coordinate system. Initially, at  $t = 0$ , the point mass is at rest at the origin. Two unit forces act on the point mass; one is chosen by  $P_1$ , the other by  $P_2$ . The directions of these forces, measured counter-clockwise with respect to the positive  $x_j$ -axis, are determined by the players and are denoted by  $u_1$  and  $u_2$ , respectively; they may in general be time-varying. At time  $t = 1$ ,  $P_1$  wants to have the point mass as far in the negative  $x_1$ -direction as possible, i.e., he wants to minimize  $x_1(1)$ ,

whereas P2 wants it as far in the positive  $x_1$ -direction as possible, i.e., he wants to maximize  $x_1(1)$ , or equivalently, to minimize  $-x_1(1)$ . (See Fig. 1.1.) The "solution" to this zero-sum game follows immediately; each player pulls in his own favorite direction, and the point mass remains at the origin—such a solution is known as the saddle-point solution. We now alter the formulation of the game slightly, so that, in the present set-up, P2 wishes to move the point mass as far in the negative  $x_2$ -direction as possible, i.e., he wants to minimize  $x_2(1)$ . P1's goal is still to minimize  $x_1(1)$ . This new game is clearly nonzero-sum. The equations of motion for the point mass are  $\ddot{x} = \cos^2(u_1) + \cos(u_2)$ ,  $x_1(0) = x_1(1) = 0$ ;  $x_2 = \sin(u_1) + \sin(u_2)$ ,  $x_2(0) = x_2(1) = 0$ . Let us now consider the pair of decisions  $\{u_1 = \pi, u_2 = -\pi/2\}$  with the corresponding values of the cost functions being  $L_1 = X_1(1) = -1$  and  $L_2 = x_2(1) = -1$ . If P2 sticks to  $u_2 = -\pi/2$ , the best thing for P1 to do is to choose  $u_1 = \pi$ ; any other choice of  $w_x(t)$  will yield an outcome which is greater than  $-1$ . Analogously, if P1 sticks to  $u_1 = \pi$ , P2 does not have a better choice than  $u_2 = -\pi/2$ . Hence, the pair  $\{u_1 = \pi, u_2 = -\pi/2\}$  exhibits an equilibrium behavior, and this kind of a solution, where one player cannot improve his outcome by altering his decision unilaterally, is called a Nash equilibrium solution, or shortly, a Nash solution.

**INTRODUCTION AND MOTIVATION 5** If both players choose  $u_1 = u_2 = 57\pi/4$ , however, then the cost values become  $L_1 = L_2 = -1/2$ , which are obviously better, for both players, than the costs incurred under the Nash solution. But, in this case, the players have to cooperate. If, for instance, P2 would stick to his choice  $u_2 = 57\pi/4$ , then P1 can improve upon his outcome by playing  $u_1 = c$ , where  $c$  is a constant with  $\pi < c < 57\pi/4$ . P1 is better off, but P2 is worse off! Therefore, the pair of strategies  $\{u_1 = c, u_2 = 57\pi/4\}$  cannot be in equilibrium in a noncooperative mode of decision making, since it requires some kind of faith (or even negotiation), and thereby cooperation, on part of the players. If this is allowable, then the said pair of strategies—known as a Pareto-optimal solution—stands out as a reasonable equilibrium solution for the game problem (which is called a cooperative game), since it features the property that no other joint decision of the players can improve the performance of at least one of them, without degrading the performance of the other.  $\square$  In this book we shall deal only with noncooperative games. The reasons for such a seeming limitation are twofold. Firstly, cooperative games can, in general, be reduced to optimal control problems by determining a single cost function to be optimized by all players, which suppresses the

"game" aspects of the problem. Secondly, the size of this book would have increased considerably by inclusion of a complete discussion on cooperative games. Actions and strategies Heretofore, we have safely talked about "decisions" made by the players, without being very explicit about what a decision really is. This will be made more precise now in terms of information available to each player. In particular, we shall distinguish between actions (also called controls) on the one hand and strategies (or, equivalently, decision rules) on the other. If an individual has to decide about what to do the next day, and the options are fishing and going to work, then a strategy is: "if the weather report early tomorrow morning predicts dry weather, then I will go fishing, otherwise I will go to my office". This is a strategy or decision rule: what actually will be done depends on quantities not yet known and not controlled by the decision maker; the decision maker cannot influence the course of the events further, once he has fixed his strategy. Any consequence of such a strategy, after the unknown quantities are realized, is called an action. In a sense, a constant strategy (such as an irrevocable decision to go fishing without any restrictions or reservations) coincides with the notion of action. In the example above, the alternative actions are to go fishing and to go to work, and the action to be implemented depends on information (the weather report) which has to be known at the time it is carried out. In general, such information can be of different types. It can, for instance, comprise the previous 3

What we have in mind here is the class of cooperative games without side payments. If side payments are allowed, we enter the territory of (cooperative) games in characteristic function form, which is an altogether different topic. (See Owen (1968, 1982) and Vorob'ev (1977).)

6 T. BAŞAR AND G. J. OLSDER initially: P1 has 100 customers P2 has 150 customers  
 $y \setminus$  P1 decides /O\* /V™ P2 decides /n y\ /n y\ 80 75 145 125 number of customers of P1 125 190 120 175 number of customers of P2 80 75 115 95 profit P1 125 160 120 145 profit P2

Figure 1.2: One-period advertising game.

actions of all the other players. As an example, consider the following sequence of actions: if he is nice to me, I will be kind to him; if he is cool, I will be cool, etc. The information can also be of a stochastic nature, such as in the fishing example. Then, the actual decision (action) is based on data not yet known and not controlled by other players, but instead determined by "nature". The next example will now elucidate the role of information in a game situation and show how it affects the solution. The information here is

deterministic; "nature" does not play a role. Example 1.2 This example aims at a game-theoretic analysis of the advertising campaigns of two competing firms during two periods of time. Initially, each firm is allotted a certain number of customers; P1 has 100 and P2 has 150 customers. Per customer, each firm makes a fixed profit per period, say one dollar. Through advertisement, a firm can increase the number of its customers (some are stolen away from the competitor and others come from outside) and thereby its profit. However, advertising costs money; an advertising campaign for one period costs thirty dollars (for each firm). The figures in this example may not be very realistic; it is, however, only the ratio of the data that matters, not the scale with respect to which it is expressed. First, we consider the one-period version of this game and assume that the game ends after the first period is over. Suppose that P1 has to decide first whether he should advertise (Yes) or not (No), and subsequently P2 makes his choice. The four possible outcomes and the paths that lead to those outcomes are depicted in Fig. 1.2, in the form of a tree diagram. At every branching, a decision has to be made as how to proceed. The objective of each firm is to maximize its profit (for this one-period game). The "best" decisions, in this case, can be found almost by inspection; whatever P1 does (Y or N), P2 will always advertise (Y), since that is more profitable to him. P1, realizing this, has

INTRODUCTION AND MOTIVATION 7 80 75 115 95 } profits 125 160 120 145 Figure 1.3: P2 chooses independently of P1's action. to choose between Y (with profit 95) and N (with profit 75), and will therefore choose Y. Note that, at the point when P2 makes his decision, he knows P1's choice (action) and therefore P2's choice depends, in principle, on what P1 has done. P2's best strategy can be described as: "if P1 chooses N, then I choose Y; if P1 chooses Y, then I choose Y", which, in this case is a constant strategy. This one-period game will now be modified slightly so that P1 and P2 make their decisions simultaneously, or, equivalently, independently of each other. The information structure associated with this new game is depicted in Fig. 1.3; a dashed curve encircling the points D and D^ has been drawn, which indicates that P2 cannot distinguish between these two points. In other words, P2 has to arrive at a decision without knowing what P1 has actually done. Hence, in this game, the strategy of P2 has a different domain of definition, and it can easily be verified that the pair {V, Y} provides a Nash solution, in the sense described before in Example 1.1, yielding the same profits as in the previous game. We now extend



the game to two periods, with the objective of the firms being maximization of their respective cumulative profits (over both periods). The complete game is depicted in Fig. 1.4 without the information sets; we shall, in fact, consider three different information structures in the sequel. First, the order of the actions will be taken as  $P1 \text{ --- } P2 \text{ --- } P1 \text{ --- } P2$ , which means that at the time of his decision each player knows the actions previously taken. Under the second information structure to be considered, the order of the actions is  $(P1, P2)-(P1, P2)$ , which means that during each period the decisions are made independently of each other, but that for the decisions of the second period the (joint) actions of the first period are known. Finally, as a third case, it will be assumed that there is no order in the actions at all;  $P_i$  has to decide on what to do during both periods without any prior knowledge of what  $P_2$  will be doing, and vice versa. We shall be looking for Nash solutions in all three games. The  $P1-P2-P1-P2$  information structure The solution is obtained by working backward in time (a la "dynamic programming"). For the decision during the second period,  $P_2$  knows to which point ( $D_j \text{ --- } D_{14}$ ) the game has proceeded. From each of these points he chooses  $Y$

T. BAŞAR AND G. J. OLSDER  $P_2$  decides  $P_1$  decides  $D$ , A ■«---  $P_2$  decides  
64 IIM 64 104 60 133 60 103 92 100 62 100 76 121 46 91 60 143 60 143 57  
171 57 141 86 128 56 98 72 155 42 125 92 100 92 100 87 128 87 98 127 96  
97 96 109 116 79 86 76 121 76 121 71 155 71 125 105 116 75 116 90 i 140 /  
customers 60 \ profit 110 J 2nd period 144 140 142 126 135 132 131 117 207  
202 212 194 171 166 170 155 i total 229 228 225 216 303 301 258 285 220  
218 216 206 266 270 261 255 / profit Figure 1.4: The two-period advertising  
game, without the information sets. or  $N$ , depending on which decision leads to  
higher profit. These decisions are denoted by an asterisk in Fig. 1.4. At the  
time of his decision during the second period,  $P_1$  knows that the game has  
proceeded to one of the points  $D^{\wedge} \text{ --- } D_q$ . He can also guess precisely what  $P_2$   
will do after he himself has made his decision (the underlying assumption  
being that  $P_2$  will behave rationally). Hence,  $P_i$  can determine what will be  
best for him. At point  $D_3$ , for instance, decision  $N$  leads to a profit of 144 (for  
 $P_1$ ) and  $Y$  leads to 142; therefore,  $P_1$  will choose  $N$  if the game would be at  
 $D_3$ . The optimal decisions for  $P_1$  at the second period are indicated by an  
asterisk also. We continue this way (in retrograde time) until the vertex of the  
tree is reached; all best actions are designated by an asterisk in Fig. 1.4. As a  
result, the actual game will evolve along a path passing through the points  $D_2$ ,

D6, and D14, and the cumulative profits for P1 and P2 will be 170 and 261, respectively. The (P1,P2)-(P1,P2) information structure During the last period, both players know that the game has evolved to one of the points D3 — D8, upon which information the players will have to base their actions. This way, we have to solve four "one-period games", one for each of the points D3 — D8. The reader can easily verify that the optimal (Nash) solutions associated with these games are: information action profit (starting point) P1 P2 P1 P2 D3 N N 144 229 D4 N N 135 303 D5 Y N 212 216 D6 NY 166 270

INTRODUCTION AND MOTIVATION 9 The profit pairs corresponding to these solutions can be attached to the respective points D3 — Dq, after which a one-period game is left to be solved. The game tree is: 144 135 212 166 profit P1 229 303 216 270 profit P2 It can be verified that both players should play Y during the first period, since a unilateral deviation from this decision leads to a smaller profit for both players. The realized cumulative profits of P1 and P2 under this information scheme are therefore 166 and 270, respectively. The no-information case Both players have four possible choices: NN, NY, YN and YY, where the first (respectively, second) symbol refers to the first (respectively, second) period. Altogether there are  $4 \times 4 = 16$  possible pairs of realizable profits, which can be written in a matrix form: P2 \ chooses P1 \ chooses NN NY YN YY NN 144 229 142 225 207 220 212 216 NY 140 228 126 216 202 218 194 206 YN 135 303 131 258 171 266 170 261 YY 132 301 117 285 166 270 155 255 If P1 chooses YN and P2 chooses YY, then the profits are 166 and 270, respectively. If any one of the players deviates from this decision, he will be worse off. (If, for example, P1 deviates, his other profit options are 132, 117 and 155, each one being worse than 166.) Therefore  $\{(YN), (YY)\}$  is the Nash solution under the "no-information"-scheme.  $\square$

10 T. BAŞAR AND G. J. OLSDER Game problems of this kind will be studied extensively in this book, specifically in Chapter 3 which is devoted to finite nonzero-sum games, i.e., games in which each player has only a finite number of possible actions available to him. Otherwise, a game is said to be infinite, such as the one treated in Example 1.1. In Chapter 3, we shall also elucidate the reasons why the realized profits are, in general, dependent on the information structure. The information structure (P1, P2)-(P1, P2) in Example 1.2 is called a feedback information structure; during each period the players

know exactly to which point ("state") the game has evolved and that information is fed back into their strategies, which then leads to certain actions. This structure is sometimes referred to as closed-loop, though later on in this book a distinction between feedback and closed-loop will be made. The no-information case is referred to as an open-loop information structure. Not every game is well-defined with respect to every possible information structure; even though Example 1.2 was well-defined under both a feedback and an open-loop information structure, this is not always true. To exemplify this situation, we provide, in the sequel, two specific games. The game to be formulated in Example 1.3 does not make much sense under an open-loop information structure, and the one of Example 1.4 can only be played under the open-loop structure. Example 1.3 "The lady in the lake." A lady is swimming in a circular lake with a maximum speed  $v_l$ . A man who wishes to intercept the lady, and who has not mastered swimming, is on the side of the lake and can run along the shore with a maximum speed  $v_m$ . The lady does not want to stay in the lake forever and wants eventually to escape. If, at the moment she reaches the shore, the man cannot intercept her, she "wins" the game since on land she can run faster than the man. Note that open-loop solutions (i.e., solutions corresponding to the open-loop information structure) do not make sense here (at least not for the man). It is reasonable to assume that each player will react immediately to what his/her opponent has done and hence the feedback information structure, whereby the current position of man and lady are fed back, is more appropriate. Example 1.4 "The princess and the monster." This game is played in complete darkness. Just like the previous example, this one is also of the pursuit- evasion type. Here, however, the players do not know where their opponent is, unless they bump into each other (or come closer to each other than a given distance  $\epsilon$ ); that moment in fact defines the termination of the game. The velocities of both players are bounded and their positions are confined to a bounded area. The monster wants to bump into the princess as soon as possible, whereas the princess wants to push this moment as far into the future as possible. Since the players cannot react to each other's actions or positions, the information structure is open-loop. The solution to Example 1.3 will be given in Chapter 8; for more infor-

INTRODUCTION AND MOTIVATION 11 mation on Example 1.4, see Foreman (1977). This latter example also leads, rather naturally, to the concept of "mixed strategies"; the optimal strategies in the "princess and monster" game

cannot be deterministic, as were the strategies in Examples 1.1 and 1.2, for instance. For, if the monster would have a deterministic optimal strategy, then the princess would be able to calculate this strategy; if, in addition, she would know the monster's initial position, this would enable her to determine the monster's path and thereby to choose for herself an appropriate strategy so that she can avoid him forever. Therefore, an optimal strategy for the monster (if it exists) should dictate random actions, so that his trajectory cannot be predicted by the princess. Such a strategy is called mixed. Equilibria in mixed strategies will be discussed throughout this book. What is optimal? In contrast to optimal control problems (one-player games) where optimality has an unambiguous meaning, in multi-person decision making, optimality, in itself, is not a well-defined concept. Heretofore we have considered the Nash equilibrium solution, which is a specific form of "optimality". We also briefly discussed Pareto-optimality. There exist yet other kinds of optimality in nonzero-sum games. Two of them will be introduced now by means of the matrix game encountered in the "no-information"-case of Example 1.2. The equilibrium strategies in the "Nash" sense were YN for P1 and YY for P2. Suppose now that P2 is a careful and defensive player and wants to protect himself against any irrational behavior on the part of the other player. If P2 sticks to YY, the worst that can happen to him is that P1 chooses YY, in which case P2's profit becomes 255 instead of 270. Under such a defensive attitude P2 might play YN, since then his profit is at least 258. This strategy (or, equivalently, action in this case) provides P2 with a lower bound for his earnings and the corresponding solution concept is called minimax. The player who adopts this solution concept basically solves a zero-sum game, even though the original game might be nonzero-sum. Yet another solution concept is the one that involves a hierarchy in decision making: one of the players, say P1, declares and announces his strategy before the other player chooses his strategy and he (i.e., P1) is in a position to enforce this strategy. In the matrix game of the previous paragraph, if P1 says "I will play YY" and irrevocably sticks to it (by the rules of the game we assume that cheating is not possible), then the best P2 can do is to choose YN, in which case P1's profit becomes 170 instead of 166 which was obtainable under the Nash solution concept. Such games in which one player (called the leader) declares his strategy first and enforces it on the other player (called the follower) are referred to as Stackelberg games. They will be discussed in Chapter 3 for finite action sets, and in Chapters 4 and 7 for infinite action sets. Static versus dynamic As the last topic of this section,

we shall attempt to answer the question: "When is a game called dynamic and when is it static?" So far, we have talked rather

12 T. BAŞAR AND G. J. OLSDER loosely about these terms; there is, in fact, no uniformly accepted separating line between static games on the one hand and dynamic games on the other. We shall choose to call a game dynamic if at least one player is allowed to use a strategy that depends on previous actions. Thus, the games treated in Example 1.2 with the information schemes  $PI - P2 - PI - P2$  and  $(P1,P2)-(P1,P2)$  are dynamic. The third game of Example 1.2, with the "no-information" scheme, should be called static, but, by an abuse of language, such a game is often also called dynamic. The reason is that the players act more than once and thus time plays a role. A game in which the players act only once and independently of each other is definitely called a static game. The game displayed in Fig. 1.3 is clearly static.

### 1.3 Outline of the Book

The book comprises eight chapters and three appendices. The present (first) chapter serves the purpose of introducing the reader to the contents of the book, and to the conventions and terminology adopted. The next three chapters constitute Part I of the text, and deal with finite static and dynamic games and infinite static games. Chapter 2 discusses the existence, derivation and properties of saddle-point equilibria in pure, mixed and behavioral strategies for two-person zero-sum finite games. It is in this chapter that the notions of normal and extensive forms of a dynamic game are introduced, and the differences between actions and strategies are delineated. Also treated is the class of two-person zero-sum finite games in which a third (chance) player with a fixed mixed strategy affects the outcome. Chapter 3 extends the results of Chapter 2 to nonzero-sum finite games under basically two different types of equilibrium solution concepts, viz. the Nash solution and the Stackelberg (hierarchical) solution. The impact of dynamic information on the structure of these equilibria is thoroughly investigated, and in this context the notion of prior and delay commitment modes of play are elucidated. Chapter 4 deals with static infinite games of both the zero-sum and nonzero-sum type. In this context it discusses the existence, uniqueness and derivation of (pure and mixed) saddle-point, Nash and Stackelberg equilibria, as well as the consistent conjectural variations equilibrium. It provides explicit solutions for some types of games, with applications in microeconomics. The remaining four chapters constitute Part II of the book, for which Chapter 5 provides a general introduction. It introduces the class of infinite dynamic games to be studied in

the remaining chapters, and also gives some background material on optimal control theory. Furthermore, it makes the notions of "representations of strategies on given trajectories" and "time consistency" precise. The major portion of Chapter 6 deals with the derivation and properties of Nash equilibria with prescribed fixed duration under different types of deterministic information patterns, in both discrete and continuous time. It also presents as a special case saddle-point equilibria in such dynamic games, with important

INTRODUCTION AND MOTIVATION 13 applications in worst-case controller designs (such as  $H^\infty$ -optimal control). Chapter 7 discusses the derivation of global and feedback Stackelberg equilibria for the class of dynamic games treated in Chapter 6, and also the relationship with the theory of incentives. Finally, Chapter 8 deals with the class of zero-sum differential games for which the duration is not fixed a priori —the so-called pursuit-evasion games— and under the feedback information pattern. It first presents some necessary and sufficient conditions for the saddle-point solution of such differential games, and then applies these to pursuit-evasion games with specific structures so as to obtain some explicit results. Each chapter (with the exception of Chapter 1) starts with an introduction section which summarizes its contents, and therefore we have kept the descriptions above rather brief. Following Chapter 8 are the three appendices, two of which (Appendices A and B) provide some background material on sets, vector spaces, matrices, optimization theory and probability theory to the extent to which these notions are utilized in the book. The third appendix, on the other hand, presents some theorems which are used in Chapters 3 and 4. The book ends with a list of references, a table that indicates the page numbers of the Corollaries, Definitions, Examples, Lemmas, Propositions, Remarks and Theorems appearing in the text and an index.

#### 1.4 Conventions, Notation and Terminology

Each chapter of the book is divided into sections, and occasionally sections are divided into subsections. Section 2.1 refers to Section 1 of Chapter 2, and Section 8.5.2 refers to subsection 2 of Section 8.5. The following items appear in this book and they are numbered per chapter, such as Prop. 3.7 referring to the seventh proposition of the third chapter: Theorem (abbreviated to Thm.) Corollary Lemma Definition (abbreviated to Def.) Problem Equation Figure (abbreviated to Fig.) Property Example Proposition (abbreviated to Prop.) Remark Unlike the numbering of other items, equation numbers appear within parentheses, such as equation (2.5) which refers to the fifth numbered equation

in Chapter 2. References to bibliographical sources (listed alphabetically at the end of the book) are made according to the Harvard system (i.e., by name(s) and date). The following abbreviations and symbols are adopted in the book, unless stated otherwise in specific contexts: RHS right-hand side LHS left-hand side w.p. a with probability a LP linear programming = equality sign

14 T. BAŞAR AND G. J. OLSDER = defined by □ end of proof, remark, example, etc. || parallel to  $r + 1$ , if  $x > 0$   $\text{sgn}(x)$   $dA$   $P_i$   $P(E)$   $U$   $J$   $1$   $N$   $N$   $ul$   $7^*$   $rf$   $K$   $K$   $M$   $X$   $x(t)$   $V$   $V$   $V$   $v_m$   $R$   $R_n$   $|x|$   $IMI$   $=<$  undefined, if  $x = 0$   $I^{-1}$ , if  $x < 0$  boundary of the set  $A$  player  $i$  pursuer (evader) cost function(al) of  $P_i$  for a game in extensive form cost function(al) of  $P_i$  for a game in normal form number of players =  $\{1, \dots, N\}$  players' set;  $i \in N$  action (decision, control) variable of  $P_i$ ;  $ul \in Ul$  strategy (decision law) of  $P_i$ ;  $7^* \in Tl$  information available to  $P_i$  number of stages (levels) in a discrete-time game  $\pm \{1, \dots, K\}$ ;  $keK$  time interval on which a differential game is defined:  $te[0, T]$  state variable;  $x(t)$  in continuous time and  $X_{fc}$  in discrete time time-derivative of  $x(t)$  value of zero-sum game (in pure strategies) upper value lower value value of a zero-sum game in mixed strategies real line  $n$ -dimensional Euclidean space Euclidean norm of a finite-dimensional vector  $x$ , i.e., norm of a vector  $x$  in an infinite-dimensional space.  $\{x'xY'2$  Other symbols or abbreviations used with regard to sets, unions, summation, matrices, optimization, random variables, etc., are introduced in Appendices A and B. A convention that we have adopted throughout the text (unless stated otherwise) is that in nonzero-sum games all players are cost-minimizers, and in two-person zero-sum games  $P_1$  is the minimizer and  $P_2$  is the maximizer. In two-person matrix games  $P_1$  is the row-selector and  $P_2$  is the column-selector. The word "optimality" is used in the text rather freely for different equilibrium solutions when there is no ambiguity in the context; optimum quantities (such as strategies, controls, etc.) are generally identified by an asterisk (\*).

## Part I

This page intentionally left blank

### Chapter 2 Noncooperative Finite Games: Two-Person Zero-Sum 2.1

Introduction This chapter deals with the class of two-person zero-sum games in which the players have a finite number of alternatives to choose from. There

exist two different formulations for such games: the normal (matrix) form and the extensive (tree) form. The former constitutes a suitable representation of a zero-sum game when each player's information is static in nature, since it suppresses all the dynamic aspects of the decision problem. The extensive form, on the other hand, displays explicitly the evolution of the game and the existing information exchanges between the players. In the first part of the chapter (Sections 2.2 and 2.3), the normal form is introduced together with several related concepts, and then existence and computation of saddle-point equilibria are discussed for both pure and mixed strategies. In the second part of the chapter (Sections 2.4 and 2.5), extensive form description for zero-sum finite games without chance moves is introduced, and saddle-point equilibria for such games are discussed, also within the class of behavioral strategies. This discussion is first confined to single-act games in which each player is allowed to act only once, and then it is extended to multi-act games. The third part of the chapter (comprising the remaining three sections) discusses extensions in a number of directions. Section 2.6 is devoted to a brief discussion on the nature of equilibria for zero-sum games which also incorporate chance moves. Section 2.7 deals with two extensions: games with nonstandard information patterns and randomized strategies, and games in extensive 17

18 T. BAŞAR AND G. J. OLSDER form which exhibits cycles. Finally, Section 2.8 deals with games with action-dependent information sets, where the information on which the players can base future moves depends on previous moves.

2.2 Matrix Games The most elementary type of two-person zero-sum games are matrix games. There are two players, to be referred to as player 1 (P1) and player 2 (P2), and an  $(m \times n)$ -dimensional (real-valued) matrix  $A = \{a^i_j\}$ . Each entry of this matrix is an outcome of the game, corresponding to a particular pair of decisions made by the players. For P1, the alternatives are the  $m$  rows of the matrix, while for P2 the possible choices are the  $n$  columns of the same matrix. These alternatives are known as the strategies of the players. If P1 chooses the  $i$ th row and P2 the  $j$ th column, then  $a^i_j$  is the outcome of the game, and P1 pays this amount to P2. In case  $a^i_j$  is negative, this should be interpreted as P2 paying P1 the positive amount corresponding to this entry (i.e.,  $-a^i_j$ ). To regard the entries of the matrix as sums of money to be paid by one player to the other is, of course, only a convention. In a more general framework, these outcomes represent utility



transfers from one player to the other. Thus, we can view each element of the matrix  $A$  (i.e., a possible outcome of the game) as the net change in the utility of P2 for a particular play of the game, which is equal to minus the net change in the utility of P1. Then, regarded as a rational decision maker, P1 will seek to minimize the outcome of the game, while P2 will seek to maximize it, by independent decisions. Assuming that this game is to be played only once, then a reasonable mode of play for P1 is to secure his losses against any (rational or irrational) behavior of P2. Under such an incentive, P1 is forced to pick that row ( $i^*$ ) of matrix  $A$ , whose largest entry is no bigger than the largest entry of any other row. Hence, if P1 adopts the  $i^*$ th row as his strategy, where  $i^*$  satisfies the inequalities  $V_i(A) = \max_j a_{ij} < \max_{i'} V_{i'}(A)$ ,  $i = 1, \dots, m$ , (2.1) then his losses will be no greater than  $V$  which we call the loss ceiling of P1, or equivalently, the security level for his losses. The strategy "row  $i^*$ " that yields this security level will be called a security strategy for P1. Adopting a similar mode of play, P2 will want to secure his gains against any behavior of P1, and consequently, he will choose that column ( $j^*$ ) whose smallest entry is no smaller than the smallest entry of any other column. In mathematical terms,  $j^*$  will be determined from the  $n$ -tuple of inequalities  $Y_j(A) = \min_i a_{ij} > \min_{j'} Y_{j'}(A)$ ,  $j = 1, \dots, n$ . (2.2) By deciding on the  $j^*$ th column as his strategy, P2 secures his gains at the level  $V$  which we call the gain-floor of P2, or equivalently, the security level for his gains. Furthermore, any strategy for P2 that secures his gain-floor, such as "column  $j^*$ ", is called a security strategy for P2.

TWO-PERSON ZERO-SUM FINITE GAMES 19 Theorem 2.1 In every matrix game  $A = \{a^i_j\}$ , (i) the security level of each player is unique, (ii) there exists at least one security strategy for each player, (iii) the security level of P1 (the minimizer) never falls below the security level of P2 (the maximizer), i.e.,  $\min_j Y_j(A) = V_-(A) < V(A) = \max_i V_i(A)$ , (2.3) where  $i^*$  and  $j^*$  denote security strategies for P1 and P2, respectively. Proof The first two parts follow readily from (2.1) and (2.2), since in every matrix game there is only a finite number of alternatives for each player to choose from. To prove part (iii), we first note the obvious set of inequalities  $\min_i a_{ik} < a_{i^*k} < \max_j a_{ij}$  which holds true for all possible  $k$  and  $i$ . Now, letting  $k = j^*$  and  $i = i^*$ , the desired result, i.e., (2.3), follows. D The third part of Thm. 2.1 says that, if played by rational players, the outcome of the game will always lie between  $V$  and  $V_-$ . It is for this reason that the security levels  $V$  and  $V_-$  are &^so called, respectively, the upper value and lower value of the matrix game. Yet another interpretation that

could be given to these values is the following. Consider the matrix game in which players do not make their decisions independently, but there is a predetermined ordering according to which the players act. First PI decides on what row he will choose, he transmits this information to P2 who subsequently chooses a column. In this game, P2 definitely has an advantage over PI, since he knows what his opponent's actual choice of strategy is. Then, it is unquestionable that the best play for PI is to choose one of his security strategies (say, row  $i^*$ ). P2's "optimal" response to this is the choice of a "column  $j^\circ$ ", where  $j^\circ$  is determined from  $d_i^* j^\circ = \max_{j \in J} a_{ij} = V = \min_{i \in I} \max_{j \in J} a_{ij}$ , (2.4) with the "min max" operation designating the order of play in this decision process. Thus, we see that the upper value of the matrix game is actually attained in this case. Now, if the roles of the players are interchanged, this time PI having a definite advantage over P2, then P2's best choice will again be one of his security strategies (say, column  $j^*$ ), and Pi's response to this will be a "row  $i^\circ$ ", where  $i^\circ$  satisfies  $a_{i^\circ j^*} = \min_{i \in I} \max_{j \in J} a_{ij} = V_-$  (2.5) with the "max min" symbol implying that the minimizer acts after the maximizer. Relation (2.5) indicates that the outcome of the game is equal to the lower value  $V_-$ , when PI observes P2's choice.

20 T. BAŞAR AND G. J. OLSDER To illustrate these facets of matrix games, let us now consider the following (3 x 4) matrix. P2 1 0 -2 3 -1 2 3 2 0 -2 1 1 Here P2 (the maximizer) has a unique security strategy, "column 3" (i.e.,  $j^* = 3$ ), securing him a gain-floor of  $V_- = 0$ . PI (the minimizer), on the other hand, has two security strategies, "row 2" and "row 3" (i.e.,  $i^* = 2, i^* = 3$ ), yielding him a loss-ceiling of  $V = \max_{j \in J} \min_{i \in I} a_{ij} = \max_{j \in J} a_{3j} = 2$  which is above the security level of P2. Now, if P2 plays first, then he chooses his security strategy "column 3", with Pi's unique response being "row 3" ( $i^\circ = 3$ ), resulting in an outcome of  $0 = V_-$ . If PI plays first, he is actually indifferent between his two security strategies. In case he chooses "row 2", then P2's unique response is "column 3" ( $j^\circ = 3$ ),<sup>4</sup> whereas if he chooses "row 3", his opponent's response is "column 2" ( $j^\circ = 2$ ), both pairs of strategies yielding an outcome of  $2 = V$ . The preceding discussion validates the argument that, when there is a definite order of play, security strategies of the player who acts first make complete sense, and actually they can be considered to be in "equilibrium" with the corresponding response strategies of the other player. By two strategies (of PI and P2) to be in equilibrium, we roughly mean here that, after the game is over and its outcome is observed, the players should

have no ground to regret their past actions. In a matrix game with a fixed order of play, for example, there is no justifiable reason for the player who acts first to regret his security strategy after the game is over. But what about the class of matrix games in which the players arrive at their decisions independently? Do security strategies have any sort of an equilibrium property, in that case? To shed some light on this question, let us consider the  $(3 \times 3)$  matrix game of (2.7) below, with the players acting independently, and the game to be played only once.

P2	4	0	1	0
-1	2	-1	3	1

Both players have unique security strategies, "row 3" for P1 and "column 1" for P2, with the upper and lower values of the game being  $V = 2$  and  $V_- = 0$ , respectively. If both players play their security strategies, then the outcome of the game is 1, which is midway between the security levels of the players. But after the game is over, P1 might think: "Well I knew that P2 was going to play his security strategy, it is a pity that I didn't choose row 2 and enjoy an outcome of 0". Thinking along the same lines, P2 might also regret that he did not play column 2 and achieve an outcome of 2. This, then, indicates that, in this matrix game, the security strategies of the players cannot possibly be optimal response strategies.

TWO-PERSON ZERO-SUM FINITE GAMES 21 possess any equilibrium property. On the other hand, if a player chooses a row or column (whichever the case is) different from the one dictated by his security strategy, then he will be taking chances, since there is always a possibility that the outcome of the game might be worse for him than his security level. For the class of matrix games with equal upper and lower values, however, such a dilemma does not arise. Consider, for example, the  $(2 \times 2)$  matrix game P2 3 -1 1 1 in which "row 2" and "column 2" are the unique security strategies of P1 and P2, respectively, resulting in the same security level  $V = V_- = 1$ . These security strategies are in equilibrium, since each one is "optimal" against the other one. Furthermore, since the security levels of the players coincide, it does not make any difference, as far as the outcome of the game is concerned, whether the players arrive at their decisions independently or in a predetermined order. The strategy pair {row 2, column 2}, possessing all these favorable features, is clearly the only candidate that could be considered as the equilibrium solution of the matrix game (2.8). Such equilibrium strategies are known as saddle-point strategies, and the matrix game, in question, is then said to have a saddle point in pure strategies. A more precise definition for these terms is

given below. Definition 2.1 For a given  $(m \times n)$  matrix game  $A = \{a_{ij}\}$ , let  $\{i^*, j^*\}$  be a pair of strategies adopted by the players. Then, if the pair of inequalities  $a_{i^*j} < a_{i^*j^*} < a_{ij^*}$  (2.9) is satisfied for all  $i = 1, \dots, m$  and all  $j = 1, \dots, n$ , the strategies  $\{i^*, j^*\}$  are said to constitute a saddle-point equilibrium (or simply, they are said to be saddle-point strategies), and the matrix game is said to have a saddle point in pure strategies. The corresponding outcome  $a_{i^*j^*}$  of the game is called the saddle-point value, or simply the value, of the matrix game, and is denoted by  $V(A)$ . Theorem 2.2 Let  $A = \{a_{ij}\}$  denote an  $(m \times n)$  matrix game with  $V(A) = V(A)$ . Then, (i)  $A$  has a saddle point in pure strategies, (ii) an ordered pair of strategies provides a saddle point for  $A$  if, and only if, the first of these is a security strategy for  $P_1$  and the second one is a security strategy for  $P_2$ , (iii)  $V(A)$  is uniquely given by  $V(A) = V(A) = V_*(A)$ .

22 T. BAŞAR AND G. J. OLSDER Proof. Let  $i^*$  denote a security strategy for  $P_1$  and  $j^*$  denote a security strategy for  $P_2$ , which always exist by Thm. 2.1 (ii). Now, since  $V = V_*$ , we have  $a_{i^*j} < a_{i^*j^*} < a_{ij^*}$  for all  $i = 1, \dots, m$  and  $j = 1, \dots, n$ ; and letting  $i = i^*, j = j^*$  in this inequality, we obtain  $V = V_* = a_{i^*j^*}$ . Using this result in the same inequality yields which is (2.9). This completes the proof of (i) and the sufficiency part of (ii). We now show that the class of saddle-point strategy pairs is no larger than the class of ordered security strategy pairs. Let  $\{i^*, j^*\}$  be a saddle-point strategy pair. Then it follows from (2.9) that  $a_{i^*j} < a_{i^*j^*} < a_{ij^*}$  for all  $i = 1, \dots, m, j = 1, \dots, n$  thus proving that " $i^*$ " is a security strategy for  $P_1$ . Analogously, it can be shown that " $j^*$ " is a security strategy for  $P_2$ . This completes the proof of (ii). Part (iii) then follows readily from (ii). D We now immediately have the following property of saddle-point solutions, which follows from Thm. 2.2 (ii). Corollary 2.1 In a matrix game  $A$ , let  $\{i^1, j^1\}$  and  $\{i^2, j^2\}$  be two saddle-point strategy pairs. Then  $\{i^1, j^2\}, \{i^2, j^1\}$  are also in saddle-point equilibrium. This feature of the saddle-point strategies, given above in Corollary 2.1, is known as their ordered interchangeability property. Hence, in the case of nonunique (multiple) saddle points, each player does not have to know (or guess) the particular saddle-point strategy his opponent will use in the game, since all such strategies are in equilibrium and they yield the same value—indeed a very desirable property that an equilibrium solution is expected to possess. For the case when the security levels of the players do not coincide, however, no such

equilibrium solution can be found within the class of (pure) strategies that we have considered so far. One way of resolving this predicament is, as we have discussed earlier, to assume that one player acts after observing the decision of the other one, in which case the security level of the player who acts first is attained by an equilibrium "strategy pair". Here, of course, the strategy of the player who acts last explicitly depends on the action of the first acting player and hence the game has a "dynamic character". The precise meaning of this will be made clear in Section 2.4, where more details can be found on these aspects of zero-sum games.

TWO-PERSON ZERO-SUM FINITE GAMES 23 Mixed strategies Yet another approach to obtain an equilibrium solution in matrix games that do not possess a saddle point, and in which players act independently, is to enlarge the strategy spaces, so as to allow the players to base their decisions on the outcome of random events—thus leading to the so-called mixed strategies. This is an especially convincing approach when the same matrix game is played over and over again, and the final outcome, sought to be minimized by P1 and maximized by P2, is determined by averaging the outcomes of individual plays. To introduce the concept of a mixed strategy, let us again consider the matrix game  $A = \{a_{ij}\}$  with  $m$  rows and  $n$  columns. In this game, the "strategy space" of P1 comprises  $m$  elements, since he is allowed to choose one of the  $m$  rows of  $A$ . If  $T_1$  denotes this space of (pure) strategies, then  $T_1 = \{\text{row } 1, \dots, \text{row } m\}$ . If we allow mixed strategies for P1, however, he will pick probability distributions on  $T_1$ . That is, an allowable strategy for P1 is to choose "row 1" with probability (w.p.)  $x_1$ , "row 2" w.p.  $x_2$ , ..., "row  $m$ " w.p.  $x_m$ , where  $x_i \geq 0$  and  $x_1 + x_2 + \dots + x_m = 1$ . The mixed strategy space of P1, which we denote by  $Y$ , is now comprised of all such probability distributions. Since the probability distributions are discrete,  $Y$  simply becomes the space of all nonnegative numbers  $x_i$  that add up to 1, which is a simplex. Note that, the  $m$  pure strategies of P1 can also be considered as elements of  $Y$ , obtained by setting (i)  $x_1 = 1, x_j = 0, \forall j \neq 1$ ; (ii)  $x_2 = 1, x_i = 0, \forall i \neq 2$ ; The mixed strategy space of P2, which we denote by  $Z$ , can likewise be defined as an  $n$ -dimensional simplex. A precise definition for a mixed strategy now follows.

**Definition 2.2** A mixed strategy for a player is a probability distribution on the space of his pure strategies. Equivalently, it is a random variable whose values are the player's pure strategies. Thus, in matrix games, a mixed strategy for each player can be considered either as an element of a simplex, or as a

random variable with a discrete probability distribution. Typical mixed strategies for the players, under the latter convention, would be independent random variables  $u$  and  $v$ , defined, respectively, by  $\{1 \text{ w.p. } 2/1 \ m, \wedge, \forall i = h$   
 $\forall i > 0, (2.10a) \ m \text{ w.p. } 2/m \gg i \ 1 \ \text{W.p. } Z \setminus n, \ \xi^* i = 1, \ * >^\circ - (2-10b) \ n \text{ w.p. } z_n \ i = \setminus$   
 Now, if the players adopt these random variables as their strategies to be implemented during the course of a repeated game, they will have to base their  
 sThis definition of a mixed strategy is valid not only for matrix games but also for other types of finite games to be discussed in this and the following chapter.

24 T. BAŞAR AND G. J. OLSDER actual decisions (as to what specific row or column to choose during each play of the game) on the outcome of a chance mechanism, unless the probability distributions involved happen to be concentrated at one point—the case of pure strategies. If the adopted strategy of  $P_1$  is to pick "row 1" w.p.  $1/2$  and "row  $m$ " w.p.  $1/2$ , for example, then the player in question could actually implement this strategy by tossing a fair coin before each play of the game, playing "row 1" if the actual outcome is "head", and "row  $m$ " otherwise. It should be noted, here, that the actual play (action) dictated by  $P_1$ 's strategy becomes known (even to him) only after the outcome of the chance mechanism (tossing of a fair coin, in this case) is observed. Hence, we have a sharp distinction between a player's (proper mixed) strategy<sup>6</sup> and its implemented "value" for a particular play, where the latter explicitly depends on the outcome of a chance experiment designed so as to generate the odds dictated by the mixed strategy. Such a dichotomy between strategy and its implemented value does not exist for the class of pure strategies that we have discussed earlier in this section, since if a player adopts the strategy "to play row  $i$ ", for example, then its implementation is known for sure—he will play "row  $i$ ". Given an  $(m \times n)$  matrix game  $A$ , which is to be played repeatedly and the final outcome to be determined by averaging the scores of the players on individual plays, let us now investigate as to how this final outcome will be related to the (mixed) strategies adopted by the players. Let the independent random variables  $u$  and  $v$ , defined by (2.10a) and (2.10b), be the strategies adopted by  $P_1$  and  $P_2$ , respectively. Then, as the number of times this matrix game is played gets arbitrarily large, the frequencies with which different rows and columns of the matrix are chosen by  $P_1$  and  $P_2$ , respectively, will converge to their respective probability distributions that characterize the strategies  $u$  and  $v$ . Hence, the average value of the outcome of the game, corresponding to the strategy pair  $\{u, v\}$ , will be

equal to  $\sum_{i,j} y_i z_j = y'Az$  (2-11) where  $y$  and  $z$  are the probability distribution vectors defined by  $y = (y_1, \dots, y_m)$ ,  $z = (z_1, \dots, z_n)$ .

(2.12)  $P_1$  wishes to minimize this quantity,  $J(y, z)$ , by an appropriate choice of a probability distribution vector  $y \in Y$ , while  $P_2$  wishes to maximize the same quantity by choosing an appropriate  $z \in Z$ , where the sets  $Y$  and  $Z$  are, respectively, the  $m$ - and  $n$ -dimensional simplices introduced earlier, i.e.,  $Y = \{y \in \mathbb{R}^m : y_i \geq 0, \sum y_i = 1\}$ , (2.13a) Since, by definition, the concept of mixed strategy also covers pure strategy, we shall sometimes use the term "proper mixed" to emphasize (whenever the case is) that the underlying probability distribution is not one point.

TWO-PERSON ZERO-SUM FINITE GAMES 25  $n Z = \{z \in \mathbb{R}^n : z_j \geq 0, \sum z_j = 1\}$ .

(2.13b)  $j=i$  The following definitions now follow as obvious extensions to mixed strategies of some of the concepts introduced earlier for matrix games with strategy spaces comprised of only pure strategies. Definition 2.3 A vector  $y^* \in Y$  is called a mixed security strategy for  $P_1$ , in the matrix game  $A$ , if the following inequality holds for all  $y \in Y$ :  $V_m(A) = \max_{y^*} y^*Az < \max_{y \in Y} y'Az$ ,  $y \in Y$ .

(2.14)  $z \in Z$  Here, the quantity  $V_m$  is known as the average security level of  $P_1$  (or equivalently the average upper value of the game). Analogously, a vector  $z^* \in Z$  is called a mixed security strategy for  $P_2$ , in the matrix game  $A$ , if the following inequality holds for all  $z \in Z$ :  $V_{-m}(A) = \min_{z^*} y^*Az^* > \min_{z \in Z} y'Az$ ,  $z \in Z$ . (2.15)  $y \in Y$   $y \in Y$  Here,  $V_{-m}$  is known as the average security level of  $P_2$  (or equivalently, the average lower value of the game).

Definition 2.4 A pair of strategies  $\{y^*, z^*\}$  is said to constitute a saddle point for a matrix game  $A$ , in mixed strategies, if  $y^*Az < y^*Az^* < y'Az^*$ ,  $\forall y \in Y, z \in Z$ . (2.16) The quantity  $V_m(A) = y^*Az^*$  is known as the saddle-point value, or simply the value, of the game, in mixed strategies.

Remark 2.1 The assumptions of existence of a maximum in (2.14) and a minimum in (2.15) are justifiable since, in each case, the objective functional is linear (thereby continuous) and the simplex over which optimization is performed is closed and bounded (thereby compact). (See Appendix A.)  $\square$

As a direct extension of Thm. 2.1, we can now verify the following properties of the security levels and security strategies in matrix games when the players are allowed to use mixed strategies. Theorem 2.3 In every matrix game  $A$  in which the players are allowed to use mixed strategies, the following properties hold: (i) The average security level of each player is unique. (ii) There exists at least one mixed security strategy for each player.

26 T. BAŞAR AND G. J. OLSDER (Hi) The security levels in pure and mixed strategies satisfy the following inequalities:  $V(A) < V_m(A) < V_m(A) < V(A)$ .

(2.17) Proof, (i) Uniqueness of  $V_m$  follows directly from its definition,  $V^* = \inf_z \max_y y'Az$ , in view of Remark 2.1. Analogously, the quantity  $V_m = \sup_z \min_y y'Az$  is also unique. (ii) The quantity,  $\max_z z'y'Az$ , on the right-hand side (RHS) of (2.14) is a continuous function of  $y \in Y$ , a result which can be proven by employing standard methods of analysis. But since  $Y$  is a compact set, the minimum of this function is attained on  $Y$ , thus proving the desired result for the security strategy of P1. Proof of existence of a security strategy for P2 follows along similar lines. (iii) The middle inequality follows from the simple reasoning that the maxi- mizer's security level cannot be higher than the minimizer's security level. The other two inequalities follow, since the pure security strategies of the players are included in their mixed strategy spaces. D As an immediate consequence of part (ii) of Thm. 2.3 above, we obtain the following. Corollary 2.2 In a matrix game  $A$ , the average upper and lower values in mixed strategies are given, respectively, by  $V_m(A) = \min \max y'Az$  (2.18a) and  $V_m(A) = \max \min y'Az$ . (2.18b) Now, one of the important results of zero-sum game theory is that these upper and lower values are equal, that is,  $V_m(A) = V_m(A)$ . This result is given in Thm. 2.4, below, which is known as the "Minimax Theorem". In its proof, we shall need the following lemma.

Lemma 2.1 Let  $A$  be an arbitrary  $(m \times n)$ -dimensional matrix. Then, either (i) there exists a nonzero vector  $y \in R^m$ ,  $y > 0$  such that  $A'y < 0$ , or (ii) there exists a nonzero vector  $z \in R^n$ ,  $z > 0$  such that  $Az > 0$ . Proof. Consider in  $R^n$  the unit vectors  $e_1, \dots, e_n$ , together with the rows of  $A$ , to be indicated by  $a_i, i = 1, \dots, m$ . Denote the convex hull of these  $n + m$  vectors by  $C$ . Two possibilities exist: either  $0 \in C$  or  $0 \notin C$ . Assume first that  $0 \in C$ . Then there exist nonnegative coefficients  $y_i (i = 1, \dots, m)$  and  $r_j (j = 1, \dots, n)$ , normalized to 1,  $\sum_{i=1}^m y_i + \sum_{j=1}^n r_j = 1$ . such that  $m \sum_{i=1}^m y_i a_i + \sum_{j=1}^n r_j e_j = 0$ . Since all the  $y_i$ 's cannot be zero simultaneously (otherwise all  $r_j$ 's would also be zero— which is an impossibility due to the normalization), we have  $\sum_{j=1}^n r_j e_j < 0$ ,  $j = 1, \dots, n$ . Now, if we take  $y \in R^m$  as the vector whose components are the  $r_j$ 's defined above, then the validity of possibility (i) of the lemma follows. Assume next that  $0 \notin C$ . Then, there exists a hyperplane passing through the origin (characterized by the equation  $z'x = 0$ , where  $x, z \in R^n$ , and where  $x$  is

TWO-PERSON ZERO-SUM FINITE GAMES 27 Componentwise this equation reads  $\sum_{i=1}^m y_i a_{ij} + r_j = 0, j = 1, \dots, n$ . Since all the  $y_i$ 's cannot be zero simultaneously (otherwise all  $r_j$ 's would also be zero— which is an impossibility due to the normalization), we have  $\sum_{i=1}^m y_i a_{ij} < 0, j = 1, \dots, n$ . Now, if we take  $y \in R^m$  as the vector whose components are the  $r_j$ 's defined above, then the validity of possibility (i) of the lemma follows. Assume next that  $0 \notin C$ . Then, there exists a hyperplane passing through the origin (characterized by the equation  $z'x = 0$ , where  $x, z \in R^n$ , and where  $x$  is



the running variable), such that  $C$  is on one side of it; furthermore,  $z'x > 0$  for  $x \in C$ . If we choose  $x = t_i$  (which belongs to  $C$  by construction), then it follows that  $Z_i > 0$ , and hence the vector  $z$  is nonzero. Similarly, if we choose  $x = a^j$ , then  $z'a^j > 0$  and hence  $Az > 0$ , which validates the possibility (ii).  $\square$

**Theorem 2.4 (The Minimax Theorem)** In any matrix game  $A$ , the average security levels of the players in mixed strategies coincide, that is,  $V_m(A) = \min_{y \in Y} \max_{z \in Z} y'Az = \max_{z \in Z} \min_{y \in Y} y'Az = V_{-m}(A)$ . (2.19)

**Proof.** We first show, by applying Lemma 2.1, that for a given matrix game  $A$ , at least one of the following two inequalities holds:  $V_m(A) < 0$ , (i)  $V_m(A) > 0$ . (ii) Assume that the first alternative of Lemma 2.1 is valid. Then, there exists a vector  $y^\circ \in Y$  such that  $A'y^\circ < 0$ , which is equivalent to the statement that the inequality  $y^\circ'Az < 0$  holds for all  $z \in Z$ . This is further equivalent to saying  $\max_{z \in Z} y^\circ'Az < 0$  which definitely implies (by also making use of (2.18a))  $V_m(A) = \min_{y \in Y} \max_{z \in Z} y'Az < 0$ . Now, under the second alternative of Lemma 2.1, there exists a vector  $z^\circ \in Z$  such that  $Az^\circ > 0$ , or equivalently,  $y'Az^\circ > 0, \forall y \in Y$ . This is further equivalent to the inequality  $\min_{y \in Y} y'Az^\circ > 0$ ,

28 T. BAŞAR AND G. J. OLSDER which finally implies (ii), by also making use of (2.18b), since  $V_{-m}(A) = \max_{z \in Z} \min_{y \in Y} y'Az > \min_{y \in Y} y'Az^\circ$ . Thus we see that, under the first alternative of Lemma 2.1, inequality (i) holds; and under the second alternative, inequality (ii) remains valid. Let us now consider a new  $(m \times n)$ -dimensional matrix  $B = \{b_{ij}\}$  obtained by shifting all entries of  $A$  by a constant,  $c$ , that is,  $b_{ij} = a_{ij} - c$ . This results in a shift of the same amount in both  $V_m(A)$  and  $V_{-m}(A)$ , that is,  $V_m(B) = V_m(A) - c$  and  $V_{-m}(B) = V_{-m}(A) - c$ . Since matrix  $A$  was arbitrary in (i) and (ii), we replace  $A$  with  $B$ , as defined above, in these inequalities, and arrive at the following property with regard to the upper and lower values of a matrix game  $A$ , in mixed strategies. For a given matrix game  $A$  and an arbitrary constant  $c$ , at least one of the following two inequalities holds true:  $V_m(A) < c$ , (iii)  $V_m(A) > c$ . (iv) But, for this statement to be valid for arbitrary  $c$ , it is necessary that  $V_m(A) = V_{-m}(A)$ . For, otherwise, the only possibility is  $V_m(A) = V_{-m}(A) + k$ , in view of the middle inequality of (2.17), where  $k > 0$ ; and picking  $c = V_{-m}(A) + (1/2)k$  in (iii) and (iv), it readily follows that neither (iii) nor (iv) is satisfied. Thus, this completes the proof of the theorem.  $\square$

**Corollary 2.3** Let  $A$  denote an  $(m \times n)$  matrix game. Then, (i)  $A$  has a saddle point in mixed strategies, (ii) a pair of mixed strategies provides a saddle point for  $A$  if, and only if, the first of these is a mixed security strategy for  $PI$ , and the second one is a mixed security

strategy for P2, (iii)  $V_m(A)$  is uniquely given by  $V_m\{A\} = V_m(A) = Y_m(A)$ , (iv) in case of multiple saddle points, the mixed saddle-point strategies possess the ordered interchangeability property. Proof. This corollary to Thm. 2.4 is actually an extension of Thm. 2.2 and Corollary 2.1 to the case of mixed strategies, and its verification is along the same lines as that of Thm. 2.2. It should only be noted that equality of the average upper and lower values of the matrix game is a fact, in the present case, rather than a part of the hypothesis, as in Thm. 2.2. D We have thus seen that, if the players are allowed to use mixed strategies, matrix games always admit a saddle-point solution which, thereby, manifests itself as the only reasonable equilibrium solution in zero-sum two-person games of that type.

TWO-PERSON ZERO-SUM FINITE GAMES 29 2.3 Computation of Mixed Equilibrium Strategies We have seen in the previous section that two-person zero-sum matrix games always admit a saddle-point equilibrium in mixed strategies (cf. Thm. 2.4). One important property of mixed saddle-point strategies is that, for each player, the corresponding one also constitutes a mixed security strategy, and conversely, each mixed security strategy is a mixed saddle-point strategy (Corollary 2.3(ii)). This property, then, strongly suggests a possible way of obtaining the mixed saddle-point solution of a matrix game, which is to determine the mixed security strategy(ies) of each player. To illustrate this approach, let us consider the  $(2 \times 2)$  matrix game P2 3 -1 0 1 which clearly does not admit a pure-strategy saddle point, since  $V = 1$  and  $V_- = 0$ . Let the mixed strategies of P1 and P2 be denoted by  $y = (y_1, y_2)'$  and  $z = (z_1, z_2)'$ , respectively, with  $y_i > 0$ ,  $z_i > 0$  ( $i = 1, 2$ ),  $y_1 + y_2 = z_1 + z_2 = 1$ , and consider first the average security level of P1. It should be noted that, while determining the average security level of a player, we can assume, without any loss of generality, that the other player chooses only pure strategies (this follows directly from Defs. 2.3 and 2.4). Hence, in the present case, we can take P2 to play either  $\{z_1 = 1, z_2 = 0\}$  or  $\{z_1 = 0, z_2 = 1\}$ ; and under different choices of mixed strategies for P1, we can determine the average outcome of the game as shown in Fig. 2.1 by the bold line, which forms the upper envelope to the two straight lines drawn. Now, if the mixed strategy  $(y_1^*, y_2^*) = (2/5, 3/5)$  corresponding to the lowest point of that envelope is adopted by P1, then the average outcome will be no greater than  $1/5$ . For any other mixed strategy of P1, however, P2 can  $y_1 = 1, y_2 = 0$  or  $y_1 = 0, y_2 = 1$  Figure 2.1: Mixed security strategy of P1 for the matrix game (2.20).

30 T. BAŞAR AND G. J. OLSDER 3 2 1 0.6 0 -1  $z_1 = 1$   $z_2 = 1/5$   $z_3 = 0$   $Z_2 = 0$   $z_1 = 4/5$   $Z_2 = 1$  Figure 2.2: Mixed security strategy of P2 for the matrix game (2.20). obtain a higher average outcome. (If, for instance,  $z_1 = 5, z_2 = 5$ , then P2 can increase the outcome by playing  $z_1 = 1, z_2 = 0$ .) This then implies that the strategy  $(t_i = |, j/2 = \xi)$  is a mixed security strategy for P1 (and his only one), and thereby, it is his mixed saddle-point strategy. The mixed saddle-point value can easily be read off from Fig. 2.1 to be  $V_m = \xi$ . To determine the mixed saddle-point strategy of P2, we now consider his security level, this time assuming that P1 adopts pure strategies. Then, for different mixed strategies of P2, the average outcome of the game can be determined to be the bold line, shown in Fig. 2.2, which forms a lower envelope to the two straight lines drawn. Since P2 is the maximizer, the highest point on this envelope is his average security level, which he can guarantee (on the average) by playing the mixed strategy  $(z_1 = |, z_2 = |)$  which is also his saddle-point strategy. The computational technique discussed above is known as the "graphical solution" of matrix games, since it makes use of a graphical construction directly related to the entries of the matrix. Such an approach is practical not only for  $(2 \times 2)$  matrix games but also for general  $(2 \times n)$  and  $(m \times 2)$  matrix games. Consider, for example, the  $(2 \times 3)$  matrix game  $P_2 \begin{matrix} 1 & 6 & 3 \\ 2 & 0 & 7 \end{matrix}$  for which P1's average security level in mixed strategies has been determined in Fig. 2.3. Assuming again that P2 uses only pure strategies ( $z_1 = 1, z_2 = z_3 = 0$  or  $z_1 = z_3 = 0, z_2 = 1$  or  $z_1 = z_2 = 0, z_3 = 1$ ), the average outcome of the game for different mixed strategies of P1 is given by the bold line drawn in Fig. 2.3, which is again the upper envelope (this time formed by three straight lines). The lowest point on this envelope yields the mixed security strategy, and thereby the mixed saddle-point strategy, of P1, which is  $(j/j^* = |, j/| = 5)$ -

TWO-PERSON ZERO-SUM FINITE GAMES 31  $y_1 = 1$   $y_2 = 2/3$   $y_3 = 0$   $y_2^* = 1/3$   $y_2 = 1$  Figure 2.3: Mixed security strategy of P1 for the matrix game (2.21). The mixed saddle-point value, in this case, is  $V_m = |$ . To determine the mixed saddle-point strategy of P2, we first note that the average security level of P1 has been determined by the intersection of only two straight lines which correspond to two pure strategies of P2, namely "column 1" and "column 2". If P1 uses his mixed security strategy, and P2 knows that before he plays, he would of course be indifferent between these two pure strategies, but he would never play the third pure strategy, "column 3", since that will reduce the average outcome of the game. This then implies that, in the actual saddle-point

solution, he will "mix" only between the first two pure strategies. Hence, in the process of determining P2's mixed security strategies, we can consider the  $(2 \times 2)$  matrix game  $P2 \begin{matrix} 1 & 6 & 3 & 2 \end{matrix}$  which is obtained from (2.21) by deleting the last column. Using the graphical approach, P2's mixed security strategy in this matrix game can readily be determined to be  $[z^1 = \frac{1}{2}, z^2 = \frac{1}{2}]$  which is also his mixed saddle-point strategy. Relation with linear programming One alternative to the graphical solution when the matrix dimensions are high is to convert the original matrix game into a linear programming (LP) problem and make use of some of the powerful algorithms available for LP. To elucidate the close connection between a two-person zero-sum matrix game and an LP problem, let us start with an  $(m \times n)$  matrix game  $A = \{a_{ij}\}$  with all entries

32 T. BAŞAR AND G. J. OLSDER positive (i.e.,  $a_{ij} > 0 \forall i = 1, \dots, m; j = 1, \dots, n$ ). Then, the average value of this game, in mixed strategies, is given by  $V_m(A) = \min_{i'} \max_{z} i'Az = \max_{z} \min_{i'} i'Az$ , (2.23) which is necessarily a positive quantity by our positivity assumption on  $A$ . Let us now consider the min-max operation used in (2.23). Here, first a  $y \in Y$  is given, and then the resulting expression is maximized over  $Z$  that is, the choice of  $z \in Z$  can depend on  $y$ . Hence, the middle expression of (2.23) can also be written as  $\min_{i'} v_i(y), y \in Y$  where  $v_i(y)$  is a positive function of  $y$ , defined by  $v_i(y) = \max_{z \in Z} y'Az > y'Az \forall z \in Z$ . (2.24) Since  $Z$  is the  $n$ -dimensional simplex as defined by (2.13b), the inequality in (2.24) becomes equivalent to the vector inequality  $A'y < \ln v_i(y)$ , where  $\ln^i(i, \dots, i) \in \mathbb{R}^n$ . Further introducing the notation  $y = y/v(y)$  and recalling the definition of  $Y$  from (2.13a), we observe that the optimization problem faced by PI in determining his mixed security strategy is minimize  $v(y)$  over  $\mathbb{R}^m$  subject to This is subject further to  $A'y < \ln, y > 0$ ,  $y = y/v(y)$ . equivalent to the maximization problem  $\max_{y \in Y} y' \ln, y > 0$ , (2.25a) (2.25b) (2.25c) which is a standard LP problem. The solution of this LP problem will give the mixed security strategy of PI, normalized with the average saddle-point

TWO-PERSON ZERO-SUM FINITE GAMES 33 value of the game,  $V_m(A)$ . Hence, if  $y^* \in Y$  is a mixed saddle-point strategy of PI, in the matrix game  $A$ , then the quantity  $y^* = y^*/V_m(A)$  solves the LP problem (2.25a)-(2.25c), whose optimal value is in fact  $1/V_m(A)$ . Conversely, every solution of the LP problem will correspond to a mixed strategy of PI in the same manner. Now, if we instead consider the right-hand expression of (2.23), and introduce  $v_2(z) =$

$\max_{y \in Y} y'Az < y'Az \forall y \in Y$ , and  $z = z/v_2(z)$ , following similar steps and reasoning leads us to the minimization problem  $\min z'l_n$  (2.26a) subject to  $Az > l_m$ , (2.26b)  $z > 0$ , (2.26c) which is the "dual" of (2.25a)-(2.25c). We thus arrive at the conclusion that if  $z^* \in Z$  is a mixed saddle-point strategy of P2, in the matrix game A, then  $z^* - z^*/V_m(A)$  solves the LP problem (2.26a)-(2.26c), whose optimal value is again  $1/V_m(A)$ . Furthermore, the converse statement is also valid. We have thus shown that, given a matrix game with all positive entries, there exist two LP problems (duals of each other) whose solutions yield the saddle-point solution(s) of the matrix game. The positivity of the matrix A, however, is only a convention here, and can easily be dispensed with, as the following lemma shows.

**Lemma 2.2** Let A and B be two  $(m \times n)$ -dimensional matrices related to each other by the relation  $A = B + cI_n$ , (2.27) where c is some constant. Then, (i) every mixed saddle-point strategy pair for matrix game A also constitutes a mixed saddle-point solution for matrix game B, and vice versa, (ii)  $V_m(A) = V_m(B) + c$ . **Proof.** Let  $(y^*, z^*)$  be a saddle-point solution for A, thus satisfying inequalities (2.16). If A is replaced by  $B + cI_n$  in (2.16), then it is easy to see that  $(y^*, z^*)$  also constitutes a saddle-point solution for B, since  $y'l_n z = 1$  for every  $y \in Y, z \in Z$ . Since the reverse argument also applies, this completes the proof of part (i). Part (ii), on the other hand, follows from the relation  $V_m(A) = y^*[B + cI_n]z^* = y^*Bz^* + cy^*I_n z^* = V_m(B) + c$ . Q

34 T. BAŞAR AND G. J. OLSDER Because of property (i) of the above lemma, we call matrix games which satisfy relation (2.27) strategically equivalent matrix games. It should now be apparent that, given a matrix game A, we can always find a strategically equivalent matrix game with all entries positive, and thus the transformation of a matrix game into two LP problems as discussed prior to Lemma 2.2 is always valid, as far as determination of saddle-point strategies is concerned. This result is summarized below in Thm. 2.5.

**Theorem 2.5** Let B be an  $(m \times n)$  matrix game, and A be defined by (2.27) with c chosen to make all its entries positive. Introduce the two LP problems  $\max_{y \in Y} y'Az < l_n >$  "primal problem", (2.28a)  $\min_{z \in Z} z'l_n \wedge Az > l_m$  "dual problem", (2.28b) with their optimal values (if they exist) denoted by  $V_p$  and  $V_a$ , respectively. (i) Both LP problems admit a solution, and  $V_p = V_a = 1/V_m(A)$ . (ii) If  $(y^*, z^*)$  is a mixed saddle-point solution of the matrix game B, then  $y^*/V_m(A)$  solves (2.28a), and  $z^*/V_m(A)$  solves (2.28b). (iii) If  $y^*$  is a solution of (2.28a), and  $z^*$  is a solution of (2.28b), the pair

$(y^*/V_p, z^*/V_d)$  constitutes a mixed saddle-point solution for matrix game B. Furthermore,  $V_m(B) = (1/V_p) - c$ . Proof. Since A is a positive matrix and strategically equivalent to B, the theorem has already been proven prior to the statement of Lemma 2.2. We should only note that the reason why both LP problems admit a solution is that the mixed security strategies of the players always exist in the matrix game A. D Theorem 2.5 provides one method of making games "essentially" equivalent to two LP problems that are duals of each other, thus enabling the use of some powerful algorithms available for LP in order to obtain their saddle-point solutions. (The reader is referred to Dantzig (1963), Luenberger (1973), Gonzaga (1992), or Karmarkar (1984) for LP algorithms.) There are also other transformation methods available in the literature, which form different kinds of equivalences between matrix games and LP problems, but the underlying idea is essentially the same in all these techniques. For one such equivalence the reader is referred to the next chapter, to Corollary 3.2, where an LP problem, not equivalent to the one of Thm. 2.5, is obtained as a special case of a more general result on nonzero-sum matrix games. More details on computational techniques for matrix games can be found in Luce and Raiffa (1957) and Singleton and Tyndal (1974).

TWO-PERSON ZERO-SUM FINITE GAMES 35 Dominating strategies We conclude this section by introducing the concept of "dominance" in matrix games, which can sometimes be useful in reducing the dimensions of a matrix game by eliminating some of its rows and/or columns which are known from the very beginning to have no influence on the equilibrium solution. More precisely, given an  $(m \times n)$  matrix game  $A = \{a^i_j\}$ , we say that "row  $i$ " dominates "row  $k$ " if  $a^i_j < a^k_j, j = 1, \dots, n$ , and if, for at least one  $j$ , the strict inequality-sign holds. In terms of pure strategies, this means that the choice of the dominating strategy, i.e., "row  $i$ ", is at least as good as the choice of the dominated strategy, i.e., "row  $k$ ". If, in the above set of inequalities, the strict inequality-sign holds for all  $j = 1, \dots, n$ , then we say that "row  $i$ " strictly dominates "row  $k$ ", in which case, regardless of what P2 chooses, P1 does better with "row  $i$ " (strictly dominating strategy) than with "row  $k$ " (strictly dominated strategy). It therefore follows that P1 can always dispense with his strictly dominated strategies and consider only strictly undominated ones, since adoption of a strictly dominated strategy is apt to increase the security level for P1. This argument also holds for mixed strategies, and strictly dominated rows are always assigned a probability of zero in an optimal (mixed) strategy for P1.

Analogously for P2, "column  $j$ " of  $A$  is said to dominate (respectively, strictly dominate) "column  $P$ ", if  $a^j > a^P$ ,  $i = 1, \dots, m$ , and if, for at least one  $i$  (respectively, for all  $i$ ), the strict inequality-sign holds. In an optimal (mixed) strategy for P2, strictly dominated columns are also assigned a probability of zero. Mathematical verifications of these intuitively obvious assertions, as well as some other properties of strictly dominating strategies, can be found in Vorob'ev (1977). We have thus seen that, in the computation of (pure or mixed) saddle-point equilibria of matrix games, strictly dominated rows and columns can readily be deleted, since they do not contribute to the equilibrium solution—and this could lead to considerable simplifications in the graphical solution or in the solution of the equivalent LP problems, since the matrix is now of smaller dimension(s). With (nonstrictly) dominated rows and columns, however, this may not always be so. Even though every saddle-point solution of a matrix game whose dominated rows and columns are deleted is also a saddle-point solution of the original matrix game, there might be other saddle points (of the original game) which are eliminated in this process. As a specific illustration of this possibility, consider the  $(2 \times 2)$  matrix game P2 0 -1 X, X which admits two saddle points, as indicated. For P1, "row 2" is clearly dominating (but not strictly dominating), and for P2, "column 2" is strictly dominating. Thus, by eliminating dominated rows and columns, we end up with the  $(1 \times 1)$  game corresponding to the dominating strategies of the players, which trivially has a unique saddle point.

36 T. BAŞAR AND G. J. OLSDER Motivated by this result, we call the saddle-point solutions of the "reduced" game which does not have any dominated rows or columns dominant saddle-point solutions. Even though the original game might have other saddle points, it may still be reasonable to confine attention only to dominant saddle points, since any game that possesses a saddle point also possesses (by definition) a dominant saddle point, and furthermore the saddle-point value (in pure or mixed strategies) is unique; in other words, there is nothing essential to gain in seeking saddle-point equilibria which are not dominant.

#### 2.4 Extensive Forms: Single-Act Games

There exist different mathematical descriptions for two-person zero-sum games with finite strategy sets, the matrix form (also known as the normal form) being one of these. A normal form does not in general provide the full picture of the underlying decision process, since it only describes the correspondences between different ordered strategy pairs and outcomes. Some

important issues like the order of play in the decision process, information available to the players at the time of their decisions, and the evolution of the game in the case of dynamic situations are suppressed in the matrix description of a zero-sum game. An alternative to the normal form, which explicitly displays the dynamic character of the decision problem, is known as the extensive form of a two-person zero-sum game. An extensive form basically involves a tree structure with several nodes and branches, providing an explicit description of the order of play and the information available to each player at the time of his decision(s); the game evolves from the top of the tree to the tip of one of its branches. Two such tree structures are depicted in Fig. 2.4, where in each case P1 has two alternatives (branches) to choose from, whereas P2 has three alternatives, and the order of play is such that P1 acts before P2 does. The numbers at the end of the lower branches represent the pay-offs to P2 (or equivalently, losses incurred to P1) if the corresponding paths are selected by the players. The uppercase letters L, M, R stand for Left, Middle, and Right branches, respectively. 130627130627 (a) (b) Figure 2.4: Two zero-sum games in extensive form differing only in the information available to P2.

TWO-PERSON ZERO-SUM FINITE GAMES 37 It should be noted that the two zero-sum games displayed in Figs 2.4(a) and 2.4(b) are equivalent in every aspect other than the information available to P2 at the time of his play, which is indicated on the tree diagrams by dotted lines enclosing an area (known as the information set) including the relevant nodes. In Fig. 2.4(a), the two possible nodes of P2 are included in the same dotted area, implying that, even though P1 acts before P2 does, P2 does not have access to his opponent's action. That is to say, at the time of his play, P2 does not know at what node he really is. This is, of course, equivalent to the case when both players act simultaneously. Hence, simultaneous play can also be represented by a tree structure, provided that the relevant information set is chosen properly. The extensive form of Fig. 2.4(a) can now easily be converted into the normal form, with the equivalent matrix game being the one given by (2.21) in Section 2.3. As it has been discussed there, this matrix game admits a saddle-point solution in mixed strategies, which is  $(y^*L = \frac{1}{3}, y^*R = \frac{2}{3})$  with the mixed saddle-point value being  $V_m = \frac{8}{3}$ . The extensive form of Fig. 2.4(b), however, admits a different matrix game as its normal form and induces a different behavior on the players. In this case, each node of P2 is



included in a separate information set, thus implying that P2 has perfect information as to which branch of the tree P1 has chosen. Hence, if  $u_1$  denotes the actual choice of P1 and  $\sigma_2(\cdot)$  denotes a strategy for P2, as a maximizer P2's optimal choice will be  $\sigma_2^*(u_1) = \{R \text{ if } u_1 = L; L \text{ if } u_1 = R\}$ ; (2.29a) Not knowing this situation ahead of time, P1 definitely adopts the strategy  $\sigma_1^* = L$  (2.29b) with the equilibrium outcome of the game being 3, which is, in fact, the upper value of the matrix game (2.21). Thus, we have obtained a solution of the zero-sum game of Fig. 2.4(b) by directly making use of the extensive tree formulation. Let us now attempt to obtain the (same) saddle-point solution by transforming the extensive form into an equivalent normal form. To this end, we first delineate the possible strategies of the players. For P1, there are two possibilities:  $\sigma_1 = L$  and  $\sigma_1 = R$ . For P2, however, since he observes the action of P1, there exist  $2^2 = 4$  possible strategies, which are  $\sigma_2(u_1) = L$ ,  $\sigma_2(u_1) = R$ ,  $\sigma_2(u_1) = M$ ,  $\sigma_2(u_1) = R$ ,  $L \text{ if } u_1 = L, M \text{ if } u_1 = R, R \text{ if } u_1 = L, L \text{ if } u_1 = R$ ,

38 T. BAŞAR AND G. J. OLSDER (2.29a) (2.29b) (2.29c) where the subscripts  $i = 1, \dots, 9$  denote a particular ordering (labeling) of the possible strategies of P2. Hence, the equivalent normal form of the zero-sum game of Fig. 2.4(b) is the  $2 \times 9$  matrix game  $\begin{matrix} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ L & 1 & 7 & 1 & 6 & 3 & 2 & 0 & 7 & 1 \\ R & 1 & 2 & 0 & 6 & 7 & 1 & 2 & 0 & 6 \end{matrix}$  which admits two saddle points, as indicated, with the dominant one being  $\{L, \text{column } 7\}$ . It is this dominant saddle-point solution that corresponds to the one given by (2.29a)-(2.29b), and thus we observe that the derivation outlined earlier, which utilizes directly the extensive tree structure, could cancel out dominated saddle points. This, however, does not lead to any real loss of generality, since, if a matrix game admits a pure-strategy saddle point, it also admits a dominant pure-strategy saddle point, and furthermore, the saddle-point value is unique (regardless of the number of pure-strategy equilibria). We should now note that the notion of "(pure) strategy" introduced above within the context of the zero-sum game of Fig. 2.4(b) is somewhat different from (and, in fact, more general than) the one adopted in Section 2.2. This difference arises mainly because of the dynamic character of the decision problem of Fig. 2.4(b)—P2 being in a position to know exactly how P1 acts. A (pure) strategy for P2, in this case, is a rule that tells him which one of his alternatives to choose, for each possible action of P1. In mathematical language, it is a mapping from the collection of his information sets into the set of his alternatives. We thus see that there is a rather sharp distinction between a strategy and the actual action

dictated by that strategy, unless there is only one information set (which is the case in Fig. 2.4(a)). In the case of a single information set, the notions of strategy and action definitely coincide, and that is why we have used these two terms interchangeably in Section 2.2 while discussing matrix games with pure-strategy solutions. We are now in a position to give precise definitions of some of the concepts introduced above.

**Definition 2.5** An extensive form of a two-person zero-sum finite game without chance moves is a finite tree structure with (i) a specific vertex indicating the starting point of the game, (ii) a partition of the nodes of the tree into two player sets (to be denoted by  $N_1$  and  $N_2$  for  $P_1$  and  $P_2$ , respectively), (iii) a subpartition of each player set  $N_i$  into information sets  $\{I_i^r\}$ , such that the same number of immediate branches emanates from every node belonging to the same information set, and no node follows another node in the same information set.

**Remark 2.2** This definition of an extensive form covers more general types of two-person zero-sum games than the ones discussed heretofore since it also allows a player to act more than once in the decision process, with the information sets being different at different levels of play. Such games are known as multi-act games, and they will be discussed in Section 2.5. In the remaining parts of this section we shall confine our analysis to single-act games—games in which each player acts only once.  $\square$

**Remark 2.3** It is possible to extend Def. 2.5 so that it also allows for chance moves by a third party called "nature". Such an extension will be incorporated in Section 2.6. In the present section, and also in Section 2.5, we only consider the class of zero-sum finite games which do not incorporate any chance moves.  $\square$

**Definition 2.6** Let  $N_i^0$  denote the class of all information sets of  $P_i$ , with a typical element designated as  $r_i$ . Let  $U_i^r$  denote the set of alternatives of  $P_i$  at the nodes belonging to the information set  $r_i$ . Define  $U_i = \bigcup_{r_i \in N_i^0} U_i^r$  where the union is over  $r_i \in N_i^0$ . Then, a strategy  $Y$  for  $P_i$  is a mapping from  $N_i^0$  into  $U_i$ , assigning one element in  $U_i^r$  for each set in  $N_i^0$ , and with the further property that  $Y(r_i) \in U_i^r$  for each  $r_i \in N_i^0$ . The set of all strategies for  $P_i$  is called his strategy set  $S_i$ , and it is denoted by  $P_i$ .

**Remark 2.4** For the starting player, it is convenient to take  $T^0 = \{v\}$  a singleton, to eliminate any possible ambiguity.  $\square$

**Example 2.1** In the extensive form of Fig. 2.4(b),  $P_2$  has two information sets

( $N_2 = \{(\text{node } 1), (\text{node } 2)\}$ ) and three alternatives ( $U_2 = \{L, M, R\}$ ). Therefore, he has nine possible strategies which have actually been listed earlier during the normalization phase of this extensive form. Here, by an abuse of notation,  $r_2 = \{u_1\}$ .  $\square$  For tree structures in which an information set contains more than one node from any directed path in the tree or for games in which the underlying graph contains cycles (and hence is not a tree anymore), the reader is referred to subsection 2.7.1.

40 T. BAŞAR AND G. J. OLSDER Definition 2.7 A strategy  $J(-)$  of  $P_i$  is said to be constant if its value does not depend on its argument  $r \setminus i$ . Saddle-point equilibria of single-act games The pure-strategy equilibrium solution of single-act two-person zero-sum games in extensive form can be defined in a way analogous to Def. 2.1, with appropriate notational modifications. To this end, let  $J(r_1, r_2)$  denote the numerical outcome of a game in extensive form, interpreted as the loss incurred to  $P_1$  when  $P_1$  and  $P_2$  employ the strategies  $r_1 \in T_1$  and  $r_2 \in T_2$ , respectively. This loss function in fact defines the correspondences between different ordered strategy pairs and outcomes, and thus describes the equivalent normal form of the zero-sum game. Then, we have the following. Definition 2.8 A pair of strategies  $\{r_1^* \in T_1, r_2^* \in T_2\}$  is in saddle-point equilibrium if the following set of inequalities is satisfied for all  $r_1 \in T_1, r_2 \in T_2$ :  $J(r_1^*, r_2) \leq J(r_1^*, r_2^*) \leq J(r_1, r_2^*)$  (2-30) The quantity  $J(r_1^*, r_2^*)$  is known as the saddle-point value of the zero-sum game. Since a saddle-point equilibrium is defined in terms of the normal form of a game, a direct method for solving two-person single-act games in extensive form would be first to convert them into equivalent normal form and solve for the saddle-point solution of the resulting matrix game, and then to interpret this solution in terms of the original game through the one-to-one correspondences that exist between the rows and columns of the matrix and strategies of the players. This method has already been illustrated in this section in conjunction with the solution of the zero-sum game of Fig. 2.4(b), and as we have observed there, a major disadvantage of such a direct approach is that one has to list down all possible strategies of the players and consider a rather high dimensional matrix game, especially when the second-acting player has several information sets. An alternative to this approach exists, which makes direct use of the extensive form description of the zero-sum game. This, in fact, is a recursive procedure, and it obtains the solution without necessarily considering all strategy combinations, especially when the second-acting

player has more than one information set. The strategies not considered in this procedure are the dominated ones, and this has actually been observed earlier in this section when we employed the technique to arrive at the equilibrium strategy pair (2.29a)- (2.29b) for the zero-sum game of Fig. 2.4(b). Before providing a general outline of the steps involved in this derivation, let us first consider another, somewhat more elaborate, example which is the zero-sum game whose extensive form is depicted in Fig. 2.5. To determine the saddle-point strategies associated with this zero-sum game, we first note that if  $u_1 = R$ , then P2 should choose  $u_2(i) = L$ , to result in a loss

TWO-PERSON ZERO-SUM FINITE GAMES 41 3 1-1 12 0 Figure 2.5: An extensive form that admits a saddle-point equilibrium in pure strategies. of 2 to PI. If PI picks L or M, however, P2 cannot differentiate between these two, since they belong to the same information set; and hence now a matrix game has to be solved for that part of the tree. The equivalent normal form is P2 3 -1 1 W L R which clearly admits the unique saddle-point solution  $u_1 = M$ ,  $u_2 = R$ , yielding a cost of 1 for PI. Hence, in the actual play of the game, PI will always play  $u_1 = u_1 = M$ , and P2's optimal response will be  $u_2 = R$ . However, the saddle-point strategy of P2 is and not the constant strategy  $u_2(u_1) = R$ , since if P2 adopts this constant strategy, then PI can switch to R to a loss of 0 instead of 1. The strategy for PI that is in equilibrium with (2.31) is still  $u_1^* = M$ , (2.32) since AT1 is a singleton. The reader can now easily check that the pair (2.31)- (2.32) satisfies the saddle-point inequalities (2.30). The preceding analysis now readily suggests a method of obtaining the pure- strategy saddle-point solution of single-act zero-sum games in extensive form, provided that it exists. The steps involved are as follows. Derivation of pure-strategy saddle-point solutions of single-act games in extensive form (1) For each information set of the second-acting player, solve the corresponding matrix game (assuming that each such matrix game admits a pure- strategy saddle point). For the case in which an information set is a

42 T. BAŞAR AND G. J. OLSDER singleton, the matrix game involved is a degenerate one with one of the players having a single choice—but such (essentially one-person) games always admit well-defined pure-strategy solutions. (2) Record the saddle-point value of each of these matrix games. If PI is the starting player, then the lowest of these is the saddle-point value of the extensive form; if P2 is the starting player, then the highest one is. (3) The

saddle-point strategy of the first-acting player is his saddle-point strategy obtained for the matrix game whose saddle-point value corresponds to the saddle-point value of the extensive game. (4) For the second-acting player, the saddle-point strategy is comprised of his saddle-point solutions in all the matrix games considered, by appropriately identifying them with each of his information sets. We leave it to the reader to verify that any pair of equilibrium strategies obtained by following the preceding procedure does, indeed, constitute a saddle-point strategy pair that satisfies (2.30). The following conclusions can now readily be drawn. Proposition 2.1 A single-act zero-sum two-person finite game in extensive form admits a (pure-strategy) saddle-point solution if, and only if, each matrix game corresponding to the information sets of the second-acting player has a saddle point in pure strategies. Proposition 2.2 Every single-act zero-sum two-person finite game in extensive form, in which the information sets of the second-acting player are singletons,<sup>8</sup> admits a pure-strategy saddle-point solution. If the matrix game corresponding to an information set does not admit a saddle-point solution in pure strategies, then it is clear that the strategy spaces of the players have to be enlarged, in a way similar to the introduction of mixed strategies in Section 2.2 within the context of matrix games. To pave the way for such an enlargement in the case of games in extensive form, let us first consider the single-act two-person zero-sum game whose extensive form is depicted in Fig. 2.6. Here, P2 is the second-acting player and he has two information sets. If his opponent picks R, then he can observe that choice, and his best response (strategy) in that case is L, yielding an outcome of 1. If P1 chooses L or M, however, then P2 is ignorant about that choice, since both of these nodes are included in the same information set of P2. The relevant matrix game, in this case, is P2

	L	M	R
L	3	-1	0
M	1	1	1

8The node corresponding to each such information set is known as the "state" of the game at that level (stage) of play. Games of this type are known as perfect information games.

TWO-PERSON ZERO-SUM FINITE GAMES 43

Figure 2.6: A game in extensive form that admits a mixed (behavioral)-strategy saddle-point solution. which is the matrix game (2.20) considered earlier in Section 2.3, and it is known to admit the mixed saddle-point solution  $(x^*, y^*) = (\frac{2}{5}, \frac{3}{5})$  for P1 and  $(z^*, w^*) = (\frac{2}{5}, \frac{3}{5})$  for P2, yielding an average outcome of  $\frac{1}{5}$ . Since  $\frac{1}{5} < 1$ , it is clear that P1 will prefer to stay on that part of the tree and thus stick to the strategy L w.p.  $\frac{2}{5}$ , M w.p.  $\frac{3}{5}$ , (2.33a) R w.p. 0, which is "mixed" in nature.

P2's equilibrium strategy will then be  $\gamma_1 = L$  w.p.  $\frac{1}{5}$ ,  $\gamma_2 = R$  w.p.  $\frac{4}{5}$  if  $u_1 = R$ , otherwise (2.33b) since he also has to consider the possibility of P1 playing R (by definition of a strategy). The pair (2.33a)-(2.33b) now yields the average outcome of  $\frac{1}{5}$ . In order to declare (2.33a)-(2.33b) as the equilibrium solution pair of the extensive form of Fig. 2.6, we have to specify the spaces in which it possesses such a property. To this end, we first note that P1 has three possible pure strategies,  $\gamma_1 = L$ ,  $\gamma_2 = M$  and  $\gamma_3 = R$ , and in view of this, (2.33a) is indeed a mixed strategy for P1 (see Def. 2.2). For P2, however, there exist four possible pure strategies, namely  $\gamma_1 = L$  if  $u_1 = R$ ,  $\gamma_2 = R$  otherwise,  $\gamma_3 = R$  if  $u_1 = R$ ,  $\gamma_4 = L$  otherwise. Hence, a mixed strategy for P2 is a probability law according to which these four strategies will be "mixed" during the play of the game, and (2.33b) provides

44 T. BAŞAR AND G. J. OLSDER one such probability distribution which assigns a weight of  $\frac{1}{5}$  to the first pure strategy and a weight of  $\frac{4}{5}$  to the third one. We have thus verified that both (2.33a) and (2.33b) are well-defined mixed strategies for P1 and P2, respectively. Now, if  $J(\gamma_1, \gamma_2)$  denotes the average outcome of the game when the mixed strategies  $\gamma_1$  and  $\gamma_2$  are adopted by P1 and P2, respectively, the reader can easily check that the pair (2.33a)-(2.33b) indeed satisfies the relevant set of saddle-point inequalities  $J(\gamma_1, \gamma_2) < J(\gamma_1, \gamma_2^*) < J(\gamma_1^*, \gamma_2)$  and thus constitutes a saddle-point solution within the class of mixed strategies. However, the class of mixed strategies is, in general, an unnecessarily large class for the second-acting player. In the extensive form of Fig. 2.6, for example, since P2 can tell exactly whether P1 has played R or not, there is no reason why he should mix between his possible actions in case of  $u_1 = R$  and his actions otherwise. That is, P2 can do equally well by considering probability distributions only on the alternatives belonging to the information set including two nodes, and (2.33b) is, in fact, that kind of a strategy. Such strategies are known as behavioral strategies, a precise definition of which is given below.

**Definition 2.99** For a given two-person zero-sum finite game in extensive form, let  $Y_{v_1}$  denote the set of all probability distributions on  $U_1$ , where the latter is the set of all alternatives of P1 at the nodes belonging to the information set  $\mathcal{I}_1$ . Analogously, let  $Z_{v_2}$  denote the set of all probability distributions on  $U_2$ . Further define  $Y = \bigcup_{\mathcal{I}_1} Y_{v_1}$ ,  $Z = \bigcup_{\mathcal{I}_2} Z_{v_2}$ . Then, a behavioral strategy  $\gamma_1$  for P1 is a mapping from the class of all his information sets ( $\mathcal{I}_1$ ) into  $Y$ , assigning one element in  $Y$  for each set in  $\mathcal{I}_1$ , such that  $\gamma_1(\mathcal{I}_1) \in Y_{v_1}$  for each  $\mathcal{I}_1 \in \mathcal{I}_1$ . A typical behavioral strategy  $\gamma_2$

for  $P_2$  is defined, analogously, as a restricted mapping from  $N_2$  into  $Z$ . The set of all behavioral strategies for  $P_i$  is called his behavioral strategy set, and it is denoted by  $P$ . Remark 2.5 Every behavioral strategy is a mixed strategy, but every mixed strategy is not necessarily a behavioral strategy, unless the player has a single information set.  $\square$  For games in extensive form, the concept of a saddle-point equilibrium in behavioral strategies can now be introduced as in Def. 2.8, but with some slight obvious modifications. If  $J(\gamma_1, \gamma_2)$  denotes the average outcome of the game resulting from the strategy pair  $\{\gamma_1 \in T_1, \gamma_2 \in T_2\}$ , then we have the following. Definition 2.10 A pair of strategies  $\{\gamma_1 \in T_1, \gamma_2 \in T_2\}$  is said to constitute a saddle point in behavioral strategies for a zero-sum game in extensive form, if the set of inequalities  $J(\gamma_1, \gamma_2) < J(\gamma_1, \gamma_2^*) < J(\gamma_1^*, \gamma_2)$  (2-34) This definition is valid not only for single-act games, but also for multi-act games which are treated in section 2.5.

TWO-PERSON ZERO-SUM FINITE GAMES 45 is satisfied for all  $\gamma_1 \in T_1, \gamma_2 \in T_2$ . The quantity  $J(\gamma_1^*, \gamma_2^*)$  is known as the saddle-point value of the game in behavioral strategies. For a given single-act zero-sum game in extensive form, derivation of a behavioral-strategy saddle-point solution basically follows the four steps outlined earlier for pure-strategy saddle points, with the only difference being that this time mixed equilibrium solutions of the matrix games are also allowed in this process. The reader should note that this routine extension has already been illustrated within the context of the extensive form of Fig. 2.6. Since every matrix game admits a saddle point in mixed strategies (cf. Corollary 2.3), and further, since the above derivation involves only a finite number of comparisons, it readily follows that every single-act zero-sum game in extensive form admits a saddle point in behavioral (or, equivalently, mixed) strategies. We thus conclude this section with a precise statement of this property, and by recording another immediate feature of the saddle-point solution in extensive games. Corollary 2.4 In every single-act two-person zero-sum finite game, (i) there exists at least one saddle point in behavioral strategies, (ii) if there exist more than one saddle-point solution in behavioral strategies, then they possess the ordered interchangeability property, and (iii) every saddle point in behavioral strategies is also a saddle point in the larger class of mixed strategies. We should note that a finite single-act game in extensive form might also admit a mixed-strategy saddle point (which is not behavioral) in addition to a behavioral saddle point. The corresponding average saddle-point values, however, will all be the same,

because of the interchangeability property, and therefore we may confine our attention only to behavioral strategies, particularly in view of Corollary 2.4(i).

2.5 Extensive Games: Multi-Act Games Zero-sum games in which at least one player is allowed to act more than once and with possibly different information sets at each level of play, are known as multi-act zero-sum games.<sup>10</sup> In the study of extensive forms of such games, and in accordance with Def. 2.5, we consider the case when the number of alternatives available to a player at each information set is finite. This leads to a finite tree structure, incorporating possibly different information sets for each <sup>10</sup>Such games are also referred to as "multi-stage zero-sum games" in the literature. They may be considered also as "dynamic games", since the information sets of a player do, in general, have a dynamic character, providing information concerning the past actions of the players. In this respect, the single-act games of the previous section may also be referred to as "dynamic games" if the second-acting player has at least two information sets.

46 T. BAŞAR AND G. J. OLSDER player at each level of play. A (pure) strategy of a player is again defined as a restricted mapping from the collection of his information sets to the finite set of all his alternatives, exactly as in Def. 2.6, and the concept of a saddle-point solution then follows the lines of Def. 2.8. This general framework, however, suppresses the dynamic nature of the decision process and is not constructive as far as the saddle-point solution is concerned. An alternative to this set-up exists, which brings out the dynamic nature of the problem, for an important class of multi-act zero-sum games known as "feedback games", a precise definition of which is given below.

Definition 2.11 A multi-act two-person zero-sum game in extensive form is called a two-person zero-sum feedback game in extensive form, if (i) at the time of his action, each player has perfect information concerning the current level of play, i.e., no information set contains nodes of the tree belonging to different levels of play, (ii) information sets of the first-acting player at every level of play are singletons, and the information sets of the second-acting player at every level of play are such that none of them include nodes corresponding to branches emanating from two or more different information sets of the other player, i.e., each player knows the state of the game at every level of play.

Figure 2.7: Two multi-act zero-sum games which are not feedback games. Figure 2.7 displays two multi-act zero-sum games in extensive form, which are not feedback games, since the first one violates (i)



above, and the second one violates (ii). The one displayed in Fig. 2.8, however, is a feedback game. Now, let the number of levels of play in a zero-sum feedback game be  $K$ . Then, a typical strategy  $Y$  of  $P_i$  in such a game can be viewed as composed of  $K$  components  $(\gamma_j, \dots, \gamma_j)$ , where  $\gamma_j$  stands for the corresponding strategy of  $P_i$  at his  $j$ 'th level of action. Moreover, because of the nature of a feedback game,  $\gamma_j$  can be taken, without any loss of generality, to have as its domain only those information sets of  $P_i$  which pertain to the  $j$ th level of play. Let us denote the collection of all such strategies for  $P_i$  at level  $j$  by  $r_j$ . Then, we can rewrite the saddle-point inequality (2.30), for a feedback game in extensive

## TWO-PERSON ZERO-SUM FINITE GAMES 47 1023230-12134102-2

Figure 2.8: A zero-sum feedback game in extensive form, as  $A \gg \dots$ ,  $1.2^* 2, N (2-35) \gamma_j \in r_j, i = 1, 2; j = 1, \dots, K$ . Such a decomposition of the strategies of the players now allows a recursive procedure for the determination of the saddle-point solution  $\{\gamma_1^*, \gamma_2^*\}$  of the feedback game, if we impose some further restrictions on the pair  $\{\gamma_1^*, \gamma_2^*\}$ . Namely, let the pair  $\{\gamma_1^*, \gamma_2^*\}$  also satisfy (recursively) the following set of  $K$  pairs of inequalities for  $i = 1, 2; j = 1, \dots, K$ :

$$\gamma_j^* \in r_j, i = 1, 2; j = 1, \dots, K. \quad (2.36)$$

Definition 2.12 For a two-person zero-sum feedback game in extensive form with  $K$  levels of play, let  $\{\gamma_1^*, \gamma_2^*\}$  be a pair of strategies satisfying (2.35) and (2.36) for all  $\gamma_j \in V, i = 1, 2; j = 1, \dots, K$ . Then,  $\{\gamma_1^*, \gamma_2^*\}$  is said to constitute a pair of (pure) feedback saddle-point strategies for the feedback game. Proposition 2.3 Every pair  $\{\gamma_1^*, \gamma_2^*\}$  that satisfies the set of inequalities (2.36) also satisfies the pair of saddle-point inequalities (2.35). Hence, the requirement for satisfaction of (2.35) in Def. 2.12 is redundant, and it can be deleted without any loss of generality.

48 T. BAŞAR AND G. J. OLSDER Proof. Replace  $\gamma_{i,j}^*$  on the LHS inequalities of the set (2.36) by  $\gamma_{i,j}^* \in r_j$  respectively. Then, the LHS inequalities of (2.36) read  $J\{1, \dots, 1, 1, \dots, n, 2K\} < J\{1^*, \dots, 1, 1, \dots, 1, 2K\}, J(j^*, \dots, 1, 1, \dots, 1, 2K) < J(h^*, \dots, 1, 1, \dots, 1, 2K)$  for all  $\gamma_j \in V, j = 1, \dots, K$ . But this set of inequalities reduces to the single inequality

$J(711^*, \dots, 7^*; 7^?, \dots, 71:)^{<^{\wedge}(7i1^*, \dots, 7j<: *; 7r, \dots, 7^{\wedge})}$ , since they characterize the recursive dynamic programming conditions for a one-person maximization problem (Bellman (1957)). This then establishes the validity of the LHS inequality of (2.35) under satisfaction of (2.36). Validity of the RHS inequality of (2.35) can likewise be verified by considering the RHS inequalities of (2.36).

D The appealing feature of a feedback saddle-point solution is that it can be computed recursively, by solving a number of static games at each level of play. Details of this recursive procedure readily follow from the set of inequalities (2.36) and the specific structure of the information sets of the game, as it is described below. A recursive procedure to determine the (pure) feedback saddle-point strategies of a feedback game

- (1) Starting at the last level of play,  $K$ , solve each single-act game corresponding to the information sets of the first-acting player at that level  $K$ . The resulting (pure) saddle-point strategies are the saddle-point strategies  $\{7^{\wedge*}, 7j\}$  sought for the feedback game at level  $K$ . Record the value of each single-act game corresponding to each of the information sets of the first-acting player, and assign these values to the corresponding nodes (states) of the extensive form at level  $K$ .
- (2) Cross out the  $K$ th level of play, and consider the resulting  $(K - 1)$  level feedback game. Now solve the last level single-act games of that extensive form, with the resulting saddle-point strategies denoted by  $\{7k-i, 7k-i\}$  and repeat the remaining deeds of step 1 with  $K$  replaced by  $K - 1$ .
- ( $K$ ) Cross out the second level of play, and consider the resulting single-act game in extensive form. Denote its saddle-point strategies by  $\{7f^*, 7i^*\}$ ,

TWO-PERSON ZERO-SUM FINITE GAMES 49 and its value by  $V$ . Then, the original feedback game in extensive form admits a saddle-point solution  $\{(7J^*, \dots, 7^{\wedge*}), (7^?^*, \dots, 7^{\wedge})\}$  and has value  $J^* = V$ .

Example 2.2 As an illustration of the above procedure, let us consider the 2-level feedback game whose extensive form is depicted in Fig. 2.8. At the last (second) level, there are four single-act games to be solved, which are all equivalent to matrix games. From left to right, we have the normal forms

$T^*$	$0$	$2$	$3$	$n^*$	$3$	$-1$	$F^{\wedge T}$	$1$	$3$	$4$
$1$	$2$	$0$	$-2$							

with the encircled entries indicating the location of the saddle-point solution in each case. The saddle-point strategies  $72^*$  and  $7I^*$  are thus given by

$1? \text{ til) } R \text{ L if } u \setminus = L, \text{ otherwise, } R, \text{ if } [ril) \text{ (2.37a)}$  where  $u \setminus$  denotes the choice of  $P_i$  at level 1. It is noteworthy that the saddle-point strategy of  $P_2$  is a constant at this level of play. Now, crossing out the second level of play, but retaining the encircled quantities in the four matrix games as costs attached to

the four nodes, we end up with the equivalent single-act game whose normal form is admitting the unique saddle-point solution  $v = L$ ,  $\tau_i = L$ , (2.37b) and with the value being  $V = 1$ , which is also the value of the original 2-level feedback game. The feedback saddle-point strategies of this feedback game are now given by (2.37a)-(2.37b), with the actual play dictated by them being  $U_i = U_0 = U_i = L$ . (2.38) D

50 T. BAŞAR AND G. J. OLSDER If a feedback game in extensive form does not admit a feedback saddle point in pure strategies, then it could still admit a saddle point in accordance with inequalities (2.35). If that is not possible either, then one has to extend the strategy spaces of the players so as to include also behavioral strategies. Definition of a behavioral strategy in multi-act games is precisely the one given by Def. 2.9 and the relevant saddle-point inequality is still (2.34), of course under the right kind of interpretation. Now, in a zero-sum feedback game with  $K$  levels of play, a typical behavioral strategy  $\tau_j$  of  $P_i$  can also be viewed as composed of  $K$  components  $(\tau_j^1, \dots, \tau_j^K)$ , where  $\tau_j^i$  stands for the corresponding behavioral strategy of  $P_i$  at his  $j$ th level of play—a feature that follows readily from Defs. 2.9 and 2.11. By the same token,  $\tau_j$  can be taken, without any loss of generality, to have as its domain only those information sets of  $P_i$ , which pertain to the  $j$ th level of play. Let us now denote the collection of all such behavioral strategies for  $P_i$  at level  $j$  by  $T_j$ . Then, since the saddle-point inequality (2.34) corresponding to a  $K$ -level feedback game can equivalently be written as (2.39)  $\sum_{j=1}^K \tau_j^i \cdot v_j \leq v$ ,  $\forall \tau_j^i \in T_j, i = 1, 2; j = 1, \dots, K$ , it follows that a recursive procedure can be devised to determine the behavioral saddle-point strategies if we further require satisfaction of a set of inequalities analogous to (2.36). Such a recursive procedure basically follows the  $K$  steps outlined in this section within the context of pure feedback saddle-point solutions, with the only difference being that now also behavioral strategies are allowed in the equilibrium solutions of the single-act games involved. Since this is a routine extension, we do not provide details of this recursive procedure here, but only illustrate it in the sequel in Example 2.3. It is worth noting at this point that since every single-act two-person zero-sum game admits a saddle-point solution in behavioral strategies (cf. Corollary 2.4), every feedback game will also admit a saddle-point solution in behavioral strategies. More precisely, we have the following proposition. Proposition 2.4 Every two-person zero-sum feedback game, which has an extensive form comprised of a finite number of

branches, admits a saddle point in behavioral strategies. Example 2.3 As an illustration of the derivation of behavioral saddle point(s) in feedback games, let us reconsider the game of Fig. 2.8 with two modifications: The outcomes of the game corresponding to the paths  $u^1 = R, u^2 = L, u^3 = L, u^4 = L$  and  $u^1 = R, u^2 = L, u^3 = L, u^4 = R$ , respectively, are now taken as 0 and  $-1$ , respectively. Then, the four single-act games to be solved at the last (second) level are equivalent to the matrix games (respectively, from left to right) which all admit pure-strategy solutions, as indicated.

TWO-PERSON ZERO-SUM FINITE GAMES 51 The saddle-point strategies at this level are, in fact, as given by (2.37a). Now, crossing out the second level of play, but retaining the encircled quantities in the four matrix games as costs attached to the four nodes, we end up with the equivalent single act game whose normal form admits a mixed saddle-point solution:  $u^1 = L$  w.p.  $1/2$ ,  $1/2$ ,  $u^2 = R$  w.p.  $1/2$ ,  $1/2$ . (2.40) Furthermore, the average saddle-point value is  $V_m = 1/2$ . Since, within the class of behavioral strategies at level 2, the pair (2.37a) can be written as  $u^1 = L, u^2 = R$  w.p.  $1/2$ ,  $1/2$  otherwise  $u^1 = R, u^2 = L$  w.p.  $1/2$ ,  $1/2$ . (2.41) (2.40) and (2.41) now constitute a set of behavioral saddle-point strategies for the feedback game under consideration, leading to a behavioral saddle-point value of  $J^* = 1/2$ . D For multi-act games which are not feedback games, neither Prop. 2.4 nor the recursive procedures discussed in this section are, in general, applicable, and there is no general procedure that would aid in the derivation of pure or behavioral saddle-point solutions, even if they exist. If we allow for mixed-strategy solutions, however, it directly follows from Corollary 2.3 that an equilibrium solution exists, since every multi-act zero-sum finite game in extensive form has an equivalent normal form which is basically a matrix game with a finite number of rows and columns. More precisely, we have the following. Proposition 2.5 Every two-person zero-sum multi-act finite game in extensive form admits a saddle-point solution in mixed strategies.

52 T. BAŞAR AND G. J. OLSDER The following example now illustrates derivation of mixed-strategy saddle-point equilibria in multi-act games, which basically involves solution of an appropriate matrix game. Example 2.4 Let us reconsider the extensive form of Fig. 2.8, but with a single information set for each player at each level of play, as depicted in Fig. 2.9.

2 Figure 2.9: A zero-sum multi-act game that does not admit a behavioral saddle-point solution. Each player has four possible ordered choices (LL,LR,RL,RR), and thus the equivalent normal form of the matrix game is  $P_1$   $P_2$  which admits the unique mixed saddle-point solution  $LL$   $LR$   $RL$   $RR$   $LL$   $1/2$   $2/3$   $LR$   $0$   $3/4$   $RL$   $2/0$   $1/2$   $RR$   $3$   $-1$   $0$   $-2$   $f$  - {  $LL$   $RR$  w.p.  $3/5$ ,  $2/5$ , (2.42a) (2.42b)  $LL$  w.p.  $4/5$ ,  $RR$  w.p.  $1/5$  with the saddle-point value in mixed strategies being  $V^* = 7/5$ . (The reader can check this result by utilizing Thm. 2.5.)  $\square$  Even though every zero-sum multi-act game in extensive form admits a saddle point in mixed strategies, such a result is no longer valid within the class of behavioral strategies, unless the mixed saddle-point strategies happen to be behavioral strategies. A precise version of this property is given below in Prop. 2.6. Proposition 2.6 Every saddle-point equilibrium of a zero-sum multi-act game in behavioral strategies also constitutes a saddle-point equilibrium in the larger class of mixed strategies for both players.

TWO-PERSON ZERO-SUM FINITE GAMES 53 Proof. Assume, to the contrary, that the behavioral saddle-point value ( $J^*$ ) is not equal to the mixed saddle-point value ( $J_m$ ), where the latter always exists by Prop. 2.5, and consider the case  $J^* < J_m$ . This implies that  $P_1$  does better with his behavioral saddle-point strategy than with his mixed saddle-point strategy. But this is not possible, since the set of mixed strategies is a much larger class. Hence, we can only have  $J^* > J_m$ . Now, repeating the same argument on this strict inequality from  $P_2$ 's point of view, it follows that only the equality  $J^* = J_m$  can hold.  $\square$  Proposition 2.6 now suggests a method for derivation of behavioral saddle-point strategies of zero-sum multi-act games in extensive form. First, determine all mixed saddle-point strategies of the equivalent normal form, and then investigate whether there exists, in that solution set, a pair of behavioral strategies. If we apply this method on the extensive form of Example 2.4, we first note that it admits a unique saddle-point solution in mixed strategies, which is the one given by (2.42a)-(2.42b). Now, since every behavioral strategy for  $P_1$  is in the form { $LL$  w.p.  $2/12/2$ ,  $LR$  w.p.  $2/1(1-2/2)$ ,  $RL$  w.p.  $2/2(1-2/1)$ ,  $RR$  w.p.  $(1-2/i)(1-3/2)$  for some  $0 < 2/1 < 1$ ,  $0 < j/2 < 1$ , and every behavioral strategy for  $P_2$  is given by  $LL$  W.p.  $2i/22$ ,  $LR$ . w.p.  $z1(1-z2)$ ,  $RL$  w.p.  $z2(1 - -zi)$ ,  $RR$  w.p.  $(i-Z1)(i-Z2)$  for some  $0 < z \setminus < 1$ ,  $0 < zi < 1$ , it follows (by picking  $2/1 = 1$ ,  $3/2 = 3/5$ ) that (2.42a) is indeed a behavioral strategy for  $P_1$ ; however, (2.42b) is not a behavioral strategy for  $P_2$ . Hence the conclusion is that the extensive form of Fig. 2.9 does not admit a saddle point

in behavioral strategies, from which we deduce the following corollary to Prop. 2.6. Corollary 2.5 A zero-sum multi-act game does not necessarily admit a saddle-point equilibrium, in behavioral strategies, unless it is a feedback game. We now introduce a special class of zero-sum multi-act games known as open-loop games, in which the players do not acquire any dynamic information throughout the decision process, and they only know the level of play that corresponds to their action. More precisely, see below. Note that  $j/i$  and  $2/2$  (and also  $z$  and  $22$ , in the sequel) are picked independently, and hence do not necessarily add up to 1.  $72 = <$

54 T. BAŞAR AND G. J. OLSDER Definition 2.13 A multi-act game is said to be an open-loop game, if, at each level of play, each player has a single information set. The multi-act game whose extensive form is depicted in Fig. 2.9 is an open-loop game. Such games can be viewed as one extreme class of multi-act games in which both players are completely ignorant about the evolution of the decision process. Feedback games, on the other hand, constitute another extreme case (with regard to the structure of the information sets) in which both players have full knowledge of their past actions. Hence, the open-loop and feedback versions of the same multi-act game could admit different saddle-point solutions in pure, behavioral or mixed strategies. A comparison of Examples 2.2 and 2.4, in fact, provides a validation of the possibility of such a situation: the extra information embodied in the feedback formulation actually helps P1 in that case. It is of course possible to devise examples in which extra information in the above sense instead makes P2 better off or does not change the value of the game at all. A precise condition for the latter property to hold is given below in Prop. 2.7 whose proof is in the same spirit as that of Prop. 2.6 and is thus left as an exercise for the reader. Proposition 2.7 A zero-sum open-loop game in extensive form admits a (pure-strategy) saddle-point solution if, and only if, its feedback version admits constant saddle-point strategies at each level of play. 2 Prior and delayed commitment models Since there is no extra information available to the players throughout the duration of an open-loop game, the players can decide on their actions at the very beginning, and then there is no real incentive for them to renege during the actual play of the game. Such games in which decisions are made at the outset, with no incentive to deviate from them later, are known as "prior commitment" games. Mixed saddle-point strategies can then be considered to constitute a reasonable equilibrium solution within such a

framework. Feedback games, on the other hand, are of "delayed commitment" type, since each player could wait until he finds out at what information set he really is, and only then announces his action. This then implies that mixed strategies are not the "right type" of strategies to be considered for such games, but behavioral strategies are. Since feedback games always admit a saddle point in behavioral strategies (cf. Prop. 2.3), equilibrium is again well defined for this class of games. There are also other classes of multi-act games of the delayed commitment type, which might admit a saddle point in behavioral strategies. One such class of multi-act zero-sum games are those in which players recall their past actions but are ignorant about the past actions of their opponent. As an illustrative example, let us reconsider the extensive form of Fig. 2.9, but under the present set-up. The information sets of the players would then look as displayed in 12It is implicit here that these constant strategies might differ from one level to another and that they might not be feedback saddle-point strategies, i.e., they might not satisfy (2.36).

TWO-PERSON ZERO-SUM FINITE GAMES 55 023230-12134102-2 Figure 2.10: A multi-act game of the delayed commitment type that admits a behavioral saddle-point solution. Fig. 2.10. Now, permissible behavioral strategies of the players in this of the form  $y_i / L$  w.p.  $y_i$   $7_i R$  w.p.  $1-2/i$ ,  $7_2 > i$ )  $L$   $R$   $L$   $R$  w.p. w.p. w.p. w.p.  $2/2$   $1-2/2$   $2/3$   $1-2/3$  if  $u \setminus = L$ , if  $u \setminus = R$ , it  $L$  w.p.  $z \setminus$ ,  $R$  w.p.  $1 - z \setminus$ ,  $-C2/2$   $7^*(\ll f) = < L$  w.p.  $z^2$   $R$  w.p.  $1 - z^2$   $L$  w.p.  $Z^3$   $i?$  w.p.  $1 - 2^3$  if  $M_i - L$ , if  $u \setminus = \#$ , where  $0 < t/i < 1$ ,  $0 < Z_i < 1$ ,  $i = 1,2,3$ , and the average outcome of corresponding to these strategies can be determined to be game are (2.43a) (2.43b) (2.43c) (2.43d) the game  $J\{V,Y\} = zlz^2\{2y_1y_2 - 2y_1y_i - 2y_i - 1\} + Z_i(52/i2/3 - 2y_x - lyxy^2 - 5t/3 + 6) + z^3(z_i - 1)(2j/i2/2 + 3j/i + 3t/3 - 32/!j/3 - 4) + 42/12/2 + 2/1 + 22/3 - 2/12/3 - 2 = F\{y_1,y_2,y_3;z_1,z_2,zz\}$ . The reader can now check that the values  $(y \setminus = \setminus, y_2 - 3/5, 2/3 = 0; z \setminus 4/5, z^2 - 1, \text{£}3 = 0)$  satisfy the inequalities  $\wedge(2/1, 2/2, 2/3; *1.*2, \wedge^3) < \wedge(2/1, 2/2, 2/3; z_l, z^2 \wedge^3) < F(y_i, \sqrt{2,2/3}; \wedge^r, 4 > 4)$  (2.44)

56 T. BAŞAR AND G. J. OLSDER for all  $0 < j/i < 1$ ,  $0 < 2j < 1$ ,  $i = 1,2,3$ ; and hence, when they are used in (2.43a)-(2.43d); the resulting behavioral strategies constitute a saddle-point solution for the multi-act game under consideration, with the average value being  $7/5$ . We thus observe that a modified version of the extensive form of Fig. 2.9, in which the players recall their past actions, does admit a saddle-point solution in behavioral strategies.

Derivation of the behavioral saddle-point solutions of multi-act games of the delayed commitment type involve, in general, solution of a pair of saddle-point inequalities of the type (2.44). For the present case (i.e., for the specific example of Fig. 2.10) it so happens that  $F$  admits a saddle-point solution which, in turn, yields the behavioral strategies sought; however, in general it is not guaranteed that it will admit a well-defined solution. The kernel  $F$  will always be a continuous function of its arguments which belong to a convex compact set (in fact a simplex), but this is not sufficient for existence of a saddle point for  $F$  (see Section 4.3). The implication, therefore, is that there will exist multi-act games of the delayed commitment type which do not admit a saddle-point solution in pure or behavioral strategies. Randomized strategies

As observed above, and also in Example 2.4, not every multi-act finite game would admit a behavioral saddle-point solution, but whenever it exists, one method for obtaining the behavioral saddle-point policies would be first to construct the average cost of the game on the set of behavioral strategies (such as the kernel  $F$  defined prior to (2.44)), and then find the saddle point of that kernel in pure strategies. The (pure) strategies in this case are in fact the probability weights attached to different actions at each information set, which belong to closed and bounded subsets of finite dimensional spaces. Let us denote these subsets for  $P_1$  and  $P_2$  by  $Y$  and  $Z$ , respectively. For the game of Example 2.4, for instance, both  $Y$  and  $Z$  are positive unit squares, whereas for the multi-act game of Fig. 2.10, they are positive unit cubes. In all cases, the kernel  $F$  will be continuous on  $Y \times Z$ , and  $Y$  and  $Z$  will further be convex. We shall study the derivation of saddle-point equilibria of such continuous-kernel games in Chapter 4, where we will see that  $F$  does not necessarily admit a pure-strategy saddle point on  $Y \times Z$  unless it is convex-concave. Clearly, neither  $F$  defined earlier in conjunction with the multi-act game of Fig. 2.10, nor the one that could be constructed for the game of Example 2.4, is convex-concave,<sup>13</sup> which is one of the reasons why these two games do not admit behavioral saddle points. The theory of Chapter 4, and particularly of Section 4.3, will tell us, however, that such continuous-kernel games defined on closed-bounded convex sets will admit mixed saddle-point equilibria, where mixed strategies in this case are probability distributions on  $Y$  and  $Z$ . For present application, since the elements of  $Y$  and  $Z$  already correspond to behavioral strategies for the original finite game <sup>13</sup>We leave it as an exercise for the reader to construct the kernel  $F$  for the game of Example 2.4, and show that it is not convex-concave (i.e., convex over  $Y$  and concave over  $Z$ ).



TWO-PERSON ZERO-SUM FINITE GAMES 57 (themselves being special types of mixed strategies), and to avoid confusion, the further mixed strategies defined on  $Y$  and  $Z$  are often called randomized strategies (for the original finite game). Hence, the message here is that every finite multi-act game admits a saddle point in randomized strategies—a full verification of which is postponed until Chapter 4. Some further discussion on randomized strategies, in this chapter, can be found in subsection 2.7.1.

### 2.6 Zero-Sum Games with Chance Moves

In this section, we briefly discuss the class of two-person zero-sum finite games which also incorporate chance moves. In such decision problems, the final outcome is determined not only by the decisions (actions) of the two players, but also by the outcome of a chance mechanism whose odds between different alternatives are known a priori by both players. One can also view this situation as a three-player game wherein the third player, commonly known as "nature", has a fixed mixed strategy. Figure 2.11: A zero-sum single-act game in extensive form that incorporates a chance move. As an illustration of such a zero-sum game, let us consider the tree structure of Fig. 2.11, where  $N$  (nature) picks  $L$  w.p.  $1/3$  and  $R$  w.p.  $2/3$ . Not knowing the realized outcome of the chance mechanism, the players each decide on whether to play  $L$  or  $R$ . Let us now assume that such a game is played a sufficiently large number of times and the final outcome is determined as the arithmetic mean of the outcomes of the individual plays. In each of these plays, the matrix of possible outcomes is either

$$P_1 \begin{matrix} & L & R \end{matrix} \begin{matrix} L \\ R \end{matrix} \begin{matrix} 3 & 1 \\ 2 & 4 \end{matrix} \quad (2.45a)$$

or

$$P_2 \begin{matrix} & L & R \end{matrix} \begin{matrix} L \\ R \end{matrix} \begin{matrix} 0 & 2 \\ 4 & 1 \end{matrix} \quad (2.45b)$$

$L$   $R$  the former occurring w.p.  $1/3$  and the latter w.p.  $2/3$ . Each of these matrix games admits a pure-strategy saddle-point solution as indicated. If  $P_1$ 's choice were known to both players, then  $P_2$  would always play  $R$ , and  $P_1$  would play  $L$  if  $P_2$ 's choice is  $L$  and play  $R$  otherwise. However, since  $P_2$ 's choice is not known a priori, and since the equilibrium strategies to be adopted by the players cannot change from one individual play to another, the matrix game of real interest is the one whose entries correspond to the possible average outcomes of a sufficiently large number of such games. This matrix game can readily be obtained from (2.45a) and (2.45b) to be

$$P_2 \begin{matrix} & L & R \end{matrix} \begin{matrix} L \\ R \end{matrix} \begin{matrix} 1 & 3 \\ 2 & 2 \end{matrix} \quad (2.46a)$$

$L$   $R$  which admits, as indicated, the unique saddle-point solution (2.46b) with an average outcome of  $1$ . This solution is, in fact, the only reasonable equilibrium solution of the zero-sum single-act game of Fig. 2.11, since it also uniquely satisfies the pair of saddle-point inequalities

$$v_1 \leq \min_j \max_i \{ a_{ij} \} \leq \max_i \min_j \{ a_{ij} \} \leq v_2 \quad (2.47)$$

for

all  $71 \in T1, 72 \in T2$ , where  $r^2 = \{1, r\}$  and  $\hat{v}(71, 72)$  denotes the expected (average) outcome of the game corresponding to a pair of permissible strategies  $\{71, 72\}$  with the expectation operation taken with respect to the probability weights assigned a priori to different choices of  $N$ . For this problem, the correspondences between  $J$  and different choices of  $(71, 72)$  are in fact displayed as the entries of matrix (2.46a). It is noteworthy that  $P2$ 's equilibrium strategy is  $L$  for the extensive form of Fig. 2.11 (as discussed above), whereas if  $N$ 's choice were known to both players (Fig. 2.12), he would have chosen the constant strategy  $R$ , as shown earlier. Such a feature might, at first sight, seem to be counter-intuitive—after all, when his information sets can distinguish between different choices of  $N$ ,  $P2$  disregards  $N$ 's choice and adopts a constant strategy ( $R$ ), whereas in the case when this information is not available to him, he plays quite a different one. Such an argument is misleading, however, because of the following reason:

TWO-PERSON ZERO-SUM FINITE GAMES 59

Figure 2.12: A modified version of the zero-sum game of Fig. 2.11, with both players having access to nature's actual choice. In the tree structure of Fig. 2.12,  $P1$ 's equilibrium strategy is not a constant but is a variant of  $N$ 's choice; whereas in Fig. 2.11 he is confined to constant strategies, and hence the corresponding zero-sum game is quite a different one. Consequently there is no reason for  $P2$ 's equilibrium strategies to be the same in both games, even if they are constants. Remark 2.6 It should be noted that the saddle-point solution of the zero-sum game of Fig. 2.12 also satisfies a pair of inequalities similar to (2.47), but this time each player has four permissible strategies, i.e.,  $V1$  is comprised of four elements. It directly follows from (2.45a)-(2.45b) that the average saddle-point value, in this case, is  $(1/3)(1) + (2/3)(1/4) = 1/2$ , which is less than the average equilibrium value of the zero-sum game of Fig. 2.11. Hence, the increase in information with regard to the action of  $N$  helps  $P1$  to reduce his average losses but works against  $P2$ .  $\square$  The zero-sum games of both Fig. 2.11 and Fig. 2.12 are in extensive form, but they do not fit the framework of Def. 2.5 since the possibility of chance moves was excluded there. We now provide below a more general definition of a zero-sum game in extensive form that also accounts for chance moves. Definition 2.14 An extensive form of a two-person zero-sum finite game that also incorporates chance moves is a tree structure with (i) a specific vertex indicating the starting point of the game, (ii) a pay-off function assigning a real number to each terminal vertex of the tree,

which determines the pay-off (respectively, loss) to P2 (respectively,  $P_1$ ) for each possible set of actions of the players together with the possible choices of nature, (in) a partition of the nodes of the tree into three player sets (to be denoted by  $N_1$  and  $N_2$  for  $P_1$  and  $P_2$ , respectively, and by  $N^\circ$  for nature),

60 T. BAŞAR AND G. J. OLSDER (iv) a probability distribution, defined at each node of  $N^\circ$ , among the immediate branches (alternatives) emanating from this node, (v) a subpartition of each player set  $J^i$  into information sets  $\{I_i^j\}$  such that the same number of immediate branches emanates from every node belonging to the same information set, and no node follows another node in the same information set. This definition of a zero-sum game in extensive form clearly covers both single-act and multi-act games, and nature's chance moves can also be at intermediate levels of play. For such a game, let us denote the strategy sets of  $P_1$  and  $P_2$  again as  $T_1$  and  $T_2$ , respectively. Then, for each pair  $\{t_1 \in T_1, t_2 \in T_2\}$  the pay-off function,  $J(t_1, t_2)$ , is a random variable, and hence the real quantity of interest is its expected value which we denote as  $E J(t_1, t_2)$ . The saddle-point equilibrium can now be introduced precisely in the same fashion as in Def. 2.8, with only  $J$  replaced by  $E J$ . Since this formulation converts the extensive form into an equivalent normal form, and since both  $T_1$  and  $T_2$  are finite sets, it follows that the saddle-point equilibrium solution of such a game problem can be obtained by basically solving a matrix game whose entries correspond to the possible values of  $E J(t_1, t_2)$ . This has actually been the procedure followed earlier in this section in solving the single-act game of Fig. 2.11, and it readily extends also to multi-act games. When a pure-strategy saddle point does not exist, it is possible to extend the strategy sets so as to include behavioral or mixed strategies and to seek equilibria in these larger classes of strategies. Proposition 2.5 clearly also holds for the class of zero-sum games covered by Def. 2.14, and so does Prop. 2.6. Existence of a saddle point in behavioral strategies is again not guaranteed, and even if such a saddle point exists there is no systematic way to obtain it, unless the players have access to nature's actions and the game possesses a feedback structure (as in Def. 2.11). For games of this specific structure, one can use an appropriate extension of the recursive derivation developed in Section 2.5 for feedback games, which we do not further discuss here. Before concluding we should mention that the notions of prior and delayed commitment introduced in the previous section also fit within the framework of games with chance moves, and so does the concept of randomized strategies, with the extensions being

conceptually straightforward. 2.7 Two Extensions In this section we discuss two possible extensions of the extensive form description of finite zero-sum games—one where the restriction of (cf. Def. 2.5 and Def. 2.14) "no nodes following another node in the same information set" is dispensed with (see subsection 2.7.1), and another one where the tree structure, viewed as a directed graph, is allowed to have cycles (see subsection 2.7.2).

TWO-PERSON ZERO-SUM FINITE GAMES 61 Figure 2.13: A zero-sum game in extensive form in which the tree enters an information set twice. 2.7.1

Games with repeated decisions We start this subsection with a specific example. Example 2.5 Consider the zero-sum game in extensive form shown in Fig. 2.13. P1 has two information sets, one of which (77J) contains two nodes from one path in the tree; P2 has a single information set. If the probability of going right from 77} is given by  $2/1$ , and going left by  $1 - 2/1$ , and if  $2/2$  and  $z\}$  are the corresponding probabilities for  $r)\}$  and  $rj\}$ , respectively, then the behavioral normal form is given by the kernel  $14 6(1 - 2/0(1 - ZI) + 2/i(1 - zi) + 2y1z1(1 - 2/2) + 6yizi y2yi$ . (2.48) First minimizing this expression with respect to  $2/2 \in [0,1]$  leads to the kernel  $A J 6(1 - 1/0(1 - zi) + 2/i(1 - 2I) + 6(2/i)22i$  if  $t/i < 1/3$ , if  $2/i > 1/3$ ,  $K\{yi,z\} - 1 6(1 - yi)(1 - Zi) + yi(i - ZI) + 2yiZI$  with a corresponding minimizing  $2/2$  being  $2/2$  if  $2/1 > 1/3$ ,  $2/i < 1/3$ . Now, the function  $K$  above is to be minimized by P1 and maximized by P2. Toward this end consider the max-min problem:  $\max_2, \min_{yi} K(y\}, z\})$ , defined on the unit square; see Fig. 2.14 for a sketch of the function  $K$ . Some calculus leads to the result that this max-min is uniquely achieved for  $z\} = (3 - (170 + 40V6))/386 = 0.694$ . This behavioral strategy is optimal for P2. To obtain 14 When there are two (or more) nodes following each other in the same information set, a behavioral strategy is defined as one where the odds for different alternatives are the same at each node (since a player cannot differentiate between two nodes belonging to the same information set), but the random mechanism that generates the actions (moves) according to these odds is independent from one node to another. Hence, moves from different nodes belonging to the same information set can be considered as independent random variables with identical probability distributions.

62 T. BAŞAR AND G. J. OLSDER Figure 2.14: The function  $K$ . the optimal strategy of P1, consider the minimization problem  $\min^{\wedge} K\{y\}, 0)$ , where the minimum is attained at only two points:  $y\} = a = 5(1 - /?)/12/3 = .184$  and  $j/i$

$= 1$ . Hence it follows that PI should use two behavioral strategies, viz.  $j/(1) = (y_1 - a, y_2 = 1)$  and  $j/(2) = (z_1 = 1, z_2 = 0)$ , in an appropriate mix. To find this right "mix", consider the  $2 \times 2$  P2 PI 6 —  $5a - 1 - 6a^2 z = 0$   $z = 1 - 2/(D - 2)$  The solution to this game is given by the following: choose  $z = 1$  with probability  $3$  and  $z = 0$  with probability  $1 - 1/3$ ; choose  $y^1$  with probability  $1/(6a^2 + 5a - 7) = 0.17$  and  $y(2)$  with probability  $0.83$ . Thus it has been established that PI mixes between two behavioral strategies, the result of which is a randomized strategy.  $\square$  The preceding example has shown features which have also surfaced in the discussion on randomized strategies in Section 2.5. In both cases the starting point was a behavioral normal form which more generally has the form  $K \sum_{i=1}^m \sum_{j=1}^n c_{ij}(k) x_i^{j(k)} y_j^{n(k)}$  where the exponents  $i(k)$  and  $j_i(k)$  are nonnegative integers. This form is sometimes called a multinomial. It is to be minimized by PI by choosing a vector  $(t_1, \dots, y_m)$  on  $[0, 1]^m$  and to be maximized by P2 by choosing  $(z_1, \dots, z_n)$  on  $[0, 1]^n$ . It is fairly easy to see that an arbitrary game in extensive form, of the

**TWO-PERSON ZERO-SUM FINITE GAMES** 63 form described in this subsection, and without chance moves, has a behavioral normal form which is a multinomial. The converse is also true: every multinomial is the behavioral normal form of an appropriately defined game in extensive form, as shown by Alpern (1988). The theory of Chapter 4 will tell us that the zero-sum game just described in the behavioral normal form does indeed have a saddle point in mixed strategies. This means that the optimal solution of the underlying game in extensive form exists in the class of randomized strategies. In Alpern (1988), an even stronger result can be found: for the saddle-point solution of finite games in randomized strategies, each player needs to average over only a finite number of behavioral strategies. Counterparts of these results do exist for games in extensive form which also include chance moves. Since the basic features are the same as above, and no new difficulties emerge, we will not treat them separately here.

**2.7.2 Extensive forms with cycles** In this subsection we consider finite zero-sum dynamic games without chance moves, in which the underlying (directed) graph is no longer a finite tree; it will contain cycles, though the number of nodes remains finite. Moreover, it is assumed that the players have perfect information, i.e., each information set consists of a single node. If the players prefer to do so, such games can go on forever by moving and staying within one (or more) cycle(s). Thus infinite play is possible. It is of course possible to 'unfold' the underlying graph with cycles so as to obtain a

tree whose nodes are paths of the original game. However, the resulting tree will be infinite and no useful general results for such trees are known. Games with terminal pay-off We now first discuss the case when the game has only a terminal pay-off. Definition 2.15 A two-person zero-sum deterministic graph game with terminal pay-off consists of a finite directed graph of which the nodes are partitioned into three sets: (i) a specific vertex indicating the starting point of the game, (ii) a pay-off function assigning a real number to each terminal vertex of the graph, which determines the pay-off (respectively, loss) to P2 (respectively, P1) for each possible set of actions of the players, (iii) a partition of the nonterminal nodes of the graph into two sets (to be denoted by  $N_1$  and  $N_2$  for P1 and P2, respectively). If the players choose their actions in such a way that the game does not end in a finite number of steps (i.e., a terminal node is never reached), the pay-off will, by definition, be zero. A pure strategy for  $P_i$  in the current context is a rule that assigns to any path that ends in  $N_i$  one of its successor nodes deterministically.

64 T. BAŞAR AND G. J. OLSDER A stationary strategy is a pure strategy that only depends on the final node of the path. It is possible to devise an algorithm that yields a pair of stationary strategies which form a saddle-point solution for the game. To introduce such an algorithm, we assign a value  $V$  to each node  $x$  of the graph, with  $V(x)$  denoting the value of the game if node  $x$  is chosen as the starting point of the game. The set of nodes for which the value has already been calculated while running the algorithm is denoted by  $H$ ; note that  $H$  will be increasing as long as the algorithm is running, until the entire graph has been covered by  $H$ . Let us also introduce the notation  $a(x)$  to denote the set of all immediate successor nodes of node  $x$ . Then, the algorithm proceeds as follows. 1. For all terminal nodes, the function  $V(x)$  equals the pay-off if the game terminates at  $x$ ; the set of all terminal nodes is denoted by  $H$ . 2. Consider all  $x \in H$  for which  $e(x) \cap N_1 \neq \emptyset$ . The value for such an  $x$  is defined as  $V(x) = \min_{a \in a(x)} V(a)$ . The argument(s) for which this minimum is achieved is then part of the saddle-point strategy for P1. Increase  $H$  with new nodes for which a value has been obtained. Repeat this as long as new nodes  $x \in N_1$  are found such that  $a(x) \cap H \neq \emptyset$ . 3. Consider all  $x \in H$  for which  $a(x) \cap N_2 \neq \emptyset$ . The value for such an  $x$  is defined as  $V(x) = \max_{a \in a(x)} V(a)$ . The argument(s) for which this maximum is achieved is then part of the saddle-point strategy for P2. Increase  $H$  with new nodes for which a value has been obtained. Repeat this as long as new nodes  $x \in N_2$  can be found such that  $a(x) \cap H \neq \emptyset$ . 4. Repeat

steps 2 and 3 until no new nodes can be found. 5. Consider those  $x \in H$  for which  $a(x) \in N_1 \cap H \neq \emptyset$ . Define node  $y$  to be that node (or those nodes) of the set  $a(x) \cap H$  which has (have) the smallest value, i.e.,  $V(y) = \min_{z \in a(x) \cap H} V(z)$ . If  $V(y) < 0$ , then increase  $H$  by node  $x$  and define  $V(x) = V(y)$ . The arc between  $x$  and  $y$  then constitutes part of the saddle-point strategy for P1. (If  $V(y) > 0$ , then  $H$  is not increased: P1 prefers nontermination with value 0 over a positive pay-off.) Repeat this as long as  $H$  can be increased. 6. Consider those  $x \in H$  for which  $a(x) \cap N_2 \cap H \neq \emptyset$ . Define node  $y$  to be that node (or those nodes) of the set  $a(x) \cap H$  which has (have) the largest value, i.e.,  $V(y) = \max_{z \in a(x) \cap H} V(z)$ . If  $V(y) > 0$ , then increase  $H$  by node  $x$  and define  $V(x) = V(y)$ . The arc between  $x$  and  $y$  then constitutes part of the saddle-point strategy for P2. (If  $V(y) < 0$ , then  $H$  is not increased: P2 prefers nontermination with value 0 to a negative pay-off.) Repeat this as long as  $H$  can be increased. 7. Repeat steps 5 and 6 until no new nodes can be found. 8. Repeat steps 2-7 until no new nodes can be found.

TWO-PERSON ZERO-SUM FINITE GAMES 65 9. The remaining nodes  $x$  get assigned the value 0. Any arc departing from such an  $x$  to another node not belonging to  $H$  is part of the saddle-point strategies. (The players do not have any desire to terminate the game once they reach any one of these remaining nodes.) Intuitively it should be clear that this algorithm indeed yields the saddle-point strategies for the game (the algorithm is in the spirit of dynamic programming and is identical to the recursive procedures introduced earlier in the context of feedback games, but here one also takes into account the possibility of nontermination). A formal proof is given in Washburn (1990).

Games with local pay-offs So far we have discussed only games with terminal pay-offs. Other pay-off structures are also possible, such as assigning a 'local' pay-off to each arc of the underlying graph. For games in extensive form in which the underlying graph is a finite tree, this is not a real generalization, since the local pay-offs assigned to arcs can be moved to the terminal nodes (and added together so as to obtain the 'total' pay-off at each terminal node). For extensive games with cycles in which pay-offs are assigned to arcs, this is not possible and other means are necessary to study saddle-point solutions (if they exist). For such games it brings in no loss of generality to assume that the underlying directed graph does not have terminal nodes (games with some terminal nodes can be modeled with a loop at those nodes: a terminal pay-off can then be assigned to the corresponding loop). Thus all plays are infinite.

The total pay-off associated with a play of the game is defined to be the long term average of the local pay-offs obtained during the play. If necessary one can consider Cesaro limits and/or take the Urn in}' or Urn sup. Note that if a 'terminal' node is reached (if such a node existed in the original game) to which a loop has been added, then the total pay-off converges to the local pay-off corresponding to that loop. For a theory of zero-sum games on graphs without terminal nodes, but with perfect information and with local pay-offs, the reader is referred to Ehrenfeucht and Mycielski (1979) and Alpern (1991). The theory and results are built around an auxiliary game defined on the same graph and with the same rules except for the fact that it ends as soon as a node is repeated. This auxiliary game has therefore a finite number of moves.

### 2.8 Action-Dependent Information Sets

The classes of zero-sum dynamic games we have seen heretofore all carried one common feature, which is that the information sets were fixed a priori. There are other important classes of games, however, where this side condition does not hold, which means that the information sets change with the actions of the players. In this section we discuss the theory behind such games, within the context of two specific examples—a class of duels, and a searchlight game.

66 T. BAŞAR AND G. J. OLSDER

#### 2.8.1 Duels

Consider a duel in which two men (P1 and P2) walk towards each other, taking one step at a time, at time instants  $t = 0, 1, \dots, N$ . If nothing intervenes, then they will meet at  $t = N$ . Each player has a pistol with exactly one bullet in it, and he may fire the pistol at any time  $t \in \{0, 1, \dots, N\}$ . If one of them fires and hits the other, the duel is immediately over and the one who has fired successfully is declared the winner. If neither one fires successfully or if both fire simultaneously and successfully, the duel becomes a stand-off. The probability of hitting the opponent after the pistol has been fired is a function of the distance between the two men. If a player fires at  $t \in \{0, 1, \dots, N\}$ , then the probability of hitting the opponent is  $t/N$ . To complete the formulation of the game, we have to specify whether the game is 'silent' or 'noisy'. In the silent duel a player does not know whether his opponent has already fired, unless, of course, he is hit. In a noisy duel a player knows when his opponent has fired and if he has not yet fired himself and is not hit he will wait until  $t = N$  to fire his own pistol, because at  $t = N$  he is certain to hit his opponent. Suppose that P1 fires at  $t_1$ . Then, if  $t_1 < t_2$ , the probability that P2 is hit is  $t_1/N$ , and the probability that P1 will be hit at  $t_2 > t_1$  is  $(1 - t_1/N)t_2/N$  for the silent duel and is  $1 - t_1/N$  for the



noisy duel (in which case P2 fires at  $t_2 - N$ ). Mutatis mutandis similar results are obtained if  $t_1 > t_2$ . With these probabilities one can construct a matrix  $A$  of size  $(N + 1) \times (N + 1)$ , the  $(t_1 + 1, t_2 + 1)$ -th element of which, with  $t_1 < t_2$  and  $t_i \in \{0, 1, \dots, N\}$ , is  $t_1/N - (1 - t_1/N)t_2/N$  for the silent case and  $(N - 2t_1)/N$  for the noisy case. The  $(t_1 + 1, t_2 + 1)$ -th element with  $t_1 > t_2$  is  $t_2/N - (1 - t_2/N)t_1/N$  for the silent case and  $1 - 2t_2/N$  for the noisy case. The diagonal elements of  $A$  equal 0 in both cases. Finding the optimal strategies for the duel now boils down to obtaining a saddle-point solution for the matrix game characterized by matrix  $A$  in which the minimizer (P1) chooses the row vector and the maximizer (P2) the column vector. We now note that for the extensive-form description of the noisy game, the extent of a player's knowledge on the evaluation of the game depends explicitly on his own and/or his opponent's actions. Hence, the noisy duel is a game with action-dependent information sets.

### 2.8.2 A searchlight game

In this game, two players, P1 and P2, move in a certain closed domain and are not aware of each other's position unless one of them flashes a searchlight which illuminates an area of known shape around him. By flashing his searchlight a player thus discloses his own position to the other player, regardless of their relative positions. Termination of the game occurs only if a player flashes his searchlight and the other player finds himself trapped in the area illuminated. Each player's objective is to catch the other player in his searchlight before he himself is caught. Therefore flashing has two competing consequences: in order for a player to win, he must flash; if, however, during such a flash the other

TWO-PERSON ZERO-SUM FINITE GAMES 67 player is not caught, then the flashing player is in a more vulnerable position because he has disclosed his location to the other player. We will provide here some explicit results for the case when the players are confined to  $n$  points on the circumference of a circle, which makes the underlying decision process a finite game. At each time step (the time is considered here to be discrete) a player can either move to one of the two adjacent nodes or stay where he is. Even if the players occupy the same position, they will not be aware of this, unless of course one of them (or both) flashes (flash). Flashing illuminates the position of the player who flashes, as well as the two adjacent positions. In the specific situation when only one of the players (say P1) has a searchlight at his disposal, the game becomes a survival game: P1 wants to catch the other player (P2), whereas P2 wants to avoid capture. A practical motivation for this game might be the

following. P2 is a smuggler who wants to steer his boat to the shore and P1 is a police boat. For obvious reasons P2 wants to avoid P1, and therefore he performs his practices only during the night, and that is exactly the reason why the police boat will use a searchlight. Of course in this case the search domain will not be the circumference of a circle, but rather a two-dimensional plane. The basic game The basic game to be formulated below is a building block for the more general game described above. Initially both players know each other's position on the circle (but not thereafter if mixed strategies are involved); P1 is at position  $p_1$  and P2 is at position  $p_2, p_i \in \{1, 2, \dots, n\}$ . Player P1 can flash only once during the time instants  $t = 1, 2, \dots, T$ , where  $T$  is the final time which is fixed and known to both players. P2 does not have any flashes at his disposal. Once the time proceeds the players do not acquire any new information on each other's new position (unless P1 flashes). The optimal strategies of this basic game can be found by solving the matrix game  $\min \max_{y^2} \sum_{t=1}^T \sum_{p_1} z_{p_1}^t$ , where the notation is as follows. At each instant of time, P2 can choose from three options (move to one of the two adjacent nodes or stay where he is) and therefore he has  $3^T$  pure strategies. Each component of  $y^2$  denotes the probability with respect to which one of the  $3^T$  pure strategies will be chosen. The set  $S^2$  is the simplex to which the  $3^T$ -vector  $y^2$  belongs. The components of  $z^1$ : on the other hand, indicate P1's pure strategies, determined by when and where to flash. After one time step P1 can flash at three different positions, after two time steps at five different positions, etc. Hence the number of pure strategies equals  $3 + 5 + \dots + (2T + 1) = T^2 + 2T$ , provided that  $T < \lfloor n/2 \rfloor$ . (For  $T > \lfloor n/2 \rfloor$ , all points on the circle can be reached by P1.) Vector  $z^1$  has  $d^1 = T^2 + 2T$  components and  $S^{d^1}$  denotes the simplex from which  $z^1$  must

68 T. BAŞAR AND G. J. OLSDER be chosen. The size of  $A_{p_1, p_2, T}$  is  $3^T \times d^1$  and the elements of this matrix are either 1 (corresponding to the cases when for the particular strategies capture takes place) or 0 (if no capture takes place). The first index of  $A$  refers to P1's initial position, the second index to P2's initial position, and the third index is the  $T$  introduced earlier. If the saddle-point value of this matrix game is denoted by  $J_{p_1, p_2, T}$ , then it should be intuitively clear (and this has been rigorously proven in Olsder and Papavassilopoulos (1988a)) that  $J_{p_1, p_2, 1} \wedge \dots \wedge J_{p_1, p_2, 2} \dots \wedge J_{p_1, p_2, \lfloor n/2 \rfloor} = J_{p_1, p_2, \lfloor n/2 \rfloor + 1} \dots \wedge J_{p_1, p_2, T} >$  for all  $p_1, p_2 \in \{1, 2, \dots, n\}$ . Hence, in the solution of the game, we can restrict the time horizon to  $\lfloor n/2 \rfloor$  steps, without any loss of

generality. The idea behind the proof of this result is that in  $\lfloor n/2 \rfloor$  steps each player can reach any point on the circle with an arbitrary underlying probability distribution. Now, consider the following generalization of the basic game. At  $t = 0$  P2 still knows where P1 is (at node  $p$ ), but P1 only knows the probability distribution of P2's position. The probability for P2 to be at node  $P_2$  at  $t = 0$  is  $v_{P_2}$ . The probability vector  $v = (v_1, \dots, v_n)'$  is assumed to be known to both players. Of course P2 knows his own position at  $t = 0$ . Such a generalization arises if one considers the following game: two players move on the circumference of a circle and do not know each other's initial position (other than in a probabilistic way). Each player has exactly one flash at his disposal. Suppose that P1 flashes first and that he does not capture P2 in this flash. Then P2 knows where P1 is at the time of the flash, but P1 only knows that P2 is not in the area just illuminated. If P1 does not have any other information about P2's position, then he may assume a uniform distribution on the nodes which were not illuminated at the time of the flash. This is the starting point for this generalization (with all  $V_i$ 's corresponding to nonilluminated nodes being equal and adding up to one). This generalization can again be formulated as a matrix game:  $\min \max V_{ij}$ . This minimax problem is not in the standard form treated earlier in Section 2.3, but it can also be solved by means of LP, viz.,  $\max_j \sum_i (j/i)j$

TWO-PERSON ZERO-SUM FINITE GAMES 69 subject to  $(y_1, \dots, y_n) \geq 0$ ;  $\sum_{j=1}^n (V_{ij})j = \sum_{j=1}^n (f_{0j})j$ ,  $i = 1, \dots, n$ ;  $\{V_{ij}\} > 0$ ,  $i = 1, \dots, n$ ;  $j = 1, \dots, n$ . The saddle-point value of this game when  $v_t = 0$  for  $i = 1, 2, 3$  and  $v_i = 1/(n - 3)$  for  $i = 4, \dots, n$  will be indicated by  $J_t$ , tacitly assuming that  $n > 3$ . Both players have one flash each. We will now solve the problem where each player has exactly one flash at his disposal and the players do not know each other's initial position, other than that they are given by the uniform distribution on the  $n$  nodes with no correlation between the two initial positions. If both players flash at  $t = 1$ , then the probability with which P2 captures P1 is  $3/n$  (since three positions are illuminated) and so is the probability for P1 to capture P2. Thus the pay-off is  $(3/n) - (3/n) = 0$ . Let us now consider the situation where P2 flashes at  $t = 1$  and P1 does not. Then the pay-off for P2 is  $(3/n) - (1 - 3/n)J_{t-1}$ . In other words, the pay-off equals the probability that P2 captures P1 minus the probability that P2 does not capture P1, times the probability that P1 captures P2 during the remaining  $T - 1$  period game. Similarly, if P2 flashes at  $t = 2$  and P1 does not flash at  $t = 1$  or  $t = 2$ , then the pay-off for P2 is

$(3/ra)-(1-3/n) jr_2$ , etc. Define  $ck = (3/n)-(1-3/ra)Jr_{fc}$ , and consider the matrix  $M = \begin{pmatrix} 0 & -c_1 & \dots & -c_1 & 0 & -c_2 & \dots & -c_2 & 0 & -c_3 & \dots & -c_3 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$ . Let  $s = (s_1, \dots, s_t)'$  and  $r = (r_1, \dots, r_t)'$  represent probability vectors according to which P1 and P2 choose to flash at times  $t = 1, \dots, T$ . Then the players face the following game:  $\min \max r' M s$ . The solution(s) to this latter matrix game provides the solution to our problem. Since the matrix  $M$  is skew-symmetric, it turns out that the value of the

70 T. BAŞAR AND G. J. OLSDER game is 0 (see Problem 9 in Section 2.9). Some further analysis regarding the signs and monotonicity properties of the elements  $C_k$  shows that the optimal strategies as when to flash are pure strategies. This is no longer true, however, if the illuminated areas for both players have different sizes. See Olsder and Papavassilopoulos (1988a) for further details and extensions. In conclusion, for the searchlight game as described above to be solved, one must solve a hierarchy of matrix games. The first matrix game determines the coefficients  $c^k$ , which are subsequently used as elements in the second matrix game defined by means of the matrix  $M$ .

2.9 Problems 1. Obtain the security strategies and the security levels of the players in the following matrix game. Does it admit a (pure) saddle-point solution? P2 P1  
 Obtain the mixed security strategies and the average security levels of the players in each of the following matrix games.

2 2 1 -1 1 3 2 1 -1 -1 3 0 1 2 4  
 1 P2 P1 4 0 3 0 2 1 P2 P1 1 3 4 -1 3. Verify that the quantity  $\sup^x \min_y y'Az$  is unique. 4. Prove that the quantity  $\max^x y'Az$  is a continuous function of  $y \in Y$ . 5. Determine graphically the mixed saddle-point solutions of each of the following matrix games. P2 P1 1 -3 0 3 -2 2 -1 2 -2 2 1 1 P2 P1 2 -1 -3 0 1 -1 4 6. Convert the following matrix games into linear programming problems and thereby numerically evaluate their saddle-point solutions. P2 P2 P1 3 2 1 1 2 3 P1 0 1 -1 0 -1 1 0 1 0 2 1 0 1 3 2 1 0

TWO-PERSON ZERO-SUM FINITE GAMES 71 7. The space of all real  $(m \times n)$  matrices  $A = \{a^k\}$  can be considered to be isomorphic to the  $mn$ -dimensional Euclidean space  $R^{mn}$ . Let  $S^{mn}$  denote the set of all real  $(m \times n)$  matrices which, considered as the pay-off matrix of a zero-sum game, admit pure strategy saddle-point equilibria. Then  $S^{mn}$  is isomorphic to a subset of  $R^{mn}$ , to be denoted by  $S^{m,n}$ . (i) Is  $S^{m,n}$  a subspace? (ii) Is  $S^{m,n}$  closed, convex? (iii) Describe  $S^{2,3}$  explicitly. 8. Let  $T(A)$  denote the set of all mixed saddle-point strategies of P1 in a matrix game  $A$ . Show that  $T(A)$  is

nonempty, convex, closed and bounded. 9. A matrix game  $A$ , where  $A$  is a skew-symmetric matrix, is called a symmetric game. For such a matrix game prove, using the notation of Problem 8, that  $T(A) = T^2(A)$ , and  $V_m(A) = 0$ . 10. Obtain the pure or behavioral saddle-point solutions of the following single-211210 10011221 11. Obtain a feedback saddle-point solution for the feedback game in extensive form depicted in Fig. 2.15. What is the value of the game? What is the actual play dictated by the feedback saddle-point solution? Show that these constant strategies also constitute a saddle-point solution for the feedback game. 12. Show that the feedback game in extensive form depicted in Fig. 2.16 does not admit a pure-strategy feedback saddle-point solution. Obtain its behavioral saddle-point solution and the behavioral saddle-point value. Compare the latter with the value of the game of Problem 11. What is the actual play dictated by the feedback saddle-point solution? 13. What is the open-loop version of the feedback game of Fig. 2.16? Obtain its saddle-point solution and value. 14. Prove that if a feedback game in extensive form admits a feedback saddle-point solution with the value  $J$ , and its open-loop version (if it exists)

72 T. BAŞAR AND G. J. OLSDER 0-1 15-2-3-5-423341-1 15 Figure 2.15:

Feedback game of Problem 11. 0-1 15-1-2-2-323341-102 Figure 2.16:

Feedback game of Problem 12. admits a unique saddle-point solution with the value  $J_0$ , then  $J = J_0$ . Furthermore, show that the open-loop saddle-point solution is the actual play dictated by the feedback saddle-point solution.

(Hint: make use of Prop. 2.7 and the interchangeability property of saddle-point strategies.) 15. Verify the following conjecture: If a feedback game in

extensive form admits a behavioral saddle-point solution which actually dictates a random choice at least at one level of play, then its open-loop version (if it exists) cannot admit a pure-strategy saddle-point equilibrium. 16.

Investigate whether the following multi-act game of the delayed commitment type admits a behavioral saddle-point solution or not. 2-30417230-1-5210-12

TWO-PERSON ZERO-SUM FINITE GAMES 73 17. Determine the pure or behavioral saddle-point solution of the game in extensive form depicted in Fig. 2.17. 0 1-10 0 110 Figure 2.17: The two-person dynamic game of Problem 17.

18. Determine the pure or behavioral saddle-point solution of the game in extensive form depicted in Fig. 2.18, which also incorporates a chance move. 09093993 Figure 2.18: The game of Problem 18. 19. Determine the solutions

to the previous problem in the cases in which (i) PI has access to nature's

choice; (ii) P2 has access to nature's choice; and (iii) both players have access to nature's choice. What is the value of this extra information to individual players in each case? 20. Construct an example of a matrix with only real eigenvalues, which, when considered as a zero-sum game, has a value in mixed or pure strategies that is smaller than the minimum eigenvalue. 21. A matrix game is said to be completely mixed if it admits a mixed saddle-point solution which assigns positive probability weight to all pure strategies of the players. Show that in a completely mixed game it does not matter whether P1 is minimizing and P2 is maximizing or the other way around, i.e., P1 is maximizing and P2 is minimizing. In both cases the saddle-point value and the strategies are the same. (For further results on completely mixed games the reader is referred to Bapat and Raghavan (1997).)

74 T. BAŞAR AND G. J. OLSDER 0-1-1 o Figure 2.19: The game of Problem 18. 22. Obtain for P1 his minimum guaranteed (average) costs in mixed, behavioral and randomized strategies, for the game with incomplete information, described by the tree of Fig. 2.19. (Answers: In mixed strategies, P1 can guarantee an average (minimum) outcome of  $-3/4$ , in behavioral strategies his security level is  $-25/64$ , and in randomized strategies it is  $-9/16$ .) 23. Consider the following matrix game. P2  $11+a; 5 \ 7 \ 9 + 3/$  If the parameters  $x$  and  $y$  would be known to both players, the matrix game is a standard one. Parameter  $x$  is known to P1 only, however, and  $y$  is only known to P2. What both players do know is that both parameters are stochastic, and independently and uniformly distributed on the interval  $[0, T]$ , with  $T \rightarrow 1$ . Show that the optimal solutions are  $1^*$  / first row, if  $0 < x < c_1$ , [ second row, if  $c_1 < x < 1$ ,  $2^*$  / first column, if  $0 < y < c_2$ , [ second column, if  $c_2 < y < 1$ , where the real constants  $c_1$  are (uniquely) determined by  $0 < c_1 < 1$  and  $2c_2(8 + c_1) + (c_2)^2 - 5 = 0$ ,  $2c_1(8 + c_2) + (c_1)^2 - 8 - 2c_2 = 0$ . For  $x = c_1$  and/or  $y = c_2$  the solution is not unique. Investigate the solution for  $\lim T \rightarrow 0$  and compare the limit behavior with the solution of the original matrix game with  $x = y = 0$ .

TWO-PERSON ZERO-SUM FINITE GAMES 75 2.10 Notes Section 2.2. The theory of finite zero-sum games dates back to Borel in the early 1920s whose work on the subject was later translated into English (Borel, 1953). Borel introduced the notion of a conflicting decision situation that involves more than one decision maker, and the concepts of pure and mixed strategies, but he did

not really develop a complete theory of zero-sum games. He even conjectured that the minimax theorem was false. It was Von Neumann who first came up with a proof of the minimax theorem and laid down the foundations of game theory as we know of today (Von Neumann, 1928, 1937). His pioneering book with Morgenstern definitely culminates his research on the theory of games (Von Neumann and Morgenstern, 1947). A full account of the historical development in finite zero-sum games, as well as possible applications in social sciences, is given in Luce and Raiffa (1957). The reader is also referred to McKinsey (1952) and the two edited volumes by Kuhn and Tucker (1950, 1953). The original proof of the minimax theorem given by Von Neumann is nonelementary and rather complicated. Since then, the theorem has been proven in several different ways, some being simpler and more illuminating than others. The proof given here seems to be the simplest one available in the literature.

Section 2.3. Several illustrative examples of the graphical solution of zero-sum matrix games can be found in the nontechnical book by Williams (1954). Graphical solution and the linear programming approach are not the only two applicable in the computation of mixed saddle-point solutions of matrix games; there is also the method of Brown and Von Neumann (1950) which relates the solution of symmetric games (see Problem 9 in Section 2.9 for a definition) to the solution of a particular differential equation. Furthermore, there is the iterative solution method of Brown (1951) by fictitious play (see also (Robinson, 1951)). A good account of these two numerical procedures, as well as another numerical technique of Von Neumann (1954) can be found in Luce and Raiffa (1957, Appendix 6). For a more recent development in this area, see Lakshminarayanan and Narendra (1981). Sections 2.4, 2.5 and 2.6. Description of single- and multi-act games in extensive form again goes back to the original work of Von Neumann and Morgenstern (1947), which was improved upon considerably by Kuhn (1953). For a critical discussion of prior and delayed commitment approaches to the solution of zero-sum games in extensive form the reader is referred to Aumann and Maschler (1972). Section 2.7. The theory presented in this section is based on Alpern (1988, 1991), Isbell (1957), and Washburn (1990). Washburn (1990) also discusses computational complexity of the solutions of zero-sum games on finite graphs with cycles and terminal pay-off. Extensions to discounting are also given in that reference, but extensions to games with random moves lead to complications in the solution algorithm in an essential way. Alpern (1991), on (not necessarily zero-sum) games with perfect information on finite graphs

with cycles and with local pay-offs, gives a constructive way to find optimal stationary strategies. See also Ehrenfeucht and Mycielski (1979) for earlier work on this topic. The former reference also discusses  $m$ -automated strategies; such strategies can be played by an automaton with  $m$  internal states and where inputs and outputs are coded to the nodes of the graph.

76 T. BAŞAR AND G. J. OLSDER Section 2.8. The two duels presented, the silent and the noisy one, are discretized versions (discretization with respect to time) of duels played in continuous time as given in Karlin (1959). For further discussion on extensions of game of timing see the Notes section of Chapter 4, with reference to Section 4.7. The theory of the searchlight game as presented here has been taken from Olsder and Papavassilopoulos (1988a, 1988b). Such games with a finite state space belong to the realm of Markov chain games; for a survey see Parthasarathy and Stern (1977), and Raghavan and Filar (1991). For other types of search games, see Gal (1980). Section 2.9. Problem 20: For more results on the relation between games and their spectra, see Weil (1968). Problem 22 has been taken from Isbell (1957).

### Chapter 3 Noncooperative Finite Games: TV-Person Nonzero-Sum 3.1

Introduction This chapter develops a general theory for static and dynamic nonzero-sum finite games under two different types of noncooperative solution concepts—those named after J. Nash and H. von Stackelberg. The analysis starts in Section 3.2 with bimatrix games which basically constitute a normal form description of two-person nonzero-sum finite games, and in this context the Nash equilibrium solution (in both pure and mixed strategies) is introduced and its properties and features are thoroughly investigated. The section concludes with some computational aspects of mixed-strategy Nash equilibria in bimatrix games. Section 3.3 extends the analysis of Section 3.2 to  $iV$ -person games in normal form and includes a proof for an important theorem (Thm. 3.2) which states that every  $iV$ -person finite game in normal form admits Nash equilibria in mixed strategies. In Section 3.4, the computational aspects of mixed-strategy Nash equilibria are briefly discussed, and in particular the close connection between mixed-strategy Nash equilibria in bimatrix games and the solution of a nonlinear programming problem is elucidated. Section 3.5 deals with properties and derivation of Nash equilibria (in pure, behavioral and mixed strategies) for dynamic TV-person games in extensive form. The analysis is first confined to single-act games and then extended to multi-act



games. In both contexts, the concept of "informational inferiority" between two finite games in extensive form is introduced, which readily leads to "informational nonuniqueness" of Nash equilibria in dynamic games. To eliminate this feature, the additional concept of "delayed commitment" type equilibrium strategy is introduced, and such equilibria of special types of single-act and 77

78 T. BAŞAR AND G. J. OLSDER multi-act finite games are obtained through recursive procedures. Other refinement schemes discussed in this section are the "perfectness" and "properness" of Nash equilibria. Section 3.6 is devoted to an extensive discussion of hierarchical equilibria in TV-person finite games. First, the two-person games are treated, and in this context both the (global) Stackelberg and feedback Stackelberg solution concepts are introduced, also within the class of mixed and behavioral strategies. Several examples are included to illustrate these concepts and the derivation of the Stackelberg solution. The analysis is then extended to many-player games. The final section of the chapter discusses extension of the results of the earlier sections to  $N$ -person finite games in extensive form which also incorporate chance moves.

3.2 Bimatrix Games Before developing a general theory of nonzero-sum noncooperative games, it is most appropriate first to investigate equilibria of bimatrix games which in fact constitute the most elementary type of decision problems within the class of nonzero-sum games. In spite of their simple structure, bimatrix games still carry most of the salient features and intricacies of noncooperative decision making. A bimatrix game can be considered as a natural extension of the matrix game of Section 2.2, to cover situations in which the outcome of a decision process does not necessarily dictate the verdict that what one player gains the other one has to lose. Accordingly, a bimatrix game is comprised of two  $(m \times n)$ -dimensional matrices,  $A = \{a_{ij}\}$  and  $B = \{b_{ij}\}$ , with each pair of entries  $(a_{ij}, b_{ij})$  denoting the outcome of the game corresponding to a particular pair of decisions made by the players. As in Section 2.2, we call the alternatives available to the players strategies. If  $P_1$  adopts the strategy "row  $i$ " and  $P_2$  adopts the strategy "column  $j$ ", then  $a_{ij}$  (respectively,  $b_{ij}$ ) denotes the loss incurred to  $P_1$  (respectively,  $P_2$ ). Being a rational decision maker, each player will strive for an outcome which provides him with the lowest possible loss. Stipulating that there exists no cooperation between the players and that they make their decisions independently, let us now investigate which pair(s) of strategies to declare as equilibrium pair(s) of

a given bimatrix game; in other words, how do we define a noncooperative equilibrium solution in bimatrix games? To recall the basic property of a saddle-point equilibrium solution in zero-sum games would definitely help in this investigation: "A pair of strategies is in saddle-point equilibrium if there is no incentive for any unilateral deviation by any one of the players" (Section 2.2). This is, in fact, a most natural property that an equilibrium solution is expected to possess, and surely it finds relevance also in nonzero-sum games. Hence, we have the following definition. Definition 3.1 A pair of strategies  $\{row\ i^*,\ column\ j^*\}$  is said to constitute a noncooperative (Nash) equilibrium solution to a bimatrix game  $(A = \{a_{ij}\}, B =$

$\{b_{ij}\})$  if the following pair of inequalities is satisfied for all  $i = 1, \dots, m$  and all  $j = 1, \dots, n$ :  $b_i^* j^* < b_{i,j}$ . (3.1a) (3.1b) Furthermore, the pair  $(a_{j^*j^*}, b_{i^*j^*})$  is known as a noncooperative (Nash) equilibrium outcome of the bimatrix game. Example 3.1 Consider the following  $(2 \times 2)$  bimatrix game  $P_2$   $P_2 = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$   $P_1 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ . It admits two Nash equilibria, as indicated:  $\{row\ 1,\ column\ 1\}$  and  $\{row\ 2,\ column\ 2\}$ . The corresponding equilibrium outcomes are  $(1,2)$  and  $(-1,0)$ .  $\square$  We now readily observe from the preceding example that a bimatrix game can admit more than one Nash equilibrium solution, with the equilibrium outcomes being different in each case. This then raises the question of whether it would be possible to order different Nash equilibrium solution pairs among themselves so as to declare only one of them as the most favorable equilibrium solution. This, however, is not completely possible, since a total ordering does not exist between pairs of numbers; but, notions of "betterness" and "admissibility" can be introduced through a partial ordering. Definition 3.2 A pair of strategies  $\{row\ i^1,\ column\ j^1\}$  is said to be better than another pair of strategies  $\{row\ i^2,\ column\ j^2\}$  if  $a_{i^1j^1} < a_{i^2j^2}$ , and if at least one of these inequalities is strict. Definition 3.3 A Nash equilibrium strategy pair is said to be admissible if there exists no better Nash equilibrium strategy pair. In Example 3.1, out of a total of two Nash equilibrium solutions only one of them is admissible ( $\{row\ 2,\ column\ 2\}$ ), since it provides uniformly lower costs for both players. This pair of strategies can therefore be declared as the most "reasonable" noncooperative equilibrium solution of the bimatrix game. In the case when a bimatrix game admits more than one admissible Nash equilibrium, however, such a clean choice cannot be made. Consider, for example, the  $(2 \times 2)$  bimatrix game  $P_2$   $P_2 =$

80 T. BAŞAR AND G. J. OLSDER which admits two admissible Nash equilibrium solutions: {row 1, column 1}, {row 2, column 2}, with the equilibrium outcomes being  $(-2, -1)$  and  $(-1, -2)$ , respectively. These Nash equilibria are definitely not interchangeable, and hence a possibility for the outcome of the game to be a nonequilibrium one in an actual play emerges. In effect, since there is no cooperation between the players by the nature of the problem, P1 might stick to one equilibrium strategy (say, row 1) and P2 might adopt the other one (say, column 2), thus yielding an outcome of  $(1, 1)$  which is unfavorable to both players. This is indeed one of the dilemmas of noncooperative nonzero-sum decision making (if a given problem does not admit a unique admissible equilibrium), but there is really no remedy for it unless one allows some kind of communication between the players, at least to ensure that a nonequilibrium outcome will not be realized. In fact, a possibility of collusion saves only that aspect of the difficulty in the above example, since the "best" choice still remains indeterminate. One way of resolving this indeterminacy is to allow a repetition and to choose one or the other Nash equilibrium solution at consecutive time points. The average outcome is then clearly  $(-3/2, -3/2)$ , assuming that the players make their choices dependent on each other. However, this favorable average outcome is not an equilibrium outcome in mixed strategies, as we shall see in the last part of this section. The solution proposed above is acceptable in the "real" decision situation, known as the battle of the sexes, that the bimatrix game (3.2) represents. The story goes as follows. A couple has to make a choice out of two alternatives for an evening's entertainment. The alternatives are a basketball game which is preferred by the husband (P1) and a musical comedy which is preferred by the wife (P2). But they would rather go together, instead of going separately to their top choices. This then implies, also in view of the bimatrix game (3.2), that they will eventually decide to go together either to the game or to the comedy. It should be noted that there is a flavor of cooperation in this example, but the solution is not unique in any sense. If repetition is possible, however, then they can go one evening to the game, and another evening to the comedy, thus resolving the problem in a "fair" way. Such a solution, however, does not carry any game-theoretic meaning as we shall see in the last part of this section. The message here is that if a bimatrix game admits more than one admissible Nash equilibrium solution, then a valid approach would be to go back to the actual decision process that the bimatrix game represents, and see whether a version of the bimatrix game that also accounts for repetition admits

a solution which is favored by both parties. This, of course, is all valid provided that some communication is allowed between the players, in which case the "favorable" solution thus obtained will not be a noncooperative equilibrium solution. Thus, we have seen that if a bimatrix game admits more than one admissible Nash equilibrium solution, then the equilibrium outcome of the game becomes rather ill-defined—mainly due to the fact that multiple Nash equilibria are in general not interchangeable. This ambiguity disappears, of course, whenever the equilibrium strategies are interchangeable, which necessarily requires the corresponding outcomes to be the same. Since zero-sum matrix games are special types of bimatrix games, in which case the equilibrium solutions are known to

TV-PERSON NONZERO-SUM FINITE GAMES 81 be interchangeable (Corollary 2.1), it follows that there exists some nonempty class of bimatrix games whose equilibrium solutions possess such a property. This class is, in fact, larger than the set of all  $(m \times n)$  zero-sum games, as it is shown below.

**Definition 3.4** Two  $(m \times n)$  bimatrix games  $(A, B)$  and  $(C, D)$  are said to be strategically equivalent if there exist positive constants  $a_1, a_2$ , and scalars  $0_1, 0_2$ , such that  $a_1 - i_j = 0_1 d_j + 0_1$ , (3.3a)  $b_{ij} = a_2 d_{ij} + 0_2$ , (3.3b) for  $a_{1i} = 1, \dots, m$ ;  $j = 1, \dots, n$ .

**Remark 3.1** The reader can easily verify that the pair (3.3a)-(3.3b) is indeed an equivalence relation since it is symmetric, reflexive and transitive.  $\square$

**Proposition 3.1** All strategically equivalent bimatrix games have the same Nash equilibria. **Proof.** Let  $(A, B)$  be a bimatrix game with a Nash equilibrium solution  $\{\text{row } i^*, \text{column } j^*\}$ . Then, by definition, inequalities (3.1a)-(3.1b) are satisfied. Now let  $(C, D)$  be any other bimatrix game that is strategically equivalent to  $(A, B)$ . Then there exist  $Q_i > 0, a_2 > 0, 0_1, 0_2$  such that (3.3a)-(3.3b) are satisfied for all possible  $i$  and  $j$ . Using these relations in (3.1a) and (3.1b), we obtain, in view of the positivity of  $a_1$  and  $a_2$ , for all  $i = 1, \dots, m$  and all  $j = 1, \dots, n$ , which proves the proposition since  $(A, B)$  was an arbitrary bimatrix game.  $\square$

**Proposition 3.2** Multiple Nash equilibria of a bimatrix game  $(A, B)$  are interchangeable if  $(A, B)$  is strategically equivalent to  $(A, -A)$ . **Proof.** The result readily follows from Corollary 2.1 since, by hypothesis, the bimatrix game  $(A, B)$  is strategically equivalent to a zero-sum game  $A$ .  $\square$

The undesirable features of the Nash equilibrium solutions of bimatrix games, that we have witnessed, might at first sight lead to questioning appropriateness and suitability of the Nash equilibrium solution concept for such noncooperative multi-person decision problems. Such a verdict, however,

is not fair, since the undesirable features detected so far are due to the noncooperative nature of the decision problem under consideration, rather than to the weakness of the Nash equilibrium solution concept. To supplement our argument, let us consider one extreme case (the so-called "team" problem) that involves two players, identical goals (i.e.,  $A = B$ ), and noncooperative decision making. Note that, since

82 T. BAŞAR AND G. J. OLSDER possible outcomes are single numbers in this case, a total ordering of outcomes is possible, which then makes the admissible Nash equilibrium solution the only possible equilibrium solution for this class of decision problems. Moreover, the admissible Nash equilibrium outcome will be unique and be equal to the smallest entry of the matrix that characterizes the game. Note, however, that even for such a reasonable class of decision problems, the players might end up at a nonequilibrium point in case of multiple equilibria; this being mainly due to the noncooperative nature of the decision process which requires the players to make their decisions independently. To illustrate such a possibility, let us consider the identical goal game (team problem) with the cost matrices  $P_2 A = B = EE - ]P_1$ , (3.4) which both parties try to minimize, by appropriate choices of strategies. There are clearly two equilibria,  $\{\text{row 1, column 1}\}$  and  $\{\text{row 2, column 2}\}$ , both yielding the minimum cost of 0 for both players. These equilibria, however, are not interchangeable, thus leaving the players at a very difficult position if there is no communication between them. (Perhaps tossing a fair coin and thereby securing an average cost of  $\frac{1}{2}$  would be preferred by both sides to any pure strategy which could easily lead to an unfortunate cost of 1.) This then indicates that even the well-established team-equilibrium solution could be frowned upon in case of multiple equilibria when the decision mode is noncooperative. One can actually make things look even worse by replacing matrix (3.4) with a nonsymmetric one  $P_2 \wedge = B = PrTirlpi >$  (3-5) which admits the same two equilibria as in the previous matrix game. But what are the players actually going to play in this case? To us, there seems to be only one logical play for each player:  $P_1$  would play "row 2", hoping that  $P_2$  would play his second equilibrium strategy, but at any rate securing a loss ceiling of 1.  $P_2$ , reasoning along similar lines, would rather play "column 1". Consequently, the outcome of their joint decision is 1 which is definitely not an equilibrium outcome. But, in an actual play of this matrix game, this outcome is more likely to occur than the equilibrium outcome. One might counter-argue at this point

and say that P2 could very well figure out P1's way of thinking and pick "column 2" instead, thus enjoying the low equilibrium outcome. But what if P1 thinks the same way? There is clearly no end to this so-called second-guessing iteration procedure if adopted by both players as a method that guides them to the "optimum" strategy.

ZV-PERSON NONZERO-SUM FINITE GAMES 83 One might wonder at this point as to why, while the multiple Nash equilibria of bimatrix games (including identical-goal games) possess all these undesirable features, the saddle-point equilibrium solutions of zero-sum (matrix) games are quite robust—in spite of the fact that both are noncooperative decision problems. The answer to this lies in the nature of zero-sum games: they are completely antagonistic, and the noncooperative equilibrium solution fits well within this framework. In other nonzero-sum games, however, the antagonistic nature is rather suppressed or is completely absent, and consequently the "noncooperative decision making" framework does not totally suit such problems. If some cooperation is allowed, or if there is a hierarchy in decision making, then equilibrium solutions of nonzero-sum (matrix) games could possess more desirable features. We shall, in fact, observe this later in Section 3.6 where we discuss equilibria in nonzero-sum games under the latter mode of decision making. One important property of the saddle-point strategies in zero-sum games was that they were also the security strategies of the players (see Thm. 2.2). In bimatrix games, one can still introduce security strategies for the players in the same way as it was done in Section 2.2, and with the properties cited in Thm. 2.1 being valid with the exception of the third one. It should of course be clear that, in a bimatrix game  $(A, B)$  the security strategies of P1 involve only the entries of matrix  $A$ , while the security strategies of P2 involve only those of matrix  $B$ . In Example 3.1, considered earlier, the unique security strategy of P1 is "row 1", and the unique security strategy of P2 is "column 1". Thus, considered as a pair, the security strategies of the players correspond to one of the Nash equilibrium solutions, but not to the admissible one, in this case. One can come up with examples of bimatrix games in which the security strategies do not correspond to any one of the equilibrium solutions, or examples of bimatrix games in which they coincide with an admissible Nash solution. As an illustration of the latter possibility, consider the following bimatrix game:

	P2	P2	w	3	0	2	W	0	3	2
--	----	----	---	---	---	---	---	---	---	---

This admits a unique pair of Nash equilibrium strategies, as indicated, which

are also the security strategies of the players. Moreover, the equilibrium outcome is determined by the security levels of the players. The preceding bimatrix game is known in the literature as the prisoners' dilemma. It characterizes a situation in which two criminals, suspected of having committed a serious crime, are detained before a trial. Since there is no direct evidence against them, their conviction depends on whether they confess or not. If the prisoners both confess, then they will be sentenced to 8 years. If neither one confesses, then they will be convicted of a lesser crime and sentenced to

84 T. BAŞAR AND G. J. OLSDER 2 years. If only one of them confesses and puts the blame on the other one, then he is set free according to the laws of the country and the other one is sentenced to 30 years. The unique Nash equilibrium in this case dictates that they both should confess. Note, however, that there is in fact a better solution for both criminals, which is that they both should refuse to confess; but implementation of such a solution would require a cooperation of some kind (and also trust), which is clearly out of the question in the present context. Besides, this second solution (which is not Nash) is extremely unstable, since a player will find it to his advantage to unilaterally deviate from this position at the last minute. Even if the security strategies might not be in noncooperative equilibrium, they can still be employed in an actual game, especially in cases when there exist two or more noninterchangeable Nash equilibria or when a player is not completely sure of the cost matrix, or even the rationality of the other player. This reasoning leads to the following second solution concept in bimatrix games. Definition 3.5 A pair of strategies  $\{\text{row } i, \text{column } j\}$  is known as a pair of minimax strategies for the players in a bimatrix game  $(A, B)$  if the former is a security strategy for  $P_1$  in the matrix game  $A$ , and the latter is a security strategy for  $P_2$  in the matrix game  $B$ . The corresponding security levels of the players are known as the minimax values of the bimatrix game. Remark 3.2 It should be noted that minimax values of a bimatrix game are definitely not lower (in an ordered way) than the pair of values of any Nash equilibrium outcome. Even if the unique Nash equilibrium strategies correspond to the minimax strategies, the minimax values could be higher than the values of the Nash equilibrium outcome, mainly because the minimax strategy of a player might not constitute optimal response to the minimax strategy of the other player.  $\square$  Mixed strategies In our discussion of bimatrix games in this section, we have so far

encountered only cases in which a given bimatrix game admits a unique or a multiple of Nash equilibria. There are other cases, however, in which Nash equilibrium strategies might not exist, as illustrated in the following example. In such a case we simply say that a Nash equilibrium does not exist in pure strategies. Example 3.2 Consider the  $(2 \times 2)$  bimatrix game

P2	3	0	2	1	1	2	0
-1							

It clearly does not admit a Nash equilibrium solution in pure strategies. The minimax strategies, however, exist (as they always do, by Thm. 2.1 (ii)) and are given as {row 1, column 2}. □

**IV-PERSON NONZERO-SUM FINITE GAMES 85** Paralleling our analysis in the case of nonexistence of a saddle point in zero-sum games (Section 2.2), we now enlarge the class of strategies so as to include all mixed strategies, defined as the set of all probability distributions on the set of pure strategies of each player (see Def. 2.2). This extension is in fact sufficient to ensure the existence of a Nash equilibrium solution (in mixed strategies, in a bimatrix game)—a result which we state below in Thm. 3.1, after making the concept of a Nash equilibrium in this extended space precise. Using the same notation as in Defs. 2.3-2.5, wherever appropriate, we first have the following definition.

**Definition 3.6** A pair  $\{y^* \in Y, z^* \in Z\}$  is said to constitute a noncooperative (Nash) equilibrium solution to a bimatrix game  $(A, B)$  in mixed strategies, if the following inequalities are satisfied for all  $y \in Y$  and  $z \in Z$ :

(3.7a)  $y^* \cdot Az^* \geq y \cdot Az^*$ ,  $y \in Y$

(3.7b)  $y^* \cdot Bz^* \leq y^* \cdot Bz$ ,  $z \in Z$

Here, the pair  $(y^* \cdot Az^*, y^* \cdot Bz^*)$  is known as a noncooperative (Nash) equilibrium outcome of the bimatrix game in mixed strategies.

**Theorem 3.1** Every bimatrix game has at least one Nash equilibrium solution in mixed strategies. *Proof.* We postpone the proof of this important result until the next section where it is proven in a more general context (Thm. 3.2).

**D** Computation of mixed Nash equilibrium strategies of bimatrix games is more involved than the computation of mixed saddle-point solutions, and in the literature there are very few generally applicable methods in that context. One of these methods converts the original game problem into a nonlinear programming problem, and it will be discussed in Section 3.4. What we intend to include here is an illustration of the computation of mixed Nash equilibrium strategies in  $(2 \times 2)$  bimatrix games by directly making use of inequalities (3.7a)-(3.7b). Specifically, consider again the bimatrix game of Example 3.2 which is known not to admit a pure-strategy Nash equilibrium. In this case, since every  $y \in Y$  can be written as  $y = (j/i, (1 - j/i))$  with  $0 < j/i < 1$ , and similarly every  $z \in Z$  as  $z = (z_i, (1 - z_i))$  with  $0 < z_i < 1$ , we first



obtain the following equivalent inequalities for (3.7a) and (3.7b), respectively, for the bimatrix game under consideration:  $-2y^i z^j + y^i < -2y^i z^j + y^i$ ,  $0 < y^i < 1$ ,  $2y^i z^j - z^j < 2y^i z^j - z^j$ ,  $0 < z^j < 1$ . Here  $(y^i, (1 - y^i))'$  and  $(z^j, (1 - z^j))'$  denote the mixed Nash equilibrium strategies of P1 and P2, respectively, which have yet to be computed so that the preceding pair of inequalities is satisfied. A way of obtaining this solution would be to find that value of  $z^j$  (if it exists in the interval  $[0,1]$ ) which would make the RHS of the first inequality independent of  $y^i$ , and also the value of  $y^i$  that

86 T. BAŞAR AND G. J. OLSDER would make the RHS of the second inequality independent of  $z^j$ . The solution then readily turns out to be  $(y^* = \hat{y}, z^* = \hat{z})$  which can easily be checked to be the unique solution set of the two inequalities. Consequently the bimatrix game of Example 3.2 admits a unique Nash equilibrium solution in mixed strategies, which is  $\{y^* = (\hat{y}, \hat{z})', z^* = (\hat{z}, \hat{y})'\}$ . Remark 3.3 In the above solution to Example 3.2 we observe an interesting, and rather counter-intuitive, feature of the Nash equilibrium solution in mixed strategies. The solution obtained, i.e.,  $\{y^*, z^*\}$ , has the property that the inequalities (3.7a) and (3.7b) in fact become independent of  $y \in Y$  and  $z \in Z$ . In other words, direction of the inequality loses its significance. This leads to the important conclusion that if the players were instead seeking to maximize their average costs, then we would have obtained the same mixed equilibrium solution. This implies that, while computing his mixed Nash equilibrium strategy, each player pays attention only to the average cost function of his co-player, rather than optimizing his own average cost function. Hence, the nature of the optimization (i.e., minimization or maximization) becomes irrelevant in this case. The same feature can also be observed in mixed saddle-point solutions, yielding the conclusion that a zero-sum matrix game  $A$  could have the same mixed saddle-point solution as the zero-sum game with matrix  $-A$ . (See Problem 21 in section 2.9.) D The following proposition now makes the conclusions of the preceding remark precise. Proposition 3.3 Let  $Y$  and  $Z$  denote the sets of inner points (interiors) of  $Y$  and  $Z$ , respectively. If a bimatrix game  $(A, B)$  admits a mixed-strategy Nash equilibrium solution  $\{y^* \in Y, z^* \in Z\}$ ,<sup>15</sup> then this also serves as a mixed-strategy Nash equilibrium solution for the bimatrix game  $(-A, -B)$ . Proof By hypothesis, inequalities (3.7a) and (3.7b) are satisfied by a pair  $\{y^* \in Y, z^* \in Z\}$ . This implies that  $y^i A z^j$  is attained by an inner point of  $Y$ , and hence, because of linearity of  $y^i A z^j$  in  $y$ , this expression has to be independent of  $y$ . Similarly,  $y^i B z^j$  is independent of  $z$ .

Consequently, (3.7a)- (3.7b) can be written as  $y^*Az^* = y'Az^*$ ,  $\forall y \in Y$ ,  $y^*Bz^* = y^*Bz$ ,  $\forall z \in Z$ , which further leads to the inequality-pair  $y^* \{-A\}z^* < y' \{-A\}z^*$ ,  $\forall y \in Y$ ,  $y^* \{-B\}z^* < y^* \{-B\}z$ ,  $\forall z \in Z$ , which verifies the assertion of the proposition. D 15 Such a solution is also known as a completely mixed Nash equilibrium solution or as an inner mixed-strategy Nash equilibrium solution.

iV-PERSON NONZERO-SUM FINITE GAMES 87 As a special case of this proposition, we now immediately have the following result for zero-sum matrix games. Corollary 3.1 Any inner mixed saddle point of a zero-sum matrix game  $A$  also constitutes an inner mixed saddle point for the zero-sum matrix game  $-A$ . Even though we have introduced the concept of a mixed noncooperative equilibrium solution for bimatrix games which do not admit a Nash equilibrium in pure strategies, it is quite possible for a bimatrix game that admits (pure) Nash equilibria to admit a mixed Nash equilibrium as well. To verify this possibility, let us again consider the bimatrix game of the "battle of the sexes", given by (3.2). We have already seen that it admits two noninterchangeable Nash equilibria with outcomes  $(-2, -1)$  and  $(-1, -2)$ . Let us now investigate whether it admits a third equilibrium, this time in mixed strategies. For this bimatrix game, letting  $y = (y_i, (1 - J/i))' \in Y$  and  $z = (z_1, (1 - z_i))' \in Z$ , we first rearrange the inequalities (3.7a)-(3.7b) and write them in the equivalent form  $-5y_j z_j + 2y_j < -5y_i z_j + 2y_i$  ( $0 < y_x < 1$ ,  $-hy^z + Zz \{ < -5t/iZ_i + 3z_i$ ,  $0 < z_i < 1$ . Now it readily follows that these inequalities admit the unique inner point solution  $\{y^* = \lambda, z^* = \mu\}$  dictating the mixed Nash equilibrium solution  $\{y^* = (\lambda, \mu)', z^* = (\mu, \lambda)'\}$ , and an average outcome of  $\{-\lambda, -\mu\}$ . Of course, the above inequalities also admit the solutions  $(y_j = 1, z_i = 1)$  and  $(y_i = 0, z_i = 0)$  which correspond to the two pure Nash equilibria which we have discussed earlier at some length. Our interest here lies in the nature of the mixed equilibrium solution. Interpreted within the context of the "battle of the sexes", this mixed solution dictates that the husband chooses to go to the game with probability  $\lambda$  and to the musical comedy with probability  $\mu$ . The wife, on the other hand, decides on the game with probability  $\mu$  and on the comedy with probability  $\lambda$ . Since, by the very nature of the noncooperative equilibrium solution concept, these decisions have to be made independently, there is a probability of  $13/25$  with which the husband and wife will have to spend the evening separately. In other words, even if repetition is allowed, the couple will be together for less than half the time, as dictated by the unique mixed Nash equilibrium solution; and this so in spite of the fact that the

husband and wife explicitly indicate in their preference ranking that they would rather be together in the evening. The main reason for such a dichotomy between the actual decision process and the solution dictated by the theory is that the noncooperative equilibrium solution concept requires decisions to be made independently, whereas in the bimatrix game of the "battle of the sexes" our way of thinking for a reasonable equilibrium solution leads us to a cooperative solution which inevitably asks for dependent decisions. (See Problem 4, Section 3.8, for more variations on the "battle of the sexes".) The conclusion we draw here is that bimatrix games with pure-strategy Nash equilibria could also admit mixed-strategy Nash solutions which, depending on

88 T. BAŞAR AND G. J. OLSDER the relative magnitudes of the entries of the two matrices, could yield better or worse (and of course also noncomparable) average equilibrium outcomes.

### 3.3 TV-Person Games in Normal Form

The class of TV-person nonzero-sum finite static games in normal form models a decision making process similar in nature to that modeled by bimatrix games, but this time with  $TV(> 2)$  interacting decision makers (players). Decisions are again made independently and out of a finite set of alternatives for each player. Since there exist more than two players, a matrix formulation on the plane is not possible for such games, thus making the display of possible outcomes and visualization of equilibrium strategies rather difficult. However, a precise formulation is still possible, as it is provided below together with the notation to be used in describing TV-person finite static games in normal form.

Formulation of an TV-person finite static game in normal form

- (1) There are  $N$  players to be denoted by  $P_1, P_2, \dots, P_N$ . Let us further denote the index set  $\{1, 2, \dots, N\}$  by  $N$ .
- (2) There is a finite number of alternatives for each player to choose from. Let  $m_j$  denote the number of alternatives available to  $P_i$ , and further denote the index set  $\{1, 2, \dots, m_j\}$  by  $M_j$ , with a typical element of  $M_j$  designated as  $n_j$ .
- (3) If  $P_j$  chooses a strategy  $n_j \in M_j$ , and this so for all  $j \in N$ , then the loss incurred to  $P_i$  is a single number  $a_i(n_1, \dots, n_N)$ . The ordered  $N$ -tuple of all these numbers (over  $i \in N$ ), i.e.,  $(a_1(n_1, \dots, n_N), \dots, a_N(n_1, \dots, n_N))$ , constitutes the corresponding unique outcome of the game.
- (4) Players make their decisions independently and each one unilaterally seeks the minimum possible loss, of course by also taking into account the possible rational choices of the other players. The noncooperative equilibrium solution concept within the context of this TV-person game can be introduced as follows as a direct extension of Def. 3.1.

**Definition 3.7** An  $N$ -tuple of strategies  $\{n_1^*, \dots, n_N^*\}$ , with  $n_i^* \in M_j$ ,  $i \in N$

$N$ , is said to constitute a noncooperative (Nash) equilibrium solution for an  $N$ -person nonzero-sum static finite game in normal form, as formulated above, if the following  $N$  inequalities are satisfied for all  $n_i \in M_j, i \in N$ :  $a_i^* A 1 < 1^\circ - \circ n_i, n_j, \wedge B \text{ ,,} 2^* A 2 < 2 \sim \text{un}'' 1, n', \dots, n'N \text{ — } "n^*, 1x2, 1X3, \langle i \rangle 5, \text{---} \rangle \langle jv' nN^* A N < N (3.8)$

AT-PERSON NONZERO-SUM FINITE GAMES 89 Here, the  $N$ -tuple  $(a_1^*, a_2^* \dots a_N^*)$  is known as a noncooperative (Nash) equilibrium outcome of the  $N$ -person game in normal form. There is actually no simple method to determine the Nash equilibrium solutions of  $iV$ -person finite games in normal form. One basically has to check exhaustively all possible combinations of  $iV$ -tuples of strategies, to see which ones provide a Nash equilibrium. This enumeration, though straightforward, could at times be rather strenuous, especially when  $N$  and/or  $m_i, i \in N$ , are large. However, given an  $iV$ -tuple of strategies asserted to be in Nash equilibrium, it is relatively simpler to verify their equilibrium property, since one then has to check only unilateral deviations from the given equilibrium solution. To get a flavor of the enumeration process and the method of verification of a Nash equilibrium solution, let us now consider the following example. Example 3.3 Consider a 3-person game in which each player has two alternatives to choose from. That is,  $N = 3$  and  $m_i = m_i = m^{\wedge} = 2$ . To complete the description of the game, the  $2^3 = 8$  possible outcomes are given as  $(f_{i1}, i, i_{a1}, i, i_{a1}, i, i) = U. -1.0)$ .  $(a_{1,2}, l_{a1,2}, l_{a1,2}, l) = (2, 1, 1)$ ,  $(a_{2,2}, l_{a2,2}, l_{a2,2}, l) = (0, -2, 1)$ ,  $(a_{i,1,2}, l_{a1,2}, l_{a1,2}, l) = (1, 1, 1)$ ,  $(a_{i,2,2}, l_{i,2,2}, l_{i,2,2}, l) = (1, 1, 1)$ ,  $(a_{ii,2,4}, i, 2ah, 2) = (0, 1, 2)$ ,  $(a_{2,2,2}, l_{2,2,2}, l_{2,2,2}, l) = (-1, 2, 0)$ . These possible outcomes can actually be displayed in the form of two  $(2 \times 2)$  matrices  $P_2 \text{ } ^3 = 1^* (1, -1, 0) (2, 0, 1) (2, 1, 1) (0, -2, 1)$  PI. (3.9a)  $P_3 = 2 (1, 1, 1) (0, 1, 2) (0, 1, 0) (-1, 2, 0)$  PI, (3.9b) where the entries of the former matrix correspond to the possible outcomes if  $P_3$ 's strategy is fixed at  $n^{\circ} = 1$ , and the latter matrix provides possible outcomes if his strategy is fixed at  $n^{\circ} = 2$ . We now assert that the "starred" entry is a Nash equilibrium outcome for this game. A verification of this assertion would involve

90 T. BAŞAR AND G. J. OLSDER three separate checks, concerning unilateral deviation of each player. If  $P_1$  deviates from this asserted equilibrium strategy  $n^{\circ} = 1$ , then his loss becomes 2 which is not favorable. If  $P_2$  deviates from  $n^{\circ} = 1$ , his loss becomes 1 which is not favorable either. Finally, if  $P_3$  deviates from  $n^{\circ} = 1$ , his loss becomes 1 which is higher than

his asserted equilibrium loss 0. Consequently, the first entry of the first matrix, i.e., (1,-1,0), indeed provides a Nash equilibrium outcome, with the corresponding equilibrium strategies being  $\{n \rightarrow n \rightarrow 1, n^1 = 1\}$ . The reader can now check by enumeration that this is actually the only Nash equilibrium solution of this 3-person game. D When the Nash equilibrium solution of an  $N$ -person nonzero-sum game is not unique, we can again introduce the concept of "admissible Nash equilibrium solution" as direct extensions of Defs. 3.2 and 3.3 to the AT-person case. It is of course possible that a given  $N$ -person game will admit more than one admissible Nash equilibrium solution which are also not interchangeable. This naturally leads to an ill-defined equilibrium outcome; but we will not elaborate on these aspects of the equilibrium solution here since they were extensively discussed in the previous section within the context of bimatrix games, and that discussion can readily be carried over to fit the present framework. Furthermore, Prop. 3.1 has a natural version in the present framework, which we quote below without proof, after introducing an extension of Def. 3.4. Proposition 3.2, on the other hand, has no direct counterpart in an  $N$ -person game with  $N > 2$ . Definition 3.8 Two nonzero-sum finite static games in normal form are said to be strategically equivalent if the following three conditions are satisfied: (i) The two games have the same number of players (say  $N$ ), (ii) each player has the same number of alternatives in both games, (iii) if  $\{(a^1, \dots, a^N) \in M_j, i \in N\}$  is the set of possible outcomes in one game, and  $\{(b^1, \dots, b^N) \in M_i, i \in N\}$  is the set of possible outcomes in the other, then there exist positive constants  $\alpha_i, i \in N$ , and scalars  $\beta_j, j \in N$ , such that  $\langle \alpha_i a^i, \dots, \alpha_i a^i \rangle = \langle \beta_j b^j, \dots, \beta_j b^j \rangle$  (3-10) for all  $i \in M_j, j \in N$ . Proposition 3.4 All strategically equivalent nonzero-sum finite static games in normal form have the same set of Nash equilibria. Minimax strategies introduced earlier within the context of bimatrix games (Def. 3.5) find applications also in AT-person games, especially when there exist more than one admissible Nash equilibrium which are not interchangeable. In the present context their definition involves a natural extension of Def. 3.5, which we do not provide here.

**N-PERSON NONZERO-SUM FINITE GAMES** 91 Example 3.4 In the 3-person game of Example 3.3, the minimax strategies of P1, P2 and P3 are  $h_1 = 1$  or 2,  $n_2 = 1$  and  $h_3 = 1$ , respectively. It should be noted that for P1 any strategy is minimax. The security levels of the players, however, are unique (as they should be) and are 2, 1 and 1, respectively. Note that these values are higher

than the unique Nash equilibrium outcome  $(1, -1, 0)$ , in accordance with the statement of Remark 3.2 altered to fit the present context. Mixed-strategy Nash equilibria Mixed noncooperative equilibrium solutions of an  $N$ -person finite static game in normal form can be introduced by extending Def. 3.6 to  $N$  players, i.e., by replacing inequalities (3.7a) and (3.7b) by an  $N$ -tuple of similar inequalities. To this end, let us first introduce the notation  $Y$  to denote the mixed strategy space of  $P_i$ , and further denote a typical element of this space by  $y_i$ , and its  $k$ th component by  $y_{ik}$ . Then, we have the following definition. Definition 3.9 An  $N$ -tuple  $\{y_i^* \in Y_i; i \in N\}$  is said to constitute a mixed-strategy noncooperative (Nash) equilibrium solution for an  $N$ -person finite static game in normal form if the following  $N$  inequalities are satisfied for all  $y_j \in Y_j; j \in N$ :

$$u_i(y_i^*, y_{-i}^*) \geq u_i(y_i, y_{-i}^*) \quad \forall y_i \in Y_i, y_{-i}^* \in \prod_{j \neq i} Y_j \quad (3.11)$$

Here, the  $N$ -tuple  $(y_1^*, \dots, y_N^*)$  is known as a noncooperative (Nash) equilibrium outcome of the  $N$ -person game in mixed strategies. One of the important results of static game theory is that every  $N$ -person game of the type discussed in this section admits a mixed-strategy Nash equilibrium solution (note that pure strategies are also included in the class of mixed strategies). We now state and prove this result which is also an extension of Thm. 3.1.

**Theorem 3.2** Every  $N$ -person static finite game in normal form admits a non-cooperative (Nash) equilibrium solution in mixed strategies.

92 T. BAŞAR AND G. J. OLSDER **Proof.** Let  $\{Y_i; i \in N\}$  denote an  $N$ -tuple of mixed strategies for the  $N$ -person game, and introduce  $u_i(y_i, y_{-i})$ .  $u_i(y_i, y_{-i}) = \sum_{j \in N} p_j u_{ij}(y_i, y_{-i})$  where  $p_j$  is the probability that player  $j$  chooses strategy  $y_j$  and  $u_{ij}(y_i, y_{-i})$  is the payoff to player  $i$  if he chooses strategy  $y_i$  and the other players choose strategies  $y_{-i}$ . Note that this function is continuous on the product set  $\prod_{i \in N} Y_i$ . Now, let  $c_{ij}$  be related to  $u_{ij}$  by  $c_{ij} = \max\{u_{ij}/4, 0\}$  for each  $j \in N$  and  $i \in N$ . It readily follows that  $c_j = (c_{j1}, \dots, c_{jN})$  is also a continuous function on  $\prod_{i \in N} Y_i$ . Then, introducing the transformation  $T(y_1, \dots, y_N) = (y_1 + c_1(y_2, \dots, y_N), \dots, y_N + c_N(y_1, \dots, y_{N-1}))$ , (3.12) and denoting it by  $T: \prod_{i \in N} Y_i \rightarrow \prod_{i \in N} Y_i$ , (3.13) we observe that  $T$  is a continuous mapping of  $\prod_{i \in N} Y_i$  into itself. Since each  $Y_i$  is a simplex of appropriate dimension (see Section 2.2), the product set  $\prod_{i \in N} Y_i$  becomes closed, bounded and convex. Hence, by Brouwer's fixed point theorem (see Appendix C), the mapping  $T$  has at least one fixed point. We now prove that every fixed point of this mapping is

necessarily a mixed-strategy Nash equilibrium solution of the game, and conversely that every mixed-strategy Nash equilibrium solution is a fixed point of  $T$ , thereby concluding the proof of the theorem. We first verify the latter assertion: If  $\{y_i^* \mid i \in N\}$  is a mixed-strategy equilibrium solution, then, by definition (i.e., from inequalities (3.11)), the function  $\phi_n(y_1^*, \dots, y_N^*)$  is nonpositive for every  $n \in M_j$  and  $i \in N$ , and thus  $c_{ij} = 0$  for all  $ij \in M_j, i \in N$ . This readily yields the conclusion  $T(y_1^*, \dots, y_N^*) = (y_1^*, \dots, y_N^*)$ , which is that  $\{y_i^* \mid i \in N\}$  is a fixed point of  $T$ . For verification of the former assertion, suppose that  $\{y_i^* \mid i \in N\}$  is a fixed point of  $T$ , but not a mixed equilibrium solution. Then, for some  $i \in N$  (say  $i = 1$ ) there exists a  $y_1 \in Y_1$  such that  $J_2(y_1, y_2^*, \dots, y_N^*) > J_2(y_1^*, y_2^*, \dots, y_N^*)$  (3-14)  $M_i M_N M_i M_N$  Now let  $h_i$  denote an index for which the quantity  $\phi_{h_i}(y_1, y_2^*, \dots, y_N^*) < \phi_{h_i}(y_1^*, y_2^*, \dots, y_N^*)$  (3-15)  $M_i M_j$ ;

**iV-PERSON NONZERO-SUM FINITE GAMES 93** attains its minimum value over  $n \in M_i$ . Then, since  $y_1$  is a mixed strategy, the RHS of (3.14) can be bounded below by (3.15) with  $m = h_i$ , thus yielding the strict inequality  $E - E(y_1, y_2^*, \dots, y_N^*) > E - E(y_1^*, y_2^*, \dots, y_N^*)$  which can further be seen to be equivalent to  $\phi_{h_i}(y_1, y_2^*, \dots, y_N^*) > 0$ . (3.16) This inequality, when used in the definition of  $c^i$ , implies that  $c^i > 0$ . But since  $c^i$  is nonnegative for all  $n \in M_i$ , the summation term  $\sum_{n \in M_i} Y_n c_n$  becomes positive. Now, again referring back to inequality (3.14), this time we let  $h_i$  denote an index for which (3.15) attains its maximum value over  $n \in M_i$ . Then, going through an argument similar to the one used in the preceding paragraph, we bound the LHS of (3.14) from above by (3.15) with  $n = h_i$ , and arrive at the strict inequality  $\phi_{h_i}(y_1, y_2^*, \dots, y_N^*) < 0$ . This then implies that  $c^i = 0$ , which, when used in (3.12) with  $i = 1$  and  $n_i = h_i$ , yields the conclusion  $\phi_{h_i}(y_1, y_2^*, \dots, y_N^*) < 0$ . But this contradicts the hypothesis that  $\{y_i^* \mid i \in N\}$  is a fixed point of  $T$ .  $\square$  The proof of Thm. 3.2, as given above, is not constructive. For a constructive proof, which also leads to a numerical scheme to obtain the equilibrium solution, the reader is referred to Scarf (1967). But, in general, it is not possible to obtain the equilibrium solution explicitly. If a given game admits an inner mixed-strategy equilibrium solution, however, then a possibility of obtaining the corresponding strategies in explicit form emerges, as the following proposition (which could also be considered as an extension of Remark 3.3 and Prop. 3.3) indicates. Proposition 3.5 Let  $Y_i$  denote the interior of  $Y_i$  with  $i \in N$ . Then, any inner mixed Nash equilibrium solution  $\{y_i^* \in Y_i \mid i \in N\}$  of an  $N$ -person finite static game in

normal form satisfies the set of equations  $\sum_{i=1}^N \lambda_i \sum_{j=1}^N v_{ij} x_j = \sum_{i=1}^N \lambda_i \sum_{j=1}^N v_{ij} x_j$  ( $\lambda_1, \dots, \lambda_N \sim a_1, a_2, \dots, a_N$ ) =  $\sum_{i=1}^N \lambda_i \sum_{j=1}^N v_{ij} x_j$ ,  $\lambda_i \geq 0$ ,  $\sum_{i=1}^N \lambda_i = 1$ ,  $\sum_{j=1}^N v_{ij} x_j = 0$   $\forall i=1, 2, \dots, N$   $\sum_{i=1}^N \lambda_i \sum_{j=1}^N v_{ij} x_j = 0$   $\forall i=1, 2, \dots, N$   $\sum_{i=1}^N \lambda_i \sum_{j=1}^N v_{ij} x_j = 0$   $\forall i=1, 2, \dots, N$   $\sum_{i=1}^N \lambda_i \sum_{j=1}^N v_{ij} x_j = 0$   $\forall i=1, 2, \dots, N$  (3.17)

94 T. BAŞAR AND G. J. OLSDER Furthermore, these mixed strategies also provide a mixed-strategy Nash equilibrium solution for the  $N$ -person game structured in a similar way but with some (or all) of the players maximizing their average costs instead of minimizing. Proof Since  $\{y_i^* \in Y_i; i \in N\}$  is an equilibrium solution,  $y_i^*$  minimizes the quantity  $E_{y_i^*} \sum_{j=1}^N v_{ij} x_j$  over  $Y_i$ , and with  $i \in N$ . Let us first take  $i = 1$ , and rewrite the preceding expression (in view of the relation  $\sum_{j=1}^N y_j^* = 1$ ) as  $\sum_{j=1}^N v_{1j} x_j$ ,  $i = 1$   $\sum_{j=1}^N v_{1j} x_j = \sum_{j=1}^N v_{1j} x_j$   $\sum_{j=1}^N v_{1j} x_j = \sum_{j=1}^N v_{1j} x_j$  Now, since the minimizing  $y_i^*$  is an inner point of  $Y_i$  by hypothesis, and since the preceding expression is linear in  $\{y_i^* \in Y_i; i \in N\}$ , it readily follows that the coefficient of  $y_i^*$  has to vanish for each  $i \in N$ ,  $i = 1, 2, \dots, N$ . This condition then yields the first set of equations of (3.17). The remaining ones can be verified analogously, by taking  $i = 2, 3, \dots, N$ . Finally, the last part of Prop. 3.5 follows, as in Prop. 3.3, from the property that at the equilibrium point  $\{y_i^* \in Y_i; i \in N\}$   $P_j$ 's average cost is actually independent of  $y_j^*$  (which has just been proven), and thus it becomes immaterial whether he minimizes or maximizes his average cost. D Since the hypothesis of Prop. 3.5 requires the mixed-strategy Nash equilibrium solution to be an inner point of the product set  $Y_1 \times \dots \times Y_N$ , the set of equations (3.17) is satisfied under quite restrictive conditions, and even more so under the conjecture of Problem 5 (Section 3.8), since then it is required that the number of alternatives of each player be the same. Nevertheless, the proposition still provides a characterization of the equilibrium solution, and enables direct computation of the corresponding strategies by merely solving coupled algebraic equations, under conditions which are not totally void. As an illustration of this point we now provide the following example. Example 3.5 Consider a 3-person game in which each player has two alternatives to choose from and in which the possible outcomes are given by  $P_1 = 1 (1, -1, 0) (2, 0, 0) (0, 1, 0) (0, 0, 1)$   $P_2 = 2 (1, 0, 1) (0, 3, 0) (0, 0, 0) (-1, 2, 0)$   $P_3 = 1 (1, -1, 0) (2, 0, 0) (0, 1, 0) (0, 0, 1)$   $P_3 = 2 (1, 0, 1) (0, 3, 0) (0, 0, 0) (-1, 2, 0)$   $P_3 = 1 (1, -1, 0) (2, 0, 0) (0, 1, 0) (0, 0, 1)$   $P_3 = 2 (1, 0, 1) (0, 3, 0) (0, 0, 0) (-1, 2, 0)$   $P_3 = 1 (1, -1, 0) (2, 0, 0) (0, 1, 0) (0, 0, 1)$   $P_3 = 2 (1, 0, 1) (0, 3, 0) (0, 0, 0) (-1, 2, 0)$

iV-PERSON NONZERO-SUM FINITE GAMES 95 Since  $M_j$  consists of two



elements for each  $i = 1, 2, 3$ , the set (3.17) involves only three equations for this game, which can be written as  $1 - V_1 - 2y_1 + 2/22/1 = 0$ ,  $2 - 2y_2 - 2y_3 + y_1y_2 = 0$ ,  $(3.18) \frac{2}{i+t/1-1} = 0$ . This coupled set of equations admits a unique solution with the property  $0 < y_i < 1$ , which is  $y_1^* = y_1/1 - \lambda$ ;  $t/22^* = 2 - v/3$ ;  $y_1^* = 1 - |V_5$ . Hence, this 3-person game admits a unique inner Nash equilibrium solution in mixed strategies, which is  $y_1^* = (2 - y/3, y_{fi} - 1)$ ,  $y_2^* = (V_5 - 1, 2 - V_5)$ ,  $y_3^* = (QV_3, 1 - \wedge)$ . It should be noted that this game admits also two pure-strategy Nash equilibria with outcomes  $(1, -1, 0)$  and  $(-1, 2, 0)$ .  $\square$

### 3.4 Computation of Mixed-Strategy Nash Equilibria in Bimatrix Games

We have discussed in Section 3.3 a possible approach toward obtaining mixed-strategy Nash equilibria of  $n$ -person games in normal form when these equilibria have the property of being inner solutions. In other words, when a mixed-strategy equilibrium solution assigns positive probabilities to all possible pure-strategy choices of a player, and this so for all players, then the corresponding equilibrium strategies can be determined by solving a set of algebraic equations (cf. Prop. 3.5). For the special case of bimatrix games, this requirement of Prop. 3.5 says that the pair of mixed Nash equilibrium strategies  $y^*$  and  $z^*$  should be elements of  $Y$  and  $Z$ , the interiors of  $Y$  and  $Z$ , respectively, which is a rather restrictive condition. If this condition is not satisfied, then one natural approach would be to set some components of the strategies  $y$  and  $z$  equal to zero and obtain algebraic equations for the other components, in the hope that the solutions of these algebraic equations will be nonnegative. If such a solution does not exist, then one can set some other components equal to zero and look for a nonnegative solution of the algebraic equation obtained from the other components, and so on. Since there is only a finite number of possibilities, it is obvious that such an approach will eventually yield mixed-strategy Nash equilibria of bimatrix games, also in view of Thm. 3.2. But, then, the natural question that comes to mind is whether this search can be done in a systematic way. Indeed it can, and this has been established by Lemke and Howson (1964) in their pioneering work. In that paper, the authors' intention has actually been to give an algebraic

96 T. BAŞAR AND G. J. OLSDER proof of the existence of mixed equilibrium solutions in bimatrix games (i.e., the result of Thm. 3.1), but, as a byproduct, they also obtained an efficient scheme for computing mixed Nash equilibria, which was thereafter referred to as the "Lemke-Howson" algorithm. The reader should consult the original work (Lemke and Howson, 1964) and the

expository article by Shapley (1974) for the essentials of this algorithm and its application in bimatrix games. Relation to nonlinear programming Yet another general method for the solution of a bimatrix game is to transform it into a nonlinear (in fact, a bilinear) programming problem, and to utilize the numerical techniques developed for solutions of nonlinear programming problems. In the sequel, we establish this equivalence between bimatrix games and a specific class of bilinear programming problems with linear constraints (see Prop. 3.6); we refer the reader to Luenberger (1973) for algorithms on the numerical solution of nonlinear programming problems. Proposition 3.6 A pair  $\{y^*, z^*\}$  constitutes a mixed-strategy Nash equilibrium solution to a bimatrix game  $(A, B)$  if, and only if, there exists a pair  $(p^*, q^*)$  such that  $\{y^*, z^*, p^*, q^*\}$  is a solution of the following bilinear programming problem:  $\min [y'Az + y'Bz + p + q]$  (3.19)  $y, z, p, q$  subject to  $Az > -plm, B'y > -qln, y > 0, z > 0, y'l m = 1, z'l n = 1$ . Proof. The constraints evidently imply that  $y'Az + y'Bz + p + q > 0$ , (i) which shows that the optimal value of the objective function is nonnegative. If  $\{y^*, z^*\}$  is an equilibrium pair, then the quadruple  $y^*, z^*, p^* = y^* A z^*, q^* = y^* B z^*$  (it) is feasible and the corresponding value of the objective function is zero. Hence (ii) is an optimal solution to the bilinear programming problem. Conversely, let  $\{y, z, p, q\}$  be a solution to the bilinear programming problem. The fact that feasible solutions to this problem exist has already been established by (ii). Specifically, by Thm. 3.2, the bimatrix game  $(A, B)$  has an equilibrium pair  $(y^*, z^*)$  and hence (ii) is a feasible solution. Furthermore, (ii) also yields the conclusion that the minimum value of the objective function is nonpositive, which, in view of (i), implies  $y'Az + y'Bz + p + q = 0$ . For any  $y > 0, z > 0$  such that  $y'l m = 1$  and  $z'l n = 1$ , we have  $y'Az > -p, y'Bz > -q$ , (3.20)

TV-PERSON NONZERO-SUM FINITE GAMES 97 and hence,  $y'Az + y'Bz > -p - q = y'Az + y'Bz$ . In particular, we have  $y'Az > -p, y'Bz > -q, y'Az + y'Bz = p + q$ . Thus  $y'Az = -p, y'Bz = -q$ , and therefore,  $y'Az > y'Az, y'Bz > y'Bz$ . This last set of inequalities verifies that  $\{y, z\}$  is indeed a mixed-strategy Nash equilibrium solution of the bimatrix game  $(A, B)$ . D For the special case of zero-sum matrix games, the following corollary now readily follows from Prop. 3.6, and establishes an equivalence between two- person zero-sum matrix games and LP problems. Corollary 3.2 A pair  $\{y^*, z^*\}$  constitutes a mixed-strategy saddle-point solution for a two-person zero-sum matrix game  $A$  if and only if, there exists a pair  $(p^*, q^*)$  such that  $\{y^*, z^*, p^*, q^*\}$  is a solution of the LP problem  $\min [p + q]$  (3.21)  $y, z, p, q$  subject to  $Az > -plm, A'y < qln, "I . .$

$y > 0, z > 0, y'_{lm} = 1, z'_{ln} = 1. j \{ ' \}$  Remark 3.4 Note that Prop. 3.6 directly extends to TV-person finite games in normal form. Furthermore, it is noteworthy that the LP problem of Corollary 3.2 is structurally different from the LP problem considered in Section 2.3 (Thm. 2.5), which also provided a means of obtaining mixed-strategy saddle-point solutions for matrix games.  $\square$

**3.5 Nash Equilibria of TV-Person Games in Extensive Form** This section is devoted to noncooperative (Nash) equilibria of TV-person finite games in extensive form, which do not incorporate chance moves. Extensive tree formulation for finite games without chance moves has already been introduced in Chapter 2 within the context of zero-sum games (cf. Def. 2.5), and such a formulation is equally valid for finite nonzero-sum games, with certain appropriate modifications. Hence first, as a direct extension of Def. 2.5, we have the following definition which covers both single-act and multi-act games.

**Definition 3.10** An extensive form of an  $N$ -person nonzero-sum finite game without chance moves is a tree structure with (i) a specific vertex indicating the starting point of the game,

98 T. BAŞAR AND G. J. OLSDER (ii)  $N$  cost functions, each one assigning a real number to each terminal vertex of the tree, where the  $i$ th cost function determines the loss to be incurred to  $P_i$ , (Hi) a partition of the nodes of the tree into  $N$  player sets, (iv) a subpartition of each player set into information sets  $\{/?\}$ , such that the same number of branches emanates from every node belonging to the same information set and no node follows another node in the same information set.

**Figure 3.1: Two typical nonzero-sum finite games in extensive form.** Two typical nonzero-sum finite games in extensive form are depicted in Fig. 3.1. The first one represents a 3-player single-act nonzero-sum finite game in extensive form in which the information sets of the players are such that both  $P_2$  and  $P_3$  have access to the action of  $P_1$ . The second extensive form of Fig. 3.1, on the other hand, represents a 2-player multi-act nonzero-sum finite game in which  $P_1$  acts twice and  $P_2$  only once. In both extensive forms, the set of alternatives for each player is the same at all information sets and it consists of two elements. The outcome corresponding to each possible path is denoted by an ordered TV-tuple of numbers  $(a_1, \dots, a_N)$ , where  $N$  stands for the number of players and  $a_1$  stands for the corresponding cost to  $P_i$ .

**Pure-strategy Nash equilibria** We now introduce the concept of a noncooperative (Nash) equilibrium solution for TV-person nonzero-sum finite games in extensive form, which covers both single-act and multi-act games. To this end,

let us first recall the definition of a strategy from Chapter 2 (cf. Def. 2.6).  
 Definition 3.11 Let  $\mathcal{N}_i$  denote the class of all information sets of  $P_i$ , with a typical element designated as  $r_i$ . Let  $A_i$  denote the set of alternatives of  $P_i$

**iV-PERSON NONZERO-SUM FINITE GAMES** 99 at the nodes belonging to the information set  $r_i$ . Define  $U_i = \bigcup_{r_i \in \mathcal{N}_i} U_i(r_i)$ , where the union is over  $r_i \in \mathcal{N}_i$ . Then, a strategy  $\gamma_i$  for  $P_i$  is a mapping from  $\mathcal{N}_i$  into  $U_i$ , assigning one element in  $U_i$  for each set in  $\mathcal{N}_i$ , and with the further property that  $\gamma_i(r_i) \in A_i$  for each  $r_i \in \mathcal{N}_i$ . The set of all strategies of  $P_i$  is called his strategy set (space), and it is denoted by  $P_i$ . Let  $J_i(\gamma_1, \dots, \gamma_N)$  denote the loss incurred to  $P_i$  when the strategies  $\gamma_1 \in P_1, \dots, \gamma_N \in P_N$  are adopted by the players. Then, the noncooperative (Nash) equilibrium solution concept for such games can be introduced as follows, as a direct extension of Def. 3.7.

Definition 3.12 An  $N$ -tuple of strategies  $\{\gamma_1^*, \gamma_2^*, \dots, \gamma_N^*\}$  with  $\gamma_i^* \in P_i, i \in N$ , is said to constitute a noncooperative (Nash) equilibrium solution for an  $N$ -person nonzero-sum finite game in extensive form, if the following  $N$  inequalities are satisfied for all  $\gamma_i^* \in P_i, i \in N$ :  $J_i(\gamma_1^*, \gamma_2^*, \dots, \gamma_N^*) \leq J_i(\gamma_i, \gamma_2^*, \dots, \gamma_N^*)$  for all  $\gamma_i \in P_i$ . The  $N$ -tuple of quantities  $\{J_1^*, \dots, J_N^*\}$  is known as a Nash equilibrium outcome of the nonzero-sum finite game in extensive form. We emphasize the word a in the last sentence of the preceding definition, since the Nash equilibrium solution could possibly be nonunique with the corresponding set of Nash values being different. This then leads to a partial ordering in the set of all Nash equilibrium solutions, as in the case of Defs. 3.2 and 3.3 whose extensions to the  $iV$ -person case are along similar lines and are therefore omitted.

Nash equilibria in mixed and behavioral strategies The concepts of noncooperative equilibria in mixed and behavioral strategies for nonzero-sum finite games in extensive form can be introduced as straightforward (natural) extensions of the corresponding ones presented in Chapter 2 within the context of saddle-point equilibria. We should recall that a mixed strategy for a player ( $P_i$ ) is a probability distribution on the set of all his pure strategies, i.e., on  $P_i$ . A behavioral strategy, on the other hand, is an appropriate mapping whose domain of definition is the class of all the information sets of the player. By denoting the behavioral strategy set of  $P_i$  by  $B_i$ , and the average loss incurred to  $P_i$  as a result of adoption of the behavioral strategy  $iV$ -tuple  $\{\gamma_1 \in B_1, \dots, \gamma_N \in B_N\}$  by  $J_i(\gamma_1, \dots, \gamma_N)$ , the definition of a Nash equilibrium (3.23)

100 T. BAŞAR AND G. J. OLSDER solution in behavioral strategies may be obtained directly from Def. 3.12 by replacing  $\Gamma$ ,  $P$  and  $J_1$  by  $\Gamma'$ ,  $P$  and  $J \setminus$  respectively. We now discuss, in the four subsections to follow, properties and derivation of these different types of Nash equilibria for finite games in extensive form—first for single-act games, and then for multi-act games.

### 3.5.1 Single-act games: Pure-strategy Nash equilibria

In single-act games, each player acts only once, and the order in which the players act could be a variant of their strategies. If the order is fixed a priori, and further if the single-act game is static in nature (i.e., if each player has a single information set), then there is no basic difference between extensive and normal form descriptions, and consequently no apparent advantage of one form over the other. Since such static games in normal form have already been extensively studied in Sections 3.2 and 3.3, our concern here will be primarily on dynamic single-act games, that is, games in which at least one of the players has access to some nontrivial information concerning the action of some other player. Figure 3.1(a), for example, displays one such game wherein both  $P_2$  and  $P_3$  have access to  $P_1$ 's action, but they are ignorant about each other's actions. Yet another single-act game with dynamic information is the 2-person nonzero-sum game depicted in Fig. 3.2, wherein  $P_2$ 's information sets can differentiate between whether  $P_1$  has played  $L$  or not, but not between his possible actions  $M$  and  $R$ .

Figure 3.2: A 2-person nonzero-sum single-act game with dynamic information. One method of obtaining Nash equilibria of a dynamic single-act finite game in extensive form is to transform it into an equivalent normal form and to make use of the theory of Sections 3.2 and 3.3. This direct method has already been discussed earlier in Section 2.4 within the context of similarly structured zero-sum games, and it has the basic drawback that one has to take into consideration all possible strategy combinations. As in the case of dynamic zero-sum games, an alternative to this direct method exists for some special types of nonzero-sum games, which makes explicit use of the extensive form description, and obtains a subclass of all Nash equilibria through a systematic (recursive) procedure. But

IV-PERSON NONZERO-SUM FINITE GAMES 101 in order to aid in understanding the essentials and the limitations of such a recursive procedure, it will be instructive first to consider the following specific example. This simple example, in fact, displays most of the intricacies of Nash equilibria in dynamic finite games. Example 3.6 Consider the 2-person nonzero-sum single-

act game whose extensive form is displayed in Fig. 3.2. The following is an intuitively appealing recursive procedure that would generate a Nash equilibrium solution for this game. At the first of his two information sets (counting from left), P2 can tell precisely whether P1 has played L or not. Hence, in case  $u_1 = L$ , the unique decision of P2 would be  $u_2 = L$ , yielding him a cost of  $-1$  which is in fact the lowest possible level P2 can hope to attain. Now, if  $u_1 = M$  or  $R$ , however, P2's information set does not differentiate between these two actions of P1, thus forcing him to play a static nonzero-sum game with P1. The corresponding bimatrix game (under the usual convention of A denoting the loss matrix of P1 and B denoting that of P2) is

$$P2 \begin{matrix} & L & R \\ A = \begin{matrix} M & R & 0 \\ P1, & B = \begin{matrix} M & R & L & R \end{matrix} \end{matrix} \end{matrix} \quad (3.24)$$

This bimatrix game admits (as indicated) a unique Nash equilibrium  $\{L, R\}$  with an equilibrium cost pair of  $(-1, 0)$ . This then readily suggests  $u_2 = L$  if  $u_1 = L$ , otherwise (3.25a) as a candidate equilibrium strategy for P2. P1, on the other hand, has two possible strategies,  $u_1 = L$  and  $u_1 = R$ , since the third possibility  $u_1 = M$  is ruled out by the preceding argument. Now, the former of these leads (under (3.25a)) to a cost of  $J_1 = 0$ , while the latter leads to  $J_1 = -1$ . Hence, he would clearly decide on playing R, i.e.,  $u_1^* = R$  (3.25b) is a candidate equilibrium strategy for P1. The cost pair corresponding to (3.25a)-(3.25b) is  $J_1^* = -1, J_2^* = 0$ . (3.25c) The strategy pair (3.25a)-(3.25b) is thus what an intuitively appealing (and rather straightforward) recursive procedure would provide us with, as a reasonable noncooperative equilibrium solution for this extensive form. It can, in fact, be directly verified by referring to inequalities (3.23) that the pair (3.25a)-(3.25b) is indeed in Nash equilibrium. To see this, let us first fix P2's strategy at (3.25a), and observe that (3.25b) is then the unique cost-minimizing decision for P1. Now, fixing P1's strategy at (3.25b), it is clear that P2's unique

102 T. BAŞAR AND G. J. OLSDER cost-minimizing decision is  $u_2 = R$  which is indeed implied by (3.25a) under (3.25b). The recursive scheme adopted surely resulted in a unique strategy pair which is also in Nash equilibrium. But is this the only Nash equilibrium solution that the extensive form of Fig. 3.2 admits? The reply is, in fact, no! To obtain the additional Nash equilibrium solution(s), we first transform the extensive form of Fig. 3.2 into an equivalent normal form. To this end, let us first note that  $T_1 = \{L, M, R\}$  and  $T_2 = \{72, 722, 732, 742\}$ , where  $72(i^2) = L, 72V) = R, 71(^2) \blacksquare \{$  if  $u_1 = L$ , otherwise  $74V) \{$  otherwise Hence, the equivalent normal form is the  $3 \times 4$

bimatrix game  $P_2$   $P_2$   $A = \begin{matrix} L & M & R \\ 0 & 3 & 2 \\ -2 & 0 & -1 \end{matrix}$   $Q = \begin{matrix} -2 & 3 & 2 \\ 1 & 1 & 3 \\ 0 & -1 & 3 \end{matrix}$   $P_I$ ,  $B = \begin{matrix} L & M & R \\ -1 & 2 & 1 \\ 1 & 1 & 3 \end{matrix}$   $P_I$ , which admits two pure-strategy Nash equilibria, as indicated. The encircled one is  $\{R; i\%$  which corresponds to the strategy pair (3.25a)-(3.25b) obtained through the recursive procedure. The other Nash equilibrium solution is the constant strategy pair  $(V) = L, (V) = L$ , with a corresponding cost pair of  $J_1 = 0, J_2 = -1$ . (3.26a) (3.26b) (3.26c) A comparison of (3.25c) and (3.26c) clearly indicates that both of these equilibria are admissible. It is noteworthy that, since (3.26a) is a constant strategy, the pair (3.26a)- (3.26b) also constitutes a Nash equilibrium solution for the static single-act game obtained from the one of Fig. 3.2 by replacing the dynamic information of  $P_2$  with static information (i.e., by allowing him a single information set). It is, in fact, the unique Nash equilibrium solution of this static game which admits the normal form (bimatrix) description  $P_2$   $P_2$   $A = \begin{matrix} L & M & R \\ 0 & 3 & 2 \\ -2 & 0 & -1 \end{matrix}$   $P_I$ ,  $B = \begin{matrix} L & M & R \\ -1 & 2 & 1 \\ 1 & 1 & 3 \end{matrix}$   $P_I$ . The actual play (R,R) dictated by (3.25a)-(3.25b), however, does not possess such a property as a constant strategy pair. D

TV-PERSON NONZERO-SUM FINITE GAMES 103 The preceding example has displayed certain important features of Nash equilibria of a single-act game, which we now list below. We will shortly see, in this section, that these features are, in fact, valid on a broader scale for both single-act and multi-act nonzero-sum games. Features of the single-act game of Example 3.6 (i) The single-act game admitted multiple (two) Nash equilibria. (ii) The (unique) Nash equilibrium solution of the static version of the single- act game (obtained by replacing the information sets of the second-acting player by a single information set) also constituted a Nash equilibrium solution for the original single-act game with the dynamic information. (iii) A recursive procedure that involves solution of only bimatrix games (which could also be degenerate games) at each information set yielded only one of the two Nash equilibria. The actual play dictated by this equilibrium solution did not constitute an equilibrium solution as constant strategies. We now show that appropriate (and more general) versions of these features are retained in a more general framework for  $N$ -person single-act games. To this end, we first introduce some terminology and also a partial ordering of extensive forms in terms of their information sets. Definition 3.13 A single-act  $N$ -person finite game in extensive form (say,  $I$ ) is the static version of a dynamic single-act  $N$ -person finite game in extensive form (say,  $I^1$ ), if  $I$  can be obtained from  $I^1$  by

replacing the information sets of each player with a single information set encompassing all the nodes pertaining to that player.<sup>16</sup> The equivalent normal form of the static game I is called the static normal form of II. Definition 3.14 Let I and II be two single-act N-person games in extensive form, and further let  $F^i$  and  $T^i$  denote the strategy sets of  $P_i$  ( $i \in N$ ) in I and II, respectively. Then, I is said to be informationally inferior to II if  $T^i \subset T^{jj}$  for all  $i \in N$ , with strict inclusion for at least one  $i$ . Remark 3.5 In Fig. 3.3, the single-act game I is informationally inferior to II and III. The latter two, however, do not admit any such comparison. □ Proposition 3.7 Let I be an N-person single-act game that is informationally inferior to some other single-act N-person game, say II. Then, (i) any Nash equilibrium solution of I also constitutes a Nash equilibrium solution for II, <sup>16</sup>It should be noted that a single-act finite game admits a static version only if the order in which the players act is fixed a priori and the possible actions of each player are the same at all of his information sets.

104 T. BAŞAR AND G. J. OLSDER 1,2 0,1 0,0 -1,2 2,-1 1,0 1,2 0,1 0,0 -1,2 2,-1 1,0 1,2 0,1 0,0 -1,2 2,-1 1,0 III Figure 3.3: Extensive forms displaying informational inferiority. (ii) if  $\{7^1, \dots, 7^N\}$  is a Nash equilibrium solution of II so that  $Y^i \in T^i$  for all  $i \in N$ , then it also constitutes a Nash equilibrium solution for I. Proof (i) If  $\{7^1, \dots, 7^N\}$  constitutes a Nash equilibrium solution for I, then inequalities (3.23) are satisfied for all  $7^i \in F^i$ ,  $i \in N$ . But, since  $T^i \subset T^{jj}$ ,  $i \in N$ , we clearly also have  $7^i \in T^{jj}$ ,  $i \in N$ . Now assume, to the contrary, that  $\{7^1, \dots, 7^N\}$  is not a Nash equilibrium solution of II. Then, this implies that there exists at least one  $i$  (say,  $i = N$ , without any loss of generality) for which the corresponding inequality of (3.23) is not satisfied for all  $7^i \in F^i$ . In particular, there exists a  $j^N \in T^N$  such that  $J^N(j^N, \dots, j^{N-1}, \dots, j^1) > J^N(7^1, \dots, 7^{N-1}, 7^N)$ . (t) Now, the  $N$ -tuple of strategies  $\{7^1, \dots, 7^{N-1}, 7^N\}$  leads to a unique path of action, and consequently to a unique outcome, in the single-act game II. Let us denote the information set of  $P^N$ , which is actually traversed by this path, by  $f^N$ , and the specific element (node) of  $f^N$  (intercepted by  $n^N$ ). Let us further denote the information set of  $P^N$  in game I, which includes the node  $n^N$ , by  $f^N$ . Then, there exists at least one element in  $F^N$  (say,  $7^N$ ) with the property  $J^N(7^1, \dots, 7^{N-1}, 7^N) > J^N(j^1, \dots, j^{N-1}, 7^N)$ . If this strategy replaces  $j^N$  on the RHS of inequality (i), the value of  $J^N$  clearly does not change, and hence we equivalently have  $J^N(7^1, \dots, 7^{N-1}, 7^N) > J^N(7^1, \dots, 7^{N-1}, 7^N)$ . But this inequality contradicts the initial hypothesis that the  $N$ -tuple of policies  $\{7^1, \dots, j^N\}$  was in Nash equilibrium for the game I.



This then completes the proof of part (i). (ii) Part (ii) of the proposition can be proven analogously. D

iV-PERSON NONZERO-SUM FINITE GAMES 105 Remark 3.6 Proposition 3.7 now verifies a general version of feature (ii) of the single-act game of Example 3.6 for  $N$ -person single-act finite games in extensive form. A counterpart of feature (i) also readily follows from this proposition, which is that such games with dynamic information will in general admit multiple Nash equilibria. A more definite statement cannot be made, since there would always be exceptional cases when the Nash equilibrium solution of a single-act dynamic game is unique and is attained in constant strategies which definitely also constitute a Nash equilibrium solution for its static version, assuming that it exists. It is almost impossible to single out all these special games, but we can comfortably say that they are "rarely" met, and existence of multiple Nash equilibria is a rule in dynamic single-act games rather than an exception. Since this sort of nonuniqueness emerges mainly because of the dynamic nature of the information sets of the players, we call it informational nonuniqueness. Q Now, to elaborate on the third (and the last) feature listed earlier, we have to impose some further structure on the "relative nestedness" of the information sets of the players, which is already implicit in the extensive form of Fig. 3.2. It is, in fact, possible to develop a systematic recursive procedure, as an extension of the one adopted in the solution of Example 3.6, if the order in which the players act is fixed a priori and the extensive form is in "ladder-nested" form—a notion that we introduce below. Definition 3.15 In an extensive form of a single-act nonzero-sum finite game with a fixed order of play, a player  $P_i$  is said to be a precedent of another player  $P_j$  if the former is situated closer to the vertex of the tree than the latter. The extensive form is said to be nested if each player has access to the information acquired by all his precedents. If, furthermore, the only difference (if any) between the information available to a player ( $P_i$ ) and his closest (immediate) precedent (say  $P_{i-1}$ ) involves only the actions of  $P_{i-1}$ , and only at those nodes corresponding to the branches of the tree emanating from singleton information sets of  $P_{i-1}$ , and this so for all players, the extensive form is said to be ladder-nested.<sup>17</sup> A single-act nonzero-sum finite game is said to be nested (respectively, ladder-nested) if it admits an extensive form that is nested (respectively, ladder-nested). Remark 3.7 The single-act extensive forms of Figs. 3.1(a) and 3.2 are both ladder-nested. If the extensive form of Fig. 3.1(a)

is modified so that both nodes of P2 are included in the same information set, then it is only nested, but not ladder-nested, since P3 can differentiate between different actions of P1 but P2 cannot. Finally, if the extensive form of Fig. 3.1(a) is modified so that this time P3 has a single information set (see Fig. 3.4(a)), then the resulting extensive form becomes non-nested, since then even though P2 is a precedent of P3 he actually knows more than P3 does. The single-act game that this

17Note that in 2-person single-act games the concepts of "nestedness" and "ladder-nestedness" coincide, and every extensive form is, by definition, ladder-nested.

106 T. BAŞAR AND G. J. OLSDER Figure 3.4: Two extensive forms of the same nested single-act game. extensive form describes is, however, nested since it also admits the extensive form description depicted in Fig. 3.4(b). D One advantage of dealing with ladder-nested extensive forms is that they can recursively be decomposed into simpler tree structures which are basically static in nature. This enables one to obtain a class of Nash equilibria of such games recursively, by solving static games at each step of the recursive procedure. Before providing the details of this recursive procedure, let us introduce some terminology. Definition 3.16 For a given single-act dynamic game in nested extensive form (say,  $I$ ), let  $r$  denote a singleton information set of  $P_i$ 's immediate follower (say  $P_j$ ); consider the part of the tree structure of  $I$ , which is cut off at  $r$ , has  $T$  as its vertex and has as immediate branches only those that enter into that information set of  $P_j$ . Then, this tree structure is called a sub-extensive form of  $I$ . (Here, we adopt the convention that the starting vertex of the original extensive form is the singleton information set of the first-acting player.) Remark 3.8 The single-act game depicted in Fig. 3.2 admits two sub-extensive forms which are as follows: The extensive form of Fig. 3.1(a), on the other hand, admits a total of four sub-extensive forms which we do not display here. It should be noted that each sub-extensive form is itself an extensive form describing a simpler game. The

AT-PERSON NONZERO-SUM FINITE GAMES 107 first one displayed above describes a degenerate 2-player game in which P1 has only one alternative. The second one again describes a 2-player game in which the players each have two alternatives. Both of these sub-extensive forms will be called static since the first one is basically a one-player game and the second one describes a static 2-player game. A precise definition follows. D

Definition 3.17 A sub-extensive form of a nested extensive form of a single-act game is static if every player appearing in this tree structure has a single information set. We are now in a position to extend the recursive procedure adopted in Example 3.6 to obtain (some of) the Nash equilibrium solutions of  $n$ -person single-act games in ladder-nested extensive form. A recursive procedure to determine pure-strategy Nash equilibria of  $n$ -person single-act games in ladder-nested extensive form

- (1) For each fixed information set of the last-acting player (say,  $P_N$ ), single out the players who have precisely the same information as  $P_N$ , determine the static sub-extensive form that includes all these players and their single information sets, and solve the static game that this sub-extensive form describes. Assuming that each of these static games admits a unique pure-strategy Nash equilibrium solution, record the equilibrium strategies and the corresponding Nash outcomes by identifying them with the players and their information sets. (If any one of the static games admits more than one Nash equilibrium solution, then this procedure as well as the following steps are repeated for each of these multiple equilibria.)
- (2) Replace each static sub-extensive form considered at step 1 with the immediate branch emanating from its vertex that corresponds to the Nash strategy of the starting player of that sub-extensive form; furthermore, attach the corresponding  $n$ -tuple of Nash equilibrium values to the end of this branch.
- (3) For the remaining game in extensive form, repeat steps 1 and 2 until the extensive form left is a tree structure comprising only a vertex and some immediate branches with an  $n$ -tuple of numbers attached to each of them.
- (4) The branch of this final tree structure which corresponds to a minimum loss for the starting player is his Nash equilibrium strategy, and the corresponding  $n$ -tuple of numbers at the end of this branch determines a Nash outcome for the original game. The corresponding Nash equilibrium strategies of the other players in the original single-act game can then be determined from the solutions of the static sub-extensive forms, by appropriately identifying them with the players and their information sets.

108 T. BAŞAR AND G. J. OLSDER Remark 3.9 Because of the ladder-nestedness property of the single-act game, it can readily be shown that every solution obtained through the foregoing recursive procedure is indeed a Nash equilibrium solution of the original single-act game. A most natural question to raise now is whether the preceding recursive procedure generates all the Nash equilibrium solutions of a given  $n$ -person single-act game in ladder-

nested extensive form. We have already seen in Example 3.6 that this is, in fact, not so, since out of the two Nash equilibrium solutions, only one of them was generated by this recursive procedure. Then, the question may be rephrased as: precisely what subclass of Nash equilibria is generated by the preceding scheme? In other words, what common feature (if any) can we attribute to the Nash equilibria that can be derived recursively? To shed some light on this matter, it will be instructive to refer again to Example 3.6 and to take a closer look at the two Nash equilibria obtained: EXAMPLE 3.6 (continued) For the single-act game of Fig. 3.2, the recursive procedure generated the solution (3.25a)-(3.25b) with Nash cost pair  $(-1, 0)$ . When compared with the other Nash cost pair which is  $(0, -1)$ , it is clear that the former is more advantageous for P1. Hence, P2 would rather prefer to play the constant strategy (3.26a) and thereby attain a favorable cost level of  $J_2 = -1$ . But, since he acts later, the only way he can ensure such a cost level is by announcing his constant strategy ahead of time and by irrevocably sticking to it. If he has the means of doing this, and further if P1 has strong reasons to believe in such an attitude on the part of P2, then  $(0, -1)$  will clearly be the only cost pair to be realized. Hence, the constant-strategy Nash solution in this case corresponds to a "prior commitment" mode of play on the part of the last acting player—a notion which has been introduced earlier in Section 2.5 within the context of multi-act zero-sum games, and which also has connections with the Stackelberg mode of play to be discussed in the next section of this chapter. Now, yet another mode of play for P2 (the second-acting player) would be to wait until he observes at what information set he is at the time of his play, and only then decide on his action. This is known as the "delayed commitment" type of attitude, which has also been introduced earlier in Section 2.5, within the context of zero-sum games and in conjunction with behavioral strategies. For the single-act game of Example 3.6, a Nash equilibrium solution of the delayed commitment type is the one given by (3.25a)-(3.25b), which has been obtained using the recursive procedure. D This discussion now leads us to the following definition. Definition 3.18 A Nash equilibrium solution of a ladder-nested single-act finite N-person game is of the delayed commitment type if it can be obtained through the recursive procedure outlined earlier, i.e., by solving only static single-act games.

AT-PERSON NONZERO-SUM FINITE GAMES 109 Remark 3.10 As a side remark, it is noteworthy that in the single-act game of Example 3.6 additional

information for P2 concerning the action of P1 could be detrimental. For, if his information set was a single one (i.e., the static version), then the game would admit the unique Nash solution (3.26a)-(3.26b) with P2's Nash cost being  $J_2 = -1$ . In the set-up of the extensive form of Fig. 3.2, however, he could end up with a higher Nash cost of  $J_2 = 0$ . Hence, in nonzero-sum games, additional information could be detrimental for the player who receives it.  $\square$  We now illustrate, giving two examples, the steps involved in the outlined recursive procedure to obtain admissible Nash equilibrium strategies of the delayed commitment type. Example 3.7 Consider the 3-person single-act game whose extensive form (of the ladder-nested type) is depicted in Fig. 3.1(a). Here, both P2 and P3 have access to the action of P1 (and to no more information), and hence they are faced with a bimatrix game at each of their information sets. Specifically, if  $u_1 = L$ , then the corresponding bimatrix game is  $\begin{matrix} & R & L \\ P_2 & 0 & 1 \\ P_3 & 1 & 0 \end{matrix}$  where the entries of matrix A denote the possible losses to P2 and the entries of matrix B denote the possible losses to P3. This bimatrix game admits a unique admissible Nash equilibrium solution which is  $\{u_2^* = L, u_3^* = L\}$ , with the corresponding Nash values being  $-1$  and  $0$  for P2 and P3, respectively. If  $u_1 = R$ , on the other hand, the bimatrix game of interest is  $\begin{matrix} & R & L \\ P_2 & 1 & 0 \\ P_3 & 0 & 2 \end{matrix}$ . This game admits a unique Nash solution which is  $\{u_2^* = R, u_3^* = L\}$  with a Nash cost pair of  $(-2, -1)$ . Hence, regardless of what P1 plays, the delayed commitment type admissible Nash equilibrium strategies of P2 and P3 in this ladder-nested game are unique, and they are given as  $u_2^*(r, 2) = u_3^*(r, 2) = L$ . (Note that P3's equilibrium strategy is a constant.) Now, if P1 picks  $u_1 = L$ , his loss under (i) will be  $1$ , but otherwise his loss is  $0$ . Hence he also has a unique equilibrium strategy  $u_1^* = R$ . (ii) Strategies (i) and (ii) now constitute the unique delayed commitment type admissible Nash equilibrium solution of the single-act game of Fig. 3.1(a), which

110 T. BAŞAR AND G. J. OLSDER can also be directly verified by again referring to the set of inequalities (3.23). The corresponding Nash equilibrium outcome is  $(0, -2, -1)$ . Q Figure 3.5: A 4-person single-act ladder-nested game in extensive form. Example 3.8 Consider the 4-person single-act game whose extensive form (of the ladder-nested type) is depicted in Fig. 3.5. The last-acting player is P4 and he has three information sets. The first one (counting from left) tells him whether P1 has actually played L or not, but nothing concerning the actions of the other players. The same is true for P2 and

P3. The corresponding three-player static game admits an equivalent normal form which is the one considered in Example 3.3, with only  $P_j$  replaced by  $P_j + 1$  in the present context. Furthermore, we now also have the costs of an additional player (PI) showing up at the end of the branches (i.e., in the present context we have a quadruple cost instead of a triple). It is already known that this game admits a unique Nash equilibrium solution which, under the present convention, is  $u_2^* = L$ ,  $u_3^* = L$ ,  $u_4^* = L$ , and this is so if PI has picked  $u_1 = L$ , with the corresponding cost quadruple being  $(1, 1, -1, 0)$ . The second information set of P4 tells him whether PI has picked R and P2 has picked L; and P3 also has access to this information. Hence, we now have a bimatrix game under consideration, which is essentially the first bimatrix game of Example 3.6, admitting the unique solution  $u_3^* = L$ ,  $u_4^* = L$ , and this is so if  $u_1 = R$ ,  $u_2 = L$ , with the corresponding cost quadruple being  $(0, 1, -1, 0)$ . The third information set of P4 provides him with the information  $u_1 = R$ ,  $u_2 = R$ . P3 has access to the same information, and thus the equivalent normal

form is the second bimatrix game of Example 3.7, which admits the unique Nash solution  $u_3^* = R$ ,  $u_4^* = L$ , with the corresponding cost quadruple being  $(0, 0, -2, 1)$ . Using all these, we can now write down the unique Nash equilibrium strategies of P3 and P4, which are given as  $(u_3, u_4) = (R, L)$ . («) Deletion of all the branches of the extensive tree already used now leads to the simpler tree structure depicted below:  $(u_1, u_2) = (L, R)$  otherwise,  $(u_1, u_2) = (R, R)$ . In this game, P2 acts only if  $u_1 = R$ , in which case he definitely picks  $u_2 = R$ , since this yields him a lower cost (0) as compared with 1 which he would obtain otherwise. This also completely determines the delayed commitment type Nash equilibrium strategy of P2 in the game of Fig. 3.5, which is unique and is given by  $(u_2, u_3) = (R, L)$ . (Hi) Then, the final equivalent form that the game takes is  $(u_1, u_2, u_3, u_4) = (L, R, L, L)$  from which it is clear that the optimal unique strategy for PI is  $u_1^* = R$ . (iv) The strategies (i)-(iv) constitute the unique delayed commitment type Nash equilibrium solution of the extensive form of Fig. 3.8 (which is clearly also admissible), and the reader is also encouraged to check this result by verifying satisfaction of inequalities (3.23). The corresponding unique Nash equilibrium outcome is  $(0, 0, -2, 1)$ . □

112 T. BAŞAR AND G. J. OLSDER Nested extensive forms which are not ladder-nested cannot be decomposed into static sub-extensive forms, and

hence the recursive derivation does not directly apply to such games. However, some nested extensive forms still admit a recursive decomposition into simpler sub-extensive forms, some of which will be dynamic in nature. Hence, an extended version of the recursive procedure applies for nested games, which involves solution of simpler dynamic sub-extensive forms. In this context, the delayed commitment mode of play also makes sense as to be introduced in the sequel.

**Definition 3.19** A nested extensive (or sub-extensive) form of a single-act game is said to be undecomposable if it does not admit any simpler sub-extensive form. It is said to be dynamic, if at least one of the players has more than one information set.

**Definition 3.20** For an  $N$ -person single-act dynamic game in nested undecomposable extensive form  $(I)$  let  $J$  denote the set of games which are informationally inferior to  $I$ . Let  $\{s_1^*, \dots, s_N^*\}$  be an  $N$ -tuple of strategies in Nash equilibrium for  $I$ , and further let  $j^*$  denote the number of games in  $J$  to which this  $N$ -tuple provides an equilibrium solution. Then,  $\{s_1^*, \dots, s_N^*\}$  is of delayed commitment type if there exists no other  $N$ -tuple of Nash equilibrium strategies of  $I$  which constitutes an equilibrium solution to a smaller (than  $j^*$ ) number of games in  $J$ .

The following example now provides an illustration of these notions as well as a derivation of delayed commitment type Nash equilibria in nested games.

**Example 3.9** Consider the 3-person single-act game whose extensive form is displayed in Fig. 3.6. In this game  $P_3$  has access to  $P_1$ 's action, but  $P_2$  does not. Hence, the game is nested but not ladder-nested. Furthermore, it is both dynamic and undecomposable, with the latter feature eliminating the possibility for a recursive derivation of any of its Nash equilibria. Then, the only method is to bring the single-act game into equivalent normal form and to obtain the Nash equilibria of this normal form. Toward this end, let us first note that  $P_1$  and  $P_2$  each have two possible strategies:  $\{s_1 = L, s_1 = R\}$ .  $P_3$ , on the other hand, has four possible strategies:  $(s_3 = L, s_3 = R), (s_3 = L, s_3 = R)$  and  $(s_3 = L, s_3 = R)$  otherwise. The reader should verify that every single-act game that admits Nash equilibria in pure strategies has at least one equilibrium solution of the delayed commitment type, i.e., the set of such equilibria is not empty unless the game does not admit any pure-strategy Nash equilibrium solution.

AT-PERSON NONZERO-SUM FINITE GAMES 113 Figure 3.6: A nested single-act game in extensive form which is dynamic and undecomposable. Using the terminology of Section 3.3, we now display the normal form as two

$(2 \times 4)$  matrices, one for each strategy of  $P_i$ :  $n \setminus = 1 : 1 \ 2 \ (0, 0, 0) \ (1, 2, 1) \ P_3$   
 $(1, -1, 1) \ (1, 1, 1) \ (0, 0, 0)^* \ (1, 2, 1) \ (1, -1, 1) \ (1, 1, 1) \ P_2, \ n_i \ 2: \ 1 \ 2 \ (-1, 1, 2) \ (1, 0, 1)$   
 $P_3 \ (1, 0, -1)^* \ (-1, 2, 0) \ (1, 0, -1) \ (-1, 2, 0) \ (-1, 1, 2) \ (1, 0, 1) \ P_2. \ 1$  Here, the  
integers  $1, 2, \dots$  identify particular strategies of the players in accordance with  
the subscripts attached to  $7^*$ 's. This normal form admits two Nash equilibria,  
as indicated, which correspond, in the original game, to the strategy triplets  $\{V$   
 $\dots 2^*$  and  $b1^* = Rn2$  There exists only one game which is informationally  
inferior to the single-act game under consideration, which is the one obtained  
by allowing a single information set to  $P_3$  (i.e., the static version). Since the  
second Nash solution dictates a constant action for  $P_3$ , it clearly also  
constitutes a Nash solution for this informationally inferior static game (cf.  
Prop. 3.7). Hence, for the second set of Nash strategies,  $j^* = 1$ , using the  
terminology of Def. 3.20. The first set of Nash strategies, on the other hand,  
does not constitute a Nash equilibrium solution to the informationally inferior  
game, and hence in this case  $j^* = 0$ . This then yields the conclusion that the  
nested single-act game of Fig. 3.6 admits a unique delayed commitment type  
Nash equilibria, given by (i) and with an outcome of  $(0, 0, 0)$ .  $\square$

114 T. BAŞAR AND G. J. OLSDER For more general types of single-act finite  
games in nested extensive form, a recursive procedure could be used to  
simplify the derivation by decomposing the original game into static and  
dynamic undecomposable sub-extensive forms. Essential steps of such a  
recursive procedure are given below. Nash equilibria obtained through this  
recursive procedure will be called the delayed commitment type. A recursive  
procedure to determine delayed commitment type pure- strategy Nash  
equilibria of  $V$ -person single-act games in nested extensive form (1) Each  
information set of the last-acting player is included in either a static sub-  
extensive form or a dynamic undecomposable sub-extensive form. Single out  
these extensive forms and obtain their delayed commitment type Nash  
equilibria. (Note that in static games every Nash equilibrium is, by definition,  
of the delayed commitment type.) Assuming that each of these games admits a  
unique pure-strategy Nash equilibrium solution of the delayed commitment  
type, record the equilibrium strategies and the corresponding Nash outcomes  
by identifying them with the players and their information sets. (If any one of  
these games admits more than one Nash equilibrium solution of the delayed  
commitment type, then this implies that the original game admits more than one  
such equilibrium, and this procedure as well as the following steps will have



to be repeated for each of these multiple equilibria.) (2) Replace each sub-extensive form considered at step 1 with the immediate branch emanating from its vertex that corresponds to the delayed commitment Nash strategy of the starting player of that sub-extensive form; furthermore, attach the corresponding  $n$ -tuple of Nash equilibrium values to the end of this branch. (3) For the remaining game in nested extensive form, repeat steps 1 and 2 until the extensive form left is a tree structure comprised of only a vertex and some immediate branches with an  $n$ -tuple of numbers attached to each of them. (4) The branch of this final tree structure which corresponds to a minimum loss for the starting player is his Nash equilibrium strategy of the delayed commitment type, and the corresponding  $n$ -tuple of numbers at the end of this branch determines a Nash outcome for the original game. The corresponding delayed commitment type Nash equilibrium strategies of the other players in the original single-act game can then be captured from the solution(s) of the sub-extensive forms considered, by appropriately identifying them with the players and their information sets. Example 3.10 To illustrate the steps of the foregoing recursive procedure, let us consider the 3-person single-act game of Fig. 3.7, which is in nested extensive

N-PERSON NONZERO-SUM FINITE GAMES 115 Figure 3.7: The single-act game of Example 3.10. form. At step 1 of the recursive procedure, we have two 2-person static sub-extensive forms and one 3-person dynamic undecomposable sub-extensive form. Counting from the left, the first static sub-extensive form admits the unique Nash equilibrium solution (from Example 3.7):  $\{u_2^* = L, u_3^* = L\}$ . (i) The second static sub-extensive form also admits a unique solution which is (again from Example 3.7):  $\{u_2^* = R, u_3^* = L\}$ . ( $\ll$ ) It should be noted that (i) corresponds to the information sets  $r_{j2} = r_f = \{u_1 = L\}$  and (ii) corresponds to  $r_{j2} = r_f = \{u_1 = M_1\}$ . Now, the dynamic undecomposable sub-extensive form is in fact the single-act game of Example 3.9, which is already known to admit the unique delayed-commitment type Nash equilibrium solution  $V^* = M_2$ ,  $v_2^* = L$ ,  $Y_3^*(n_3) = 1L$  if  $u_1 = M_2$ ,  $71, 7$   $\setminus j_i \mid R$  if  $u_1 = R$ . At step 2, therefore, the extensive form of Fig. 3.7 becomes  $1, -1, 0, 0, -2, -1, 0, 0, 0$  and it readily follows from this tree structure that PI has two Nash equilibrium strategies:  $71^* = M_1$  and  $71^* = M_2$ .

116 T. BAŞAR AND G. J. OLSDER Consequently, the original single-act game admits two Nash equilibria of the delayed commitment type:  $1^*$ ,  $2^*/ 2n f R$  if

$w_1 = M_1$ ,  $V^* = m_i$ ,  $72-(,2) = | L$  otherwise;  $73V) = \{ * R$  if  $u_1 = i?$ , otherwise  
 and  $1^* ,,,, 2^*/ 2n f \& \text{if } 11^* = M_1, !V) = \{ L$  otherwise  $3^*( 2) = r -r$  if  $\langle x=i?$ , ' [ $\wedge$   
 $L$  otherwise , with equilibrium outcomes being  $(0, -2, -1)$  and  $(0,0,0)$ ,  
 respectively. D Remark 3.11 As we have discussed earlier in this section,  
 single-act nonzero- sum finite dynamic games admit, in general, a multiple of  
 Nash equilibria, mainly due to the fact that Nash equilibria of informationally  
 inferior games also provide Nash solutions to the original game (cf. Prop. 3.7)  
 the so-called informational nonuniqueness of Nash equilibria. The delayed  
 commitment mode of play introduced for nested and ladder-nested games,  
 however, eliminates this informational nonuniqueness, and therefore  
 strengthens the Nash equilibrium solution concept for dynamic single-act  
 games. D 3.5.2 Single-act games: Nash equilibria in behavioral and mixed  
 strategies If an TV-person single-act game in extensive form does not admit a  
 Nash equilibrium solution (in pure strategies), then an appropriate approach is  
 to investigate Nash equilibria within the enlarged class of behavioral  
 strategies, as has been done in Section 2.5 for zero-sum games. If the nonzero-  
 sum single-act finite game under consideration is of the "ladder-nested" type,  
 then it is relatively simpler to obtain Nash equilibria in behavioral strategies  
 (which also fit well within a delayed commitment framework), since one still  
 follows the recursive procedure developed in this section for pure strategies,  
 but this time by also considering the mixed-strategy Nash equilibria of the  
 static sub-extensive forms. Then, these mixed strategies, obtained for each  
 equivalent normal form, will have to be appropriately concatenated and  
 written in the form of a behavioral strategy. The following example now  
 illustrates the steps in the derivation of behavioral Nash equilibria. Example  
 3.11 Consider the 3-person single-act game depicted in Fig. 3.8. This is of the  
 ladder-nested type, and both P2 and P3 have access to the

N-PERSON NONZERO-SUM FINITE GAMES 117 Figure 3.8: The 3-person  
 single-act game of Example 3.11. action of P1, but to no other information.  
 Hence, we first have to consider two bimatrix games corresponding to the  
 possible actions L and R of P1. If  $u_1 \sim L$ , then the bimatrix game is the one  
 treated in Example 3.2, which admits the unique mixed-strategy Nash  
 equilibrium  $2, J L$  w.p.  $1/2$ ,  $s R$  w.p.  $1/2$ , ,  $_ J L$  w.p.  $1/2$ ,  $R$  w.p.  $1/2$ , with the  
 corresponding average cost triple being  $(0, 1/2, 3/2)$ . If  $u_1 = R$ , on the other  
 hand, the bimatrix game is  $P_3 A = R \begin{matrix} 3 & 2 \end{matrix} P_3 P_2, B \begin{matrix} L & R \end{matrix} L R \begin{matrix} 0 & 2 \end{matrix} P_2, L R$  which  
 admits the unique Nash solution (in pure strategies)  $,.3^*$  with the corresponding

cost triple being (1,1,1). Comparing this outcome with the previous one, it is obvious that the unique equilibrium strategy of P1 is  $s_1^* = L$ . Appropriate concatenation now yields the following unique behavioral Nash equilibrium strategies for P2 and P3, respectively:  $(s_2^*, s_3^*) = (LR, LRLRLR)$  w.p. w.p. w.p. w.p. w.p. w.p.  $1/2 \ 1/2 \ 1 \ 0 \ 1/2 \ 1/2 \ 1 \ 0$

118 T. BAŞAR AND G. J. OLSDER The unique average Nash equilibrium outcome in behavioral strategies is  $(0, 1/2, 3/2)$ . D Since the recursive procedure involves solution of static games in normal form at every step of the derivation, and since every N-person nonzero-sum finite game in normal form admits a mixed strategy Nash equilibrium solution (cf. Thm. 3.2), it follows that a Nash equilibrium solution in behavioral strategies always exists for single-act games in ladder-nested extensive form. This result is summarized below in Prop. 3.8, which is the counterpart of Corollary 2.4 in the present framework. Proposition 3.8 In every single-act N-person game which has a ladder-nested extensive form comprising a finite number of branches, there exists at least one Nash equilibrium solution in behavioral strategies. Remark 3.12 It should be noted that, as opposed to Corollary 2.4, Prop. 3.8 does not claim an ordered interchangeability property in case of multiple Nash equilibria, for reasons which have extensively been discussed in Sections 3.2 and 3.3. Furthermore, the result might fail to hold if the single-act game is not ladder-nested, in which case one has to consider the even larger class of mixed strategies. □ Remark 3.13 Since the players have only a finite number of possible pure strategies, it readily follows from Thm. 3.2 that there always exists a Nash equilibrium in mixed strategies in single-act finite games of the general type. However, computation of these mixed equilibrium strategies is quite a nontrivial problem, since a recursive procedure cannot be developed for finite single-act games which are not ladder-nested. □ 3.5.3 Multi-act games: Pure-strategy Nash equilibria In the discussion of the Nash-equilibria of multi-act nonzero-sum finite games in extensive form, we shall follow rather closely the analysis of Section 2.5 which was devoted to similarly structured zero-sum games, since the results to be obtained for multi-act nonzero-sum games will be direct extensions of their counterparts in the case of zero-sum games. Not all conclusions drawn in Section 2.5, however, are valid in the present context and these important differences between the properties of saddle-point and Nash equilibrium solutions will also be elucidated in the sequel. In order to be able to develop a systematic (recursive) procedure to

obtain Nash equilibria of multi-act nonzero-sum finite games, we will have to impose (as in Section 2.5) some restrictions on the nature of the informational coupling among different levels of play. To this end, we confine ourselves to multi-act games whose Nash equilibria can be obtained by solving a sequence of single-act games and by appropriate concatenation of the equilibrium strategies determined at each level of play. This reasoning, then, brings us to the so-called

AT-PERSON NONZERO-SUM FINITE GAMES 119 "feedback games", a precise definition of which is given below (as a counterpart of Def. 2.11).

Definition 3.21 A multi-act  $N$ -person nonzero-sum game in extensive form with a fixed order of play is called an  $N$ -person nonzero-sum feedback game in extensive form, if (i) at the time of his act, each player has perfect information concerning the current level of play, i.e., no information set contains nodes of the tree belonging to different levels of play, (ii) information sets of the first-acting player at every level of play are singletons, and the information sets of the other players at every level of play are such that none of them includes nodes corresponding to branches emanating from two or more different information sets of the first-acting player, i.e., each player knows the state of the game at every level of play. If, furthermore, (iii) the single-act games corresponding to the information sets of the first-acting player at each level of play are of the ladder-nested (respectively, nested) type (cf Def. 3.15), then the multi-act game is called an  $N$ -person nonzero-sum feedback game in ladder-nested (respectively, nested,) extensive form. Remark 3.14 It should be noted that, in the case of two-person multi-act games, every feedback game is, by definition, also in ladder-nested extensive form; hence, the zero-sum feedback game introduced earlier by Def. 2.11 is also ladder-nested.  $\square$  Now, paralleling the analysis of Section 2.5, let the number of levels of play in an  $N$ -person nonzero-sum feedback game in extensive form be  $K$ , and consider a typical strategy  $\sigma^*$  of  $P_i$  in such a game to be composed of  $K$  components  $(\sigma_i^j)_{j=1}^K$ . Here,  $\sigma_i^j$  stands for the corresponding strategy of  $P_i$  at his  $j$ th level of act, and it can be taken to have as its domain the class of only those information sets of  $P_i$  which pertain to the  $j$ th level of play. Let us denote the collection of all such strategies for  $P_i$  at level  $j$  by  $T_i^j$ . Then, for such a game, the set of inequalities (3.23) can be written as (as a counterpart of (2.35))  $(\sigma_i^j \in T_i^j, \sigma_i^j \in T_i^j)_{j=1}^K$  (3.27) which are to be satisfied for all  $\sigma_i^j \in T_i^j, i = 1, \dots, N; j = 1, \dots, K$ . On any  $N$ -tuple of Nash

equilibrium strategies  $\{s_1^*, \dots, s_N^*\}$  that satisfies (3.27), let

120 T. BAŞAR AND G. J. OLSDER us impose the further restriction that it also satisfies (recursively) the following  $K$  TV-tuple inequalities for all  $i \in T_j, i = 1, \dots, N; j = 1, \dots, K: T_i f_i^A \sim, 1, \dots, 2, \dots, 2^* \dots, N, 1k-v1k, 1 > J, 17, 7i > \dots, 7i f_i - i i / c \dots > 7i > \dots > 7/f_i > 7/f; \text{level level if-i } T_2 C_{-1}, 1 \sim 1^* \wedge 2, 2^* \dots, AT, J, H_i, \dots > 7K-i > 7i f_i > \dots > 7/f-u/7/f_i \dots > 7i \dots > 1k-v1k) \wedge t_2 | \dots, 1, \dots, 1^* \dots, 2, \dots AT, \dots, N, \dots AT^* \langle J, 17i, \dots, 7/f_i, 7/f \rangle 7, \dots, 7i \dots \cdot i / f_i, 7/f, J, t_N u, 1 \dots, 1^* \sim 2 \sim 2 \dots 2^* \dots AT \dots \wedge 17i > \dots, 7/f-u/7/f, 7i, \dots, 7/f-i) 7/f, \dots, 7i > \dots > \langle J_N \{i, \dots, 1k-v1k, 1k_i I, \dots, 12k-i, 12k; \dots: 1N\} \wedge 1 \dots / M, \dots, 1, \dots, 1^* \dots, 2, \dots, 2^* \dots, 2^* \dots, 7/f-2, 7/f-i \rangle 7/f \rangle 7i, \dots, 7/f-i, 7/f_i, 7i > \dots > 7/f-2 > 7/f-n, 7K' \dots < J \dots \dots, Ik-2, 7^*-i, Ik, 57i, \dots, 7^* \dots, 7^*-i \rangle 7^* 5, 7i > \dots \cdot) 7/f \dots 2i / f-i, 7K, J, T_2 C_{-1} \dots, 1^* \wedge 1^* \dots, 2, 2, 2^* 2^* \dots J, I, 7i > \dots > 7/f-2' / f-i > 7/f) 7i > \dots > 7/f-2 > 7/f-i) 7/f > \dots > \dots_i V AT, AT^* \dots Af^* \dots 7i > \dots \dots i / f-2i / f-n / f, J < J_2(7i_1, \dots, 7x-2, 7k-i, 17x; 7?, \dots, 7x-2, 7a: -i, 7^*; \dots Af, Af, \dots, AT^* \dots, A / \dots 7i \dots \dots i / f-2' / v-i > 7/f; tWc^i \dots, 1 \dots i^* \dots i^* \dots, 2, 2 \wedge 2^* 2^* \dots J, I, 7i > \dots -i / f-2 > 7/f-i) 7/f_i 7i > \dots \cdot i / f-2 > 7/f-i \rangle 7/f? \dots \dots 1 \wedge Af \wedge Af \wedge Af^* 0 / Ar^* 1, 7i) \dots, 7K \dots, 2, 7/f-i, 7/f, J < \wedge (7j, \dots, 7^*-2, 7i? \dots, 7^*; 7i_2, \dots, \dots, 7^*-2, 7^*-1, 7^*; \dots, N, N \dots N, N^* \dots 7i, \dots \dots, 7/v-2) 7/f-i, 7/f, J, \text{level } 1 \text{ (3.28) } f, J_1(71^*; 72^*; \dots; 7w^*) < -/1(7i_1, 7^*, \dots, 7i_s?; 72^*; \dots; 7w^*) J_2(71^*; 72^*; \dots; 7^*) < \wedge 2(71^*; 7i_2, 72^*, \dots, 7^* 72^*; \dots; 7w^*) \{ J_w(71^*; 72^*; \dots; 7w^*) < \wedge (71^*; 72^*; \dots; 7i_w, 72^* V^*, \dots, 7^*) \}$  Definition 3.22 For an  $N$ -person nonzero-sum feedback game in extensive form with  $K$  levels of play, let  $\{s_1^*, \dots, s_N^*\}$  be an  $N$ -tuple of strategies satisfying (3.27) and (3.28) for all  $i \in G_j, i = 1, \dots, N; j = 1, \dots, K$ . Then,  $\{s_1^*, \dots, s_N^*\}$  constitutes a (pure) feedback Nash equilibrium for the feedback game.

**N-PERSON NONZERO-SUM FINITE GAMES** 121 Proposition 3.9 Every  $N$ -tuple  $\{s_1^*, \dots, s_N^*\}$  that satisfies the set of inequalities (3.28) also satisfies the set of  $N$  inequalities (3.27). Hence, the requirement for satisfaction of (3.27) in Def. 3.22 is redundant, and it can be dispensed with without any loss of generality. Proof. This basically follows from the lines of argument used in the proof of Prop. 2.3, with two inequalities at each level of play replaced by  $N$  inequalities in the present case. The set of inequalities (3.28) now readily suggests a recursive procedure for derivation of feedback Nash equilibria in feedback nonzero-sum games. Starting at level  $K$ , we have to solve a single-act game for each information set of the first-acting player at each level of play, and then appropriately concatenate all the equilibrium strategies (with

restricted domains) thus obtained. If the single-act games encountered in this recursive procedure admit more than one (pure-strategy) Nash equilibrium, the remaining analysis will have to be repeated for each one of these equilibria so as to determine the complete set of feedback Nash equilibria the multi-act game admits. For nested multi-act feedback games one might wish to restrict the class of Nash equilibria of interest even further, so that only those obtained by concatenation of delayed commitment type equilibria of single-act games are considered. The motivations behind such a further restriction are: (i) the delayed commitment mode of play conforms well with the notion of feedback Nash equilibria and it eliminates informational nonuniqueness, and (ii) Nash equilibria at every level of play can then be obtained by utilizing the recursive technique developed earlier in this section for single-act games (which is computationally attractive, especially for ladder-nested games). A precise definition now follows. Definition 3.23 For a nonzero-sum feedback game in extensive form, a feedback Nash equilibrium solution is of the delayed commitment type if its appropriate restriction at each level of play constitutes a delayed commitment type Nash equilibrium (cf. Defs. 3.18, 3.20) for each corresponding single-act game to be encountered in the recursive derivation. Heretofore, the discussion on multi-act feedback games has been confined only to feedback Nash equilibria (of the delayed commitment type or otherwise). However, such games could very well admit other types of Nash equilibria which satisfy inequalities (3.23) but which do not fit the framework of Def. 3.22. The main reason for this phenomenon is the "informational nonuniqueness" feature of Nash equilibria in dynamic games, which has been discussed earlier in this section within the context of single-act games. In the present context, a natural extension of Prop. 3.7 is valid, which we provide below after introducing some terminology and also the concept of informational inferiority in nonzero-sum feedback games. Definition 3.24 Let  $I$  be an  $N$ -person multi-act nonzero-sum game with a fixed order of play, which satisfies the first requirement of Def. 3.21. An  $N$ -person

122 T. BAŞAR AND G. J. OLSDER open-loop game in extensive form (say, 11) (cf Def 3.13) is said to be the static (open-loop) version of  $I$  if it can be obtained from  $I$  by replacing, at each level of play, the information sets of each player with a single information set. The equivalent normal form of the open-loop version of  $I$  is the static normal form of  $I$ . Definition 3.25 Let  $I$  and  $II$  be two  $N$ -person multi-act nonzero-sum games with fixed orders of play, which

satisfy the first requirement of Def. 3.21. Further let  $Y \setminus \text{onrfrji}$  denote the strategy sets of  $P_i$  in  $I$  and  $II$ , respectively. Then  $I$  is informationally inferior to  $II$  if  $T \setminus C Y \setminus i$  for all  $i \in N$ , with strict inclusion for at least one  $i$ . Proposition 3.10 Let  $I$  and  $II$  be two  $N$ -person multi-act nonzero-sum games as introduced in Def. 3.25, so that  $I$  is informationally inferior to  $II$ . Then, (i) any Nash equilibrium solution for  $I$  is also a Nash equilibrium solution for  $II$ , (ii) if  $\{f_1, \dots, f_N\}$  is a Nash equilibrium solution for  $II$  so that  $f_i \in T \setminus$  for all  $i = 1, \dots, N$ , then it is also a Nash equilibrium solution for  $I$ . Proof. The proof is analogous to the proof of Prop. 3.7, and is therefore omitted. D Remark 3.15 It should now be apparent why informationally nonunique Nash equilibria exist in nonzero-sum multi-act feedback games. Every nontrivial multi-act feedback game in which the players have at least two alternatives at each level of play admits several informationally inferior multi-act games. These different games in general admit different Nash equilibria which are, moreover, not interchangeable. Since every one of these also constitutes a Nash equilibrium for the original multi-act game (cf. Prop. 3.10), existence of a plethora of informationally nonunique equilibria readily follows. The further restrictions of Defs. 3.22 and 3.23, imposed on the concept of noncooperative equilibrium, clearly eliminate this informational nonuniqueness, and they stand out as providing one possible criterion according to which a further selection can be made. This "informational nonuniqueness" property of noncooperative equilibrium in dynamic games is, of course, also featured by zero-sum games, since they are special types of nonzero-sum games, in which case the Nash equilibrium solution coincides with the saddle-point solution. However, since all saddle-point strategies have the ordered interchangeability property and further since the saddle-point value is unique, nonuniqueness of equilibrium strategies does not create a problem in zero-sum dynamic games—the main reason why we have not included a discussion on informational nonuniqueness in Chapter 2. It should be noted, however, that for two-person zero-sum multi-act feedback games, Def. 3.23 is redundant and Def. 3.22 becomes equivalent to Def. 2.12, which clearly dictates a delayed commitment mode of play.  $\square$

**IV-PERSON NONZERO-SUM FINITE GAMES 123** The following example now serves to illustrate both the recursive procedure and the informational nonuniqueness of Nash equilibria in feedback nonzero-sum games. 0,1 1,0 2,-1 -1,2 3,1 4,2 1,1 5,2 0,-1/2 2,1 2,1 2,0 0,-1 -2,1 3,2 0,3 2,1 -1,0 Figure 3.9: The multi-act feedback nonzero-sum game of Example 3.12. Example 3.12

Consider the multi-act feedback game whose extensive form is depicted in Fig. 3.9. Since this is a 2-person game (of the feedback type) it is definitely also ladder-nested (or equivalently, nested), and hence derivation of its delayed commitment type feedback Nash equilibria involves solutions of only static sub-extensive forms. At the second (and the last) levels of play, there are four single-act games, one for each singleton information set of PI. The first of these (counting from left) admits the unique Nash equilibrium solution of the delayed commitment type  $2^*$ ,  $i_2^* = L$ ,  $7f(\wedge) R$   $i_1^* = L, L$  otherwise, (0 with the Nash equilibrium outcome being (1,0). The second one admits the unique Nash equilibrium solution  $2^*$ ,  $2 \setminus 72 (\wedge) = \wedge. 72^*W) = i, (\ll)$  with an outcome of (1,1). The third single-act game again has a unique Nash equilibrium, given by  $7r(\wedge) = i, i^* = L (Hi)$  with an outcome of (0,  $-1/2$ ). The last one is the single-act game of Example 3.6, which admits the unique delayed commitment type Nash equilibrium solution  $-rfW-R, 7|*(?) L$  if  $u = L, R$  otherwise (iv)

124 T. BAŞAR AND G. J. OLSDER with the Nash outcome being  $(-1,0)$ . When (i)-(iv) are collected together, we obtain the delayed commitment type Nash equilibrium solution of the feedback game at level 2 to be uniquely given as  $7i^*(\%) = 2 = L, u = L, u = L, 1 -- R, u = R, u = M$  or  $R$ , otherwise. (3.29a) Now, crossing out the second level of play, we end up with the single-act game  $pi/ P2Q II 1,0 X \setminus ru' 1,1 0, -1/2 y r -1,0$  whose equivalent normal form is the bimatrix game  $P2 P2 PI A = R \otimes 1 -1 LR 0 Q i 0 PI R LR$  which admits, as indicated, the unique Nash equilibrium solution with the corresponding outcome being (0,  $-1/2$ ). (3.29b) (3.29c) In conclusion, the feedback game of Fig. 3.9 admits a unique feedback Nash equilibrium solution of the delayed commitment type, which is given by (3.29a)- (3.29b), with the corresponding Nash outcome being (3.29c). This recursively obtained Nash equilibrium solution, although it is unique as a delayed commitment type feedback Nash equilibrium, is clearly not unique as a feedback Nash equilibrium solution, since it is already known from Example 3.6 that the fourth single-act game encountered at the first step of the recursive derivation admits another Nash equilibrium solution which is not of the delayed commitment type. This equilibrium solution is the constant strategy pair given by (3.26a)-(3.26b), that is,  $11^*\{r, \setminus) = L, it^*\{r\$) = L$  which has to replace (iv) in the recursive derivation. The Nash outcome corresponding to this pair of strategies is (0,  $-1$ ). The strategy pair (3.29a) will now



iV-PERSON NONZERO-SUM FINITE GAMES 125 be replaced by 722\*  
 (%2) = R if u\=L,u? = iJ, L otherwise , R if u\ = L,u\ = L,u\ = L, L otherwise ,  
 (3.30a) which also provides a Nash equilibrium solution at level 2. Now,  
 crossing out the second level of play, this time we have the single-act game 1,0  
 1,1 0,-1/2 0,-1 whose equivalent normal form is P2 t, n A = P2 L R 1 0 1 w PI,  
 B = L R L R 0 -1/2 1 kr1) PI, R which admits, as indicated, the unique Nash  
 equilibrium solution with the corresponding outcome being (0,-1). (3.30b)  
 (3.30c) We therefore arrive at the conclusion that (3.30a)-(3.30b) provides  
 another feedback Nash equilibrium solution for the multi-act game of Fig. 3.9,  
 which is, however, not of the delayed commitment type. It is noteworthy that  
 the unique feedback Nash solution of the delayed commitment type is in this  
 case inadmissible, which follows immediately by a comparison of (3.29c) and  
 (3.30c). To determine the other (non-feedback type) Nash equilibria of this  
 multi-act game, the only possible approach is to bring it into equivalent normal  
 form and to investigate Nash equilibria of the resulting bimatrix game. This is,  
 however, quite a strenuous exercise, since the bimatrix game is of rather high  
 dimensions. Specifically, at level 1 each player has two possible strategies,  
 while at level 2 PI has  $2^3 \times 3 = 24$  and P2 has  $2^6 = 64$  possible strategies.  
 This then implies that the bimatrix game is of dimensions  $48 \times 128$ . Two Nash  
 equilibria of this bimatrix game are the ones given by (3.29a)-(3.29b) and  
 (3.30a)-(3.30b), but there are possibly others which in fact correspond to Nash  
 equilibria of multi- act games which are informationally inferior to the one of  
 Fig. 3.9; one of these games is the open-loop game which is the static version  
 of the original feedback game. □

126 T. BAŞAR AND G. J. OLSDER 3.5.4 Multi-act games: Behavioral and  
 mixed equilibrium strategies Behavioral equilibrium strategies for an TV-  
 person feedback game in ladder- nested extensive form can be derived by  
 essentially following the recursive procedure developed in Section 3.5.3 for  
 pure-strategy feedback equilibria, but this time by also considering the  
 behavioral-strategy Nash equilibria of the single-act games to be encountered  
 in the derivation. The justification of applicability of such a recursive  
 procedure is analogous to the one of behavioral saddle-point equilibria of  
 feedback games discussed in Section 2.5. Specifically, if K denotes the number  
 of levels of play in the ./V-person feedback game of the ladder-nested type,  
 then a typical behavioral strategy  $\gamma^*$  of  $P_i$  in such a game can be decomposed  
 into K components  $\gamma^1, \dots, \gamma^K$ , where  $\gamma^j \in r_j$  is the corresponding behavioral

strategy of  $P_i$  at his  $j$ th level of play. Furthermore,  $\Gamma_j$  has, as its domain, only the class of those information sets of  $P_i$  which pertain to the  $j$ th level of play. Then, the set of inequalities that determines the behavioral Nash equilibrium solution can be written as (3.31)  $\{j \in T^*, j = 1, \dots, K; i \in N\}$ , which have to be satisfied for all  $j \in T^*, j = 1, \dots, K; i \in N$ . Under the delayed commitment mode of play, we impose the further restriction that they satisfy the  $K$  TV-tuple inequalities similar to (3.28), with only  $\Gamma_j$  replaced by  $\Gamma_j$  and  $J_1$  by  $J \setminus A$ . A counterpart of Prop. 3.9 also holds true in the present framework and this readily leads to recursive derivation of behavioral-strategy Nash equilibria in feedback games of the ladder-nested type. This procedure always leads to a Nash equilibrium solution in behavioral strategies as the following proposition states. Proposition 3.11 Every finite  $N$ -person nonzero-sum feedback game in ladder-nested extensive form admits a Nash equilibrium solution in behavioral strategies which can be determined recursively. Proof. This result is an immediate consequence of Prop. 3.8, since every feedback game in ladder-nested extensive form can be decomposed recursively into ladder-nested single-act games. D For feedback games which are not of the ladder-nested type, and also for other types of (non-feedback) multi-act games, Prop. 3.11 is not in general valid, and the Nash equilibrium in behavioral strategies (even if it exists) cannot be obtained recursively. Then, the only approach would be to bring the extensive form into an equivalent normal form, obtain all mixed-strategy Nash equilibria of this normal form, and to seek whether any one of these equilibria would constitute a Nash equilibrium in behavioral strategies. This argument is now

**N-PERSON NONZERO-SUM FINITE GAMES 127** supported by the following two propositions which are direct counterparts of Props 2.5 and 2.6, respectively. Proposition 3.12 Every  $N$ -person nonzero-sum multi-act finite game in extensive form admits a Nash equilibrium solution in mixed strategies. Proof. The result readily follows from Thm. 3.2, since every such game can be transformed into an equivalent normal form in which each player has a finite number of strategies. D Proposition 3.13 Every Nash equilibrium of a finite  $N$ -person nonzero-sum multi-act game in behavioral strategies also constitutes a Nash equilibrium in the larger class of mixed strategies. Proof. Let  $\{f_1^*, \dots, f_N^* \in T^N\}$  denote an TV-tuple of behavioral strategies stipulated to be in Nash equilibrium, and let  $f_i$  denote the mixed-strategy set of  $P_i, i \in N$ . Since  $P \in C P$ , we clearly have  $f_i^* \in f_i$  for every  $i \in N$ . Assume, to the contrary, that

$\{y_1^*, \dots, y_n^*\}$  is not a mixed-strategy Nash equilibrium solution; then this implies that there exists at least one  $i$  (say,  $i = n$ , without any loss of generality) for which the corresponding inequality of (3.31) is not satisfied for all  $y_i \in P$ . In particular, there exists a  $y_i^* \in P$  such that  $J_N(y_1^*, \dots, y_n^*) > F^*$ . (t) Now, abiding by our standard convention, let  $T_N$  denote the pure-strategy set of  $PAT$ , and consider the quantity  $F(y_1^*, \dots, y_n^*)$  defined for each  $y_i \in T_N$ . (This is well defined since  $T_M \subset T_N$ .) The infimum of this quantity over  $T_N$  is definitely achieved (say, by  $y_i^* \in T_N$ ), since  $T_N$  is a finite set. Furthermore, since  $P_N$  is comprised of all probability distributions on  $T_N$ ,  $\inf_{F \in P_N} F(y_i^*) = F(y_i^*)$ . We therefore have  $F^* = F(y_i^*)$ , (H) and also the inequality  $J_N^* > J_N(y_i^*)$  in view of (i). But this is impossible since  $J_N^* = \inf_{y \in T_N} J_N(y)$  and  $T_N \subset T_N$ , thus completing the proof of the proposition. D  
 19 Here, by an abuse of notation, we take  $J_1$  to denote the average loss to  $P_i$  under also mixed-strategy  $JV$ -tuples.

128 T. BAŞAR AND G. J. OLSDER 3.5.5 Other refinements on Nash equilibria  
 As we have seen heretofore in this chapter, existence of multiple Nash equilibria for a given dynamic game is more a rule rather than an exception, with the multiplicity arising because of the informational richness of the underlying decision problem as well as the structure of the players' cost matrices. As a means of shrinking the set of Nash equilibria in a rational way, we have introduced heretofore the notions of "delayed commitment", "informational inferiority", "feedback games" and "admissibility". There are, however, several other refinement schemes which have been introduced in the literature, some of which we discuss below. To motivate the discussion, let us start with a two-player matrix game  $(A, B)$  where the players have identical cost matrices.  $P_2 A = B \begin{matrix} L & R \\ 0 & 1 \\ 1 & 1 \end{matrix} P_1$ . (3.32)  $L \ R$  The game admits two pure-strategy Nash equilibria:  $(L, L)$  and  $(R, R)$ . Note, however, that if we perturb the entries of the two matrices slightly, and independently:  $P_2 P_2 A + \epsilon \begin{matrix} 1 & 1 \\ 1 & 2 \end{matrix} P_1, B + \epsilon \begin{matrix} 1 & 1 \\ 1 & 2 \end{matrix} P_1$ , where  $\epsilon^i, j, k = 1, 2$ , are infinitesimally small (positive or negative) numbers, then  $(L, L)$  will still retain its equilibrium property (as long as  $|\epsilon^*| < 1/2$ ), but  $(R, R)$  will not. More precisely, there will exist infinitely many perturbed versions of the original game for which  $(R, R)$  will not constitute a Nash equilibrium. Hence, in addition to admissibility,  $(L, L)$  can be singled out in this case as the Nash solution that is robust to infinitesimal perturbations in the entries of the cost matrices. Can such perturbations be induced naturally by some behavioral

assumptions imposed on the players? The answer is yes, as we discuss next. Consider the scenario where a player who intends to play a particular pure strategy (out of a set of  $n$  possible alternatives) errs and plays with some small probability one of the other  $n - 1$  alternatives. In the matrix game (3.32), for example, if both players err with equal (independent) probability  $e > 0$ , the resulting matrix game is  $(A_t, B_t)$ , where  $P_2 A_e = B_t = L R e \{2 - t\} 1 - e + e^* 1 - e + e' 1 - e^* P_1$ . Note that for all  $e \in (0, 1/2)$  this matrix game admits the unique Nash equilibrium  $(L, L)$ , with a cost pair of  $(e(2 - e), e(2 - e))$ , which converges to  $(0, 0)$

AT-PERSON NONZERO-SUM FINITE GAMES 129 as  $e \rightarrow 0$ , thus recovering one of the Nash cost pairs of the original game. A Nash equilibrium solution that can be recovered this way is known as a perfect equilibrium, which was first introduced in precise terms by Selten (1975), in the context of  $n$ -player games in extensive form. Given a game of perfect recall, denoted  $Q$ , the idea is to generate a sequence of games,  $Q, G_2, \dots, G_k, \dots$ , a limiting equilibrium solution of which (in behavioral strategies, and as  $k \rightarrow \infty$ ) is an equilibrium solution of  $Q$ . If  $Q_k$  is obtained from  $Q$  by forcing the players at each information set to choose every possible alternative with positive probability (albeit small, for those alternatives that are not optimal), then the equilibrium solution(s) of  $Q$  that are recovered as a result of the limiting procedure above is (are) called perfect equilibrium (equilibria).<sup>20</sup> Selten (1975) has shown that every finite game in extensive form with perfect recall (and as a special case in normal form) admits at least one perfect equilibrium, thus making this refinement scheme a legitimate one. The procedure discussed above, which amounts to "completely" perturbing a game with multiple equilibria, is one way of obtaining perfect equilibria; yet another one, as introduced by Myerson (1978), is to restrict the players to use completely mixed strategies (with some lower positive bound on the probabilities) at each information set. Again referring back to the matrix game  $(A, B)$  of (3.32), let the players' mixed strategies be restricted to the class  $\{y, z\}$ , where  $e < y < 1 - e, e < z < 1 - e$ , for some (sufficiently small) positive  $e$ . Over this class of strategies, the average cost functions of the players will be  $J_1 = J_2 = -2/2 + 1$ , which admits (assuming that  $0 < e < 1/2$ ) a unique Nash equilibrium:  $(L, L)$ . Such a solution is called an  $e$ -perfect equilibrium (Myerson, 1978), which in the limit as  $e \rightarrow 0$  clearly yields the perfect Nash equilibrium obtained earlier.

Myerson in fact proves, for  $N$ -person games in normal form, that every perfect equilibrium can be obtained as the limit of an appropriate  $\epsilon$ -perfect equilibrium, with the converse statement also being true. More precisely, using the notation of Section 3.3, we have the following. Proposition 3.14 For an  $N$ -person finite game in normal form, a mixed-strategy Nash equilibrium  $\{y^i \in Y^i, i \in N\}$  is a perfect equilibrium if, and only if there exist some sequences  $\{\epsilon^k, i \in N\}$  such that  $\epsilon^k \rightarrow 0$ . This is also called "trembling hand equilibrium", as the process of erring at each information set is reminiscent of a "trembling hand" making unintended choices with small probability. Here, as  $k \rightarrow \infty$ , this probability of unintended plays converges to zero.

130 T. BAŞAR AND G. J. OLSDER i)  $\epsilon^k > 0$  and  $\lim_{k \rightarrow \infty} \epsilon^k = 0$  ii)  $\{y^k, i \in N\}$  is an  $\epsilon^k$ -perfect equilibrium in)  $\lim_{k \rightarrow \infty} y^k = y^*$ ,  $i \in N$ . Furthermore, a perfect equilibrium necessarily exists, and every perfect equilibrium is a Nash equilibrium. Even though the notion of perfect equilibrium provides a refinement of the notion of Nash equilibrium with some appealing properties, it also carries some undesirable features as the following example of an identical cost matrix game (due to Myerson (1978)) exhibits:  $P^2 L A = B = M [l] l \in I$  PI. (3.33)  $0 \ 1 \ -1 \ 10 \ 1 \ 1 \ -1 \ 8 \ 10 \ 8 \ 8$  R  $\{^T M R \sim$  Note that this is a matrix game derived from (3.32) by adding a completely dominated row and a completely dominated column. It now has three Nash equilibria:  $(L,L)$ ,  $(M,M)$ ,  $(R,R)$ , the first two of which are perfect equilibria, while the last one is not.<sup>21</sup> Hence, inclusion of completely dominated rows and columns could create additional perfect equilibria not present in the original game—a feature that is clearly not desirable. To remove this shortcoming of perfect equilibrium, Myerson (1978) has introduced the notion of proper equilibrium, which corresponds to a particular construction of the sequence of strategies used in Prop. 3.14. Proper equilibrium is defined as in Prop. 3.14, with only the  $\epsilon^k$ -perfect equilibrium in ii) replaced by the notion of  $\epsilon^k$ -proper equilibrium to be introduced next. Toward this end, let  $J^i(j; y^k)$  denote the average cost to  $P_i$  when he uses his  $j$ th strategy (such as  $j$ th column or row of the matrix) in the game and all the other players use their mixed strategies  $y^k$ ,  $\epsilon^k \in N$ ,  $k \geq i$ . Furthermore, let  $y_j^k$  be the probability attached to his  $j$ th strategy under the mixed strategy  $y^k$ . Then, the  $N$ -tuple  $\{y^k, i \in N\}$  is said to be in  $\epsilon^k$ -proper equilibrium if the strict inequality  $J^i(j; y^k) > J^i(k; y^k)$  implies that  $y_j^k < \epsilon^k y_k^k$ , this being so for every  $j, k \in M$ , and every  $i \in N$ . In other words, an  $\epsilon^k$ -proper equilibrium is one in which every player is giving his better responses much

more probability weight than his worse responses (by a factor  $1/e$ ), regardless of whether those "better" responses are "best" or not. Myerson (1978) proves that such an equilibrium necessarily exists as follows. Proposition 3.15 Every finite  $N$ -player game in normal form admits at least one proper equilibrium. Furthermore, every proper equilibrium is a perfect equilibrium (but not vice versa). The reader is encouraged to verify this conclusion.

iV-PERSON NONZERO-SUM FINITE GAMES 131 Remark 3.16 The reader should verify that in the matrix game (3.33) there is only one proper equilibrium, which is  $\{L,L\}$ : the perfect equilibrium of (3.32).  $\square$  Another undesirable feature of a perfect equilibrium is that it is very much dependent on whether the game is in extensive or normal form (whereas the Nash equilibrium property is form-independent). As it has been first observed by Selten (1975), and further elaborated on by van Damme (1984), a perfect equilibrium of the extensive form of a game need not be perfect in the normal form, and conversely a perfect equilibrium of the normal form need not be perfect in the extensive form. To remove this undesirable feature, van Damme (1984) has introduced the concept of quasi-perfect equilibria for games in extensive form, and has shown that a proper equilibrium of a normal form game induces a quasi-perfect equilibrium in every extensive form game having this normal form. Quasi-perfect equilibrium is defined as a behavioral strategy combination which prescribes at every information set a choice that is optimal against mistakes ("trembling hands") of the other players; its difference from perfect equilibrium is that here in the construction of perturbed matrices each player ascribes "trembling hand" behavior to all other players (with positive probability), but not to himself. Other types of refinement have also been proposed in the literature, such as sequential equilibria (Kreps and Wilson, 1982), and strategic equilibria (Kohlberg and Mertens, 1986), which we do not further discuss here. None of these, however, are uniformly powerful, in the sense of shrinking the set of Nash equilibria to the smallest possible set. We will revisit the topic of "refinement on Nash equilibria" later in Chapters 5 and 6, in the context of infinite dynamic games and with emphasis placed on the issue of time consistency.

### 3.6 The Stackelberg Equilibrium Solution

The Nash equilibrium solution concept that we have heretofore studied in this chapter provides a reasonable noncooperative equilibrium solution for nonzero-sum games when the roles of the players are symmetric, that is to say, when no single player dominates the decision process. However, there are yet

other types of noncooperative decision problems wherein one of the players has the ability to enforce his strategy on the other player(s), and for such decision problems one has to introduce a hierarchical equilibrium solution concept. Following the original work of H. von Stackelberg (1934), the player who holds the powerful position in such a decision problem is called the leader, and the other players who react (rationally) to the leader's decision (strategy) are called the followers. There are, of course, cases of multi-levels of hierarchy in decision making, with many leaders and followers; but for purposes of brevity and clarity in exposition we will first confine our discussion here to hierarchical decision problems which incorporate two players (decision makers)—one leader and one follower. To set the stage to introduce the hierarchical (Stackelberg) equilibrium so-

132 T. BAŞAR AND G. J. OLSDER

tion concept, let us first consider the bimatrix game  $(A, B)$  displayed (under our standard convention) as

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 3/2 & 3 \\ 2 & 2 & 2 \end{pmatrix}, B = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix} \quad (3.34)$$

This bimatrix game clearly admits a unique Nash equilibrium solution in pure strategies, which is  $\{M, M\}$ , with the corresponding outcome being  $(1, 0)$ . Let us now stipulate that the roles of the players are not symmetric and P1 can enforce his strategy on P2. Then, before he announces his strategy, P1 has to take into account possible responses of P2 (the follower), and in view of this, he has to decide on the strategy that is most favorable to him. For the decision problem whose possible cost pairs are given as entries of  $A$  and  $B$ , above, let us now work out the reasoning that P1 (the leader) will have to go through. If P1 chooses  $L$ , then P2 has a unique response (that minimizes his cost) which is  $L$ , thereby yielding a cost of  $0$  to P1. If the leader chooses  $M$ , P2's response is again unique (which is  $M$ ), with the corresponding cost incurred to P1 being  $1$ . Finally, if he picks  $R$ , P2's unique response is also  $R$ , and the cost to P1 is  $2$ . Since the lowest of these costs is the first one, it readily follows that  $L$  is the most reasonable choice for the leader (P1) in this hierarchical decision problem. We then say that  $L$  is the Stackelberg strategy of the leader (P1) in this game, and the pair  $\{L, L\}$  is the Stackelberg solution with P1 as the leader. Furthermore, the cost pair  $(0, -1)$  is the Stackelberg (equilibrium) outcome of the game with P1 as the leader. It should be noted that this outcome is actually more favorable for both players than the unique Nash outcome—this latter feature, however, is not a rule in such games. If, for example, P2 is the leader in the bimatrix game (3.34), and

PI the follower, then the unique Stackelberg solution is  $\{L, R\}$  with the corresponding outcome being  $(3/2, -2/3)$  which is clearly not favorable for PI (the follower) when compared with his unique Nash cost. For P2 (the leader), however, the Stackelberg outcome is again better than his Nash outcome. The Stackelberg equilibrium solution concept introduced above within the context of the bimatrix game (3.34) is applicable to all two-person finite games in normal form, but provided that they exhibit one feature which was inherent to the bimatrix game (3.34) and was used implicitly in the derivation: the follower's response to every strategy of the leader should be unique. If this requirement is not satisfied, then there is ambiguity in the possible responses of the follower and thereby in the possible attainable cost levels of the leader. As an explicit example to demonstrate such a decision situation, consider the bimatrix game  $A = \begin{matrix} L & R \\ P2 & \begin{matrix} 0 & 2 \\ 1 & 2 \end{matrix} \\ 3 & -1 \end{matrix}$ .  $B = \begin{matrix} L & M & R \\ P2 & \begin{matrix} 0 & -1 \\ 0 & 0 \end{matrix} \\ 1 & -1 \end{matrix}$  (3.35)  $M & R$

IV-PERSON NONZERO-SUM FINITE GAMES 133 and with PI acting as the leader. Here, if PI chooses (and announces) L, P2 has two optimal responses L and M, whereas if PI picks R, P2 again has two optimal responses, L and R. Since this multiplicity of optimal responses for the follower results in a multiplicity of cost levels for the leader for each one of his strategies, the Stackelberg solution concept introduced earlier cannot directly be applied here. However, this ambiguity in the attainable cost levels of the leader can be resolved if we stipulate that the leader's attitude is toward securing his possible losses against the choices of the follower within the class of his optimal responses, rather than toward taking risks. Then, under such a mode of play, PI's secured cost level corresponding to his strategy L would be 1, and the one corresponding to R would be 2. Hence, we declare  $J_1^* = 1$  as the unique Stackelberg strategy of PI in the bimatrix game of (3.35), when he acts as the leader. The corresponding Stackelberg cost for PI (the leader) is  $J_1^* = 1$ . It should be noted that, in the actual play of the game, PI could actually end up with a lower cost level, depending on whether the follower chooses his optimal response  $J_2 = L$  or the optimal response  $J_2 = M$ . Consequently, the outcome of the game could be either  $(1,0)$  or  $(0,0)$ , and hence we cannot talk about a unique Stackelberg equilibrium outcome of the bimatrix game (3.35) with PI acting as the leader. Before concluding our discussion on this example, we finally note that the admissible Nash outcome of the bimatrix game (3.35) is  $(-1, -1)$  which is more favorable for both players than the possible



Stackelberg outcomes given above. We now provide a precise definition for the Stackelberg solution concept introduced above within the context of two bimatrix games, so as to encompass all two-person finite games of the single-act and multi-act type which do not incorporate any chance moves. For such a game, let  $T_1$  and  $T_2$  again denote the pure-strategy spaces of  $P_1$  and  $P_2$ , respectively, and  $J_i(x_1, x_2)$  denote the cost incurred to  $P_i$  corresponding to a strategy pair  $\{x_1 \in T_1, x_2 \in T_2\}$ . Then, we have the following definitions.

**Definition 3.26** In a two-person finite game, the set  $R_2^1 \subset T_2$  defined for each  $x_1 \in T_1$  by  $R_2^1(x_1) = \{x_2 \in T_2 : J_2(x_1, x_2) \leq J_2(x_1, y_2), \forall y_2 \in T_2\}$  (3.36) is the optimal response (rational reaction) set of  $P_2$  to the strategy  $x_1 \in T_1$  of  $P_1$ .

**Definition 3.27** In a two-person finite game with  $P_1$  as the leader, a strategy  $x_1^* \in T_1$  is called a Stackelberg equilibrium strategy for the leader, if  $\max_{x_1 \in T_1} J_1(x_1, R_2^1(x_1)) = J_1(x_1^*, R_2^1(x_1^*))$ . (3.37) The quantity  $J_1^*$  is the Stackelberg cost of the leader. If, instead,  $P_2$  is the leader, the same definition applies with only the superscripts 1 and 2 interchanged.

134 T. BAŞAR AND G. J. OLSDER **Theorem 3.3** Every two-person finite game admits a Stackelberg strategy for the leader. **Proof** Since  $T_1$  and  $T_2$  are finite sets, and  $R_2^1(x_1)$  is a subset of  $T_2$  for each  $x_1 \in T_1$ , the result readily follows from (3.37). **Remark 3.17** The Stackelberg strategy for the leader does not necessarily have to be unique. But nonuniqueness of the equilibrium strategy does not create any problem here (as it did in the case of Nash equilibria), since the Stackelberg cost for the leader is unique.  $\square$  **Remark 3.18** If  $R_2^1(x_1)$  is a singleton for each  $x_1 \in T_1$ , then there exists a mapping  $T_2 : T_1 \rightarrow T_2$  such that  $x_2 \in R_2^1(x_1)$  implies  $x_2 = T_2(x_1)$ . This corresponds to the case in which the optimal response of the follower (given by  $T_2$ ) is unique for every strategy of the leader, and it leads to the following simplified version of (3.37) in Def. 3.27:  $J_1(x_1^*, T_2(x_1^*)) = \min_{x_1 \in T_1} J_1(x_1, T_2(x_1)) = J_1^*$ . (3.38) Here  $J_1^*$  is no longer only a secured equilibrium cost level for the leader ( $P_1$ ), but it is the cost level that is actually attained.  $\square$  From the follower's point of view, the equilibrium strategy in a Stackelberg game is any optimal response to the announced Stackelberg strategy of the leader. More precisely, we have the following. **Definition 3.28** Let  $x_1^* \in T_1$  be a Stackelberg strategy for the leader ( $P_1$ ). Then, any element  $x_2^* \in R_2^1(x_1^*)$  is an optimal strategy for the follower ( $P_2$ ) that is in equilibrium with  $x_1^*$ . The pair  $\{x_1^*, x_2^*\}$  is a Stackelberg solution for the game with  $P_1$  as the leader, and the cost pair

$(\hat{J}_1(71^*, 72^*), \hat{J}_2(71^*, 72^*))$  is the corresponding Stackelberg equilibrium outcome. Remark 3.19 In the preceding definition, the cost level  $J_1(71^*, 72^*)$  could in fact be lower than the Stackelberg cost  $J_1^*$ —a feature which has already been observed within the context of the bimatrix game (3.35).

However, if  $R_2^1$  is a singleton, then these two cost levels have to coincide. □ For a given two-person finite game, let  $J_1^*$  again denote the Stackelberg cost of the leader (PI), and  $\hat{J}$  denote any Nash equilibrium cost for the same player. We have already seen within the context of the bimatrix game (3.35) that  $J_1^*$  is not necessarily lower than  $J_j$ , in particular, when the optimal response of the follower is not unique. The following proposition now provides one sufficient condition under which the leader never does worse in a "Stackelberg game" than in a "Nash game".

**IV-PERSON NONZERO-SUM FINITE GAMES 135** Proposition 3.16 For a given two-person finite game, let  $J_1^*$  and  $J_j$  be as defined before. If  $R_2^1$  is a singleton for each  $71 \in V_1$ , then  $J_U < 4$ . Proof Under the hypothesis of the proposition, assume to the contrary that there exists a Nash equilibrium solution  $\{710 \in T_1, \sim f_2^\circ \in T_2\}$  whose corresponding cost to PI is lower than  $J_1^*$ , i.e.,  $J_1^* > i71(710)72o$  (0 Since  $R_2^1$  is a singleton, let  $T_2 : T_2 \rightarrow T_1$  be the unique mapping introduced in Remark 3.18. Then, clearly,  $720 = T_2 71^\circ$ , and if this is used in (i), together with the RHS of (3.38), we obtain which is a contradiction.  $\min J_V.rV) = J_U > J_1(71M, V^\circ)$ , D Remark 3.20 One might be tempted to think that if a nonzero-sum game admits a unique Nash equilibrium solution and a unique Stackelberg strategy ( $71^*$ ) for the leader, and further if  $R_2^1$  is a singleton, then the inequality of Prop. 3.16 still should hold. This, however, is not true as the following bimatrix game demonstrates: P2 P2 L R  $0^*1 \_1 \ll 1 \ 2 \ PI, B = LR \ 0^*1 \ 1JV \ 2 \ 1 \ PI. \ R \ R$  Here, there exists a unique Nash equilibrium solution, as indicated, and a unique Stackelberg strategy  $71^* = L$  for the leader (PI). Furthermore, the follower's optimal response to  $71^* = L$  is unique (which is  $72 = L$ ). However,  $0 = -1$ . This counterexample indicates that the sufficient condition of  $n^* > 71 \ JN$  Prop. 3.16 cannot be relaxed any further in any satisfactory way. D Since the Stackelberg equilibrium concept is nonsymmetric as to the roles of the players, there is no recursive procedure that can be developed (as in Section 3.5) to determine the Stackelberg solution of dynamic games. The only possibility is the brute-force method, that is, to convert the original two-person finite dynamic game in extensive form into an equivalent normal form (which is basically a bimatrix game), to exhaustively

work out the possible outcomes of the game for each strategy choice of the leader, and to see which of these is the most favorable one for the leader. Since this is a standard application of Def. 3.27 and is no different from the approach adopted in solving the bimatrix games (3.34) and (3.35), we do not provide any further examples here. However, the reader is referred to Section 3.8 for problems that involve derivation of Stackelberg equilibria in single-act and multi-act dynamic games.

136 T. BAŞAR AND G. J. OLSDER The feedback (stagewise) Stackelberg solution Within the context of multi-act dynamic games, the Stackelberg equilibrium concept is suitable for the class of decision problems in which the leader has the ability to announce his decisions at all of his possible information sets ahead of time—an attitude which naturally forces the leader to commit himself irrevocably to the actions dictated by these strategies. Hence, in a sense, the Stackelberg solution is of the prior commitment type from the leader's point of view. There are yet other types of hierarchical multi-act (multi-stage) decision problems, however, in which the leader does not have the ability to announce and enforce his strategy at all levels of play prior to the start of the game, but can instead enforce his strategy on the follower(s) at every level (stage) of the game. Such a hierarchical equilibrium solution, which has the "Stackelberg" property at every level of play (but not globally), is called a feedback Stackelberg solution.<sup>22</sup> We now provide below a precise definition of this equilibrium solution within the context of two-person nonzero-sum feedback games in extensive form (cf. Def. 3.21). To this end, let us adopt the terminology and notation of Section 3.5, and consider  $I$  to be a two-person nonzero-sum feedback game in extensive form with  $K$  levels of play and with  $P_i$  being the first-acting player at each level of play. Let a typical strategy  $j_l \in P$  of  $P_i$  be decomposed as  $(\gamma_i \in r^i > \bullet \blacksquare \blacksquare i \gamma_k \in \wedge^k) >$  where  $\gamma_j$  stands for the corresponding strategy of  $P_i$  at the  $j$ th level of play. Furthermore, for a given strategy  $Y \in P$ , let  $Y^k$  and  $\gamma_{jk}$  denote two truncated versions of  $\gamma_j$ , defined as  $\gamma^k = (\gamma_1, \dots, \gamma_k)$  (3.39a) and  $\gamma_{jk}^* = (\gamma_j, \dots, \gamma_{k-i})$  (3.39b) Then, we have the following. Definition 3.29 Using the preceding notation and terminology, a pair  $\{(T_1, T_2) \in T\}$  constitutes a feedback Stackelberg solution for  $I$  with  $P_i$  as the leader if the following two conditions are fulfilled: (i)  $J_1(\gamma_{i,k}, \text{Phk}; \gamma_{2,k}, \wedge^k) = \min \max / (w_1 \wedge \wedge P_2 \wedge)$  (3-40a) fieri  $\gamma_{j \in l} J W; T_n, *+i$  for all  $\gamma^* \in n, j = 1, \dots, k - 1; i = 1, 2$ ; and with  $k = K, K - 1, \dots, 1$ , where  $R_2\{\{3; \gamma_{i,fe+i}\}$  is the optimal response set of the follower at level  $k$ , <sup>22</sup>We

shall henceforth refer to the Stackelberg equilibrium solution introduced through Defs. 3.26-3.28 as the global Stackelberg equilibrium solution whenever it is not clear from the context which type of equilibrium solution we are working with.

AT-PERSON NONZERO-SUM FINITE GAMES 137 defined by =  
 $\{ \pi_i^k / (U_i^k + T_i^k) \}$ . The feedback Stackelberg cost of the leader,  $J^k(\pi^k; \pi^k)$  (3.40b) (ii)  $R^k(\pi^k; \pi^k)$  is a singleton, with its only element being  $\pi^k$ ,  $k = 1, \dots, K$ . The quantity  $J^k(\pi^k; \pi^k)$  is the corresponding feedback Stackelberg cost of the leader. Relations (3.40a) and (3.40b) now indicate that the feedback Stackelberg solution features the "Stackelberg" property at every level of play, and this can in fact be verified recursively—thus leading to a recursive construction of the equilibrium solution. Before we go into a discussion of this recursive procedure, let us first prove a result concerning the structures of  $R^k(\pi^k; \pi^k)$  and  $J^k(\pi^k; \pi^k)$ — Proposition 3.17 For a two-person feedback game that admits a feedback Stackelberg solution with PI as the leader, (i)  $R^k(\pi^k; \pi^k)$  is not a variant of  $\pi^k$ , and hence it can be written as (ii)  $J^k(\pi^k; \pi^k)$  depends on  $\pi^k$  and  $\pi^k$  only through the singleton information sets  $\pi^k$  of PI at level  $k$ . Proof. Both of these results follow from the "feedback" nature of the two-person dynamic game under consideration, which enables one to decompose the  $K$ -level game recursively into a number of single-act games each of which corresponds to a singleton information set of PI (the leader). In accordance with this observation, let us start with  $k = K$ , in which case the proposition is trivially true. Relation (3.40a) then defines the (global) Stackelberg solution of the single-act game obtained for each information set  $\pi^k$  of PI, and the quantity on the RHS of (3.40a) defines, for  $k = K$ , the Stackelberg cost for the leader in each of these single-act games, which is actually realized because of condition (ii) of Def. 3.29 (see Remark 3.19). Then, the cost pair transferred to level  $k = K - 1$  is well defined (not ambiguous) for each information set  $\pi^k$  of PI. Now, with the  $K$ th level crossed out, we repeat the same analysis and first observe that  $J^k(\pi^k; \pi^k)$  in (3.40) depends only on the information set  $\pi^k$  of PI and on the strategies  $\pi^k$  and  $\pi^k$  of PI and P2, respectively, at level  $k = K - 1$ , since we again have essentially single-act games. Consequently,  $J^k(\pi^k; \pi^k)$  depends only on  $\pi^k$ . If the cost level of PI at  $k = K - 1$  is minimized subject to that constraint,  $\pi^k$  is by hypothesis a solution to that problem, to which a unique response of the follower corresponds. Hence, again a well-defined cost pair is determined for each game corresponding to the singleton

information sets  $r_{i-1}$  of PI, and these are attached to these nodes to provide

138 T. BAŞAR AND G. J. OLSDER cost pairs for the remaining  $(K - 2)$ -level game. An iteration on this analysis then proves the result. The foregoing proof already describes the recursive procedure involved to obtain a feedback Stackelberg solution of a two-person feedback game. The requirement (ii) in Def. 3.29 ensures that the outcomes of the Stackelberg games considered at each level are unambiguously determined, and it imposes an indirect restriction on the problem under consideration, which cannot be tested a priori. This then leads to the conclusion that a two-person nonzero-sum feedback game does not necessarily admit a feedback Stackelberg solution. There are also cases in which a feedback game admits more than one feedback Stackelberg solution. Such a situation might arise if, in the recursive derivation, one or more of the single-act Stackelberg games admit multiple Stackelberg strategies for the leader whose corresponding costs to the follower are not the same. The leader is, of course, indifferent to these multiple equilibria at that particular level where they arise, but this nonuniqueness might affect the (overall) feedback Stackelberg cost of the leader. Hence, if  $\{J_1^1, J_2^1\}$  and  $\{J_1^2, J_2^2\}$  are two feedback Stackelberg solutions of a given game, it could very well happen that  $J_1^1 < J_1^2$ , in which case the leader definitely prefers  $J_1^1$  over  $J_1^2$  and attempts to enforce the components of  $J_1^1$  at each level of the game. This argument then leads to a total ordering among different feedback Stackelberg solutions of a given game, which we formalize below.

**Definition 3.30** A feedback Stackelberg solution  $\{J_1^1, J_2^1\}$  of a two-person nonzero-sum feedback game is admissible if there exists no other feedback Stackelberg solution  $\{J_1^2, J_2^2\}$  with the property  $J_1^2 < J_1^1$ . It should be noted that if a feedback game admits more than one feedback Stackelberg solution, then the leader can definitely enforce the admissible one on the follower since he leads the decision process at each level. A unique feedback Stackelberg solution is clearly admissible. We now provide below an example of a 2-level two-person nonzero-sum feedback game which admits a unique feedback Stackelberg solution. This example also illustrates recursive derivation of this equilibrium solution.

**Example 3.13** Consider the multi-act feedback nonzero-sum game of Fig. 3.9. To determine its feedback Stackelberg solution with PI acting as the leader, we first attempt to solve the Stackelberg games corresponding to each of the four information sets of PI at level 2. The first one (counting from the left) is a trivial game,

admitting the unique equilibrium solution with the corresponding outcome being  $(1,0)$ . The next two admit, respectively, the following bimatrix representations whose unique Stackelberg equilibria are

N-PERSON NONZERO-SUM FINITE GAMES 139 indicated on the matrices:  
 $P_2$   $P_2$   $L$   $R$   $L$   $R$   $3$   $\leq$   $4$   $5$   $L$   $R$   $P_2$   $w$   $2$   $2$   $2$   $P_1$ ,  $B = P_1$ ,  $B$   $L$   $R$   $L$   $R$  &  $\pi$ ,  $I$   $7$ ?  $P_2$  ( $-$   
 $\leq$ )  $i$   $i$   $0$   $P_L$  (it) (Hi)  $L$   $R$   $L$   $R$  Finally, the fourth one admits the trivial bimatrix  
representation  $P_2$   $\wedge$   $P_2$   $P_1$ ,  $L$   $R$   $*$   $-$   $'0$   $1$   $P_1$  (iv) together with the  $2 \times 2$  bimatrix  
representation  $P_2$   $L$   $R$   $P_2$   $A$   $M$   $R$   $3$   $2$   $0$   $u$   $P_1$ ,  $B = L$   $R$   $M$   $R$   $2$   $1$   $3$   $\text{ffi}$   $P_1$  (v)  $L$   $R$   
with the unique Stackelberg solutions being as indicated, in each case. By a  
comparison of the leader's cost levels at (iv) and (v), we conclude that he  
plays  $R$  at the fourth of his information sets at level 2. Hence, in going from the  
second to the first level, we obtain from (i), (ii), (iii) and (v) the unique cost  
pairs at each of  $P_i$ 's information sets to be (counting from the left):  $(1,0)$ ,  $(1,1)$ ,  
 $(0,-1/2)$  and  $(-1,0)$ . Then, the Stackelberg game at the first level has the  
bimatrix representation  $P_2$   $P_2$   $A = L$   $R$   $1$  ( $A$   $1$   $-1$   $P_1$ ,  $B = L$   $R$   $R$   $0$   $G$   $i$   $0$   $P_i$  (vi)  $L$   
 $R$  whose unique Stackelberg solution with  $P_i$  as leader is as indicated. To  
recapitulate, the leader's unique feedback Stackelberg strategy in this feedback  
game is  $\Pi^*(r_i) = R$ ,  $721^*(\%1) = \ll i$ . and the follower's unique optimal  
response strategy is  $\{R \text{ if } u \setminus = L \text{ and } u \setminus \text{---} L, \text{ or } u \setminus = R \text{ and } u \setminus = R \text{ and } u \setminus \pm L, L$   
otherwise . The unique feedback Stackelberg outcome of the game is  $(0, \text{---}$   
 $1/2)$ , which follows from (vi) by inspection.  $\square$

140 T. BAŞAR AND G. J. OLSDER Before concluding our discussion on the  
feedback Stackelberg equilibrium solution, we should mention that the same  
concept has potential applicability to multi-act decision problems in which the  
roles of the players change from one level of play to another, that is, when  
different players act as leaders at different stages. An extension of Def. 3.29 is  
possible so that such classes of decision problems are also covered; but since  
this involves only an appropriate change of indices and an interpretation in the  
right framework, it will not be covered here. Mixed and behavioral  
Stackelberg equilibria The motivation behind introducing mixed strategies in  
the investigation of saddle-point equilibria (in Chapter 2) and Nash equilibria  
(in Section 3.2) was that such equilibria do not always exist in pure strategies,  
whereas within the enlarged class of mixed strategies one can ensure existence  
of noncooperative equilibria. In the case of the Stackelberg solution of two-  
person finite games, however, an equilibrium always exists (cf. Thm. 3.3), and

thus, at the outset, there seems to be no need to introduce mixed strategies. Besides, since the leader dictates his strategy on the follower, in a Stackelberg game, it might at first seem to be unreasonable to imagine that the leader would ever employ a mixed strategy. Such an argument, however, is not always valid, and there are cases in which the leader can actually do better (in the average sense) with a proper mixed strategy, than the best he can do within the class of pure strategies. As an illustration of such a possibility, consider the bimatrix game  $(A, B)$  displayed below:

	P2	L	R
P1	L	1, 1/2	3, 4/3
	R	1/2, 1	1, 1/2

If P1 acts as the leader, then the game admits two pure-strategy Stackelberg equilibrium solutions, which are  $\{L, L\}$  and  $\{R, R\}$ , the Stackelberg outcome in each case being  $(1, 1/2)$ . However, if the leader (P1) adopts the mixed strategy which is to pick L and R with equal probability  $1/2$ , then the average cost incurred to P1 will be equal to  $1/2$ , quite independent of the follower's (pure or mixed) strategy. This value  $J_1 = 1/2$  is clearly lower than the leader's Stackelberg cost in pure strategies, which can further be shown to be the unique Stackelberg cost of the leader in mixed strategies, since any deviation from  $(1/2, 1/2)$  for the leader results in higher values for  $J_1$ , by taking into account the optimal responses of the follower. The preceding result then establishes the significance of mixed strategies in the investigation of Stackelberg equilibria of two-person nonzero-sum games, and demonstrates the possibility that a proper mixed-strategy Stackelberg solution could lead to a lower cost level for the leader than the Stackelberg cost level in pure strategies. To introduce the concept of mixed-strategy Stackelberg equilibrium in mathematical terms, we take the two-person nonzero-sum

**TWO-PERSON NONZERO-SUM FINITE GAMES** 141

finite game to be in normal form (without any loss of generality) and associate with it a bimatrix game  $(A, B)$ . Abiding by the notation and terminology of Section 3.2, we let  $Y$  and  $Z$  denote the mixed-strategy spaces of P1 and P2, respectively, with their typical elements designated as  $y$  and  $z$ . Then, we have the following.

**Definition 3.31** For a bimatrix game  $(A, B)$ , the set  $R_2(y) = \{z \in Z : y'Az \leq y'Bz, \forall z \in Z\}$  (3.42) is the optimal response (rational reaction,) set of P2 in mixed strategies to the mixed strategy  $y \in Y$  of P1.

**Definition 3.32** In a bimatrix game  $(A, B)$  with P1 acting as the leader, a mixed strategy  $y^* \in Y$  is called a mixed Stackelberg equilibrium strategy for the leader if  $\max_{y \in Y} y'Az = \inf_{z \in R_2(y^*)} y^*Az = J_1^*$ . (3.43)

The quantity  $J_1^*$  is the Stackelberg cost of the leader in mixed strategies. It should be noted that the "maximum" in (3.43)

always exists since, for each  $y \in Y$ ,  $y'Az$  is continuous in  $z$ , and  $R_2(y)$  is a closed and bounded subset of  $Z$  (which is a finite dimensional simplex). Hence,  $J_1^*$  is a well-defined quantity. The "infimum" in (3.43), however, cannot always be replaced by a "minimum", unless the problem admits a mixed Stackelberg equilibrium strategy for the leader. The following example now demonstrates the possibility that a two-person finite game might not admit a mixed-strategy Stackelberg strategy even though  $J_1^* < J_1^*$ . Example 3.14

Consider the following modified version of the bimatrix game of (3.41):  $P_2$   $P_1$   
 $A = \begin{pmatrix} L & R \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ ,  $B = \begin{pmatrix} L & R \\ 1/2 & 1 \\ 1 & 1/3 \end{pmatrix}$ . With  $P_1$  as the leader, this bimatrix game also admits two pure-strategy Stackelberg equilibria, which are  $\{L, L\}$  and  $\{R, R\}$ , the Stackelberg cost for the leader being  $J_1^* = 1$ . Now, let the leader adopt the mixed strategy  $y = (y_i, (1 - j/i))'$ , under which  $J_2$  is  $J_2(y, z) = y'Bz = -y_i + 2/1 + 2/1 + z/1 + 3'$  where  $z = (z, (1 - z))'$  denotes any mixed strategy of  $P_2$ . Then, the mixed-strategy optimal response set of  $P_2$  can readily be determined as  $\{z = (1, 0)\}$  if  $j/i > 4/7$ ,  $R_2(y) = \{U = (0, 1)\}$  if  $2/1 < 4/7$ ,  $Z$  if  $y_i = 4/7$ .

142 T. BAŞAR AND G. J. OLSDER Hence, for  $j/i > 4/7$ , the follower chooses "column 1" with probability 1, and this leads to an average cost of  $J_1 = y_i$  for  $P_1$ . For  $y_i < 4/7$ , on the other hand,  $P_2$  chooses "column 2" with probability 1, which leads to an average cost level of  $J_1 = (1 - 2/i)$  for  $P_1$ . Then, clearly, the leader will prefer to stay in this latter region; in fact, if he employs the mixed strategy  $y = (4/7 - e, 3/7 + e)'$  where  $e > 0$  is sufficiently small, his realized average cost will be  $J_1 = 3/7 + e$ , since then  $P_2$  will respond with the unique pure-strategy  $z = R$ . Since  $e > 0$  can be taken as small as possible, we arrive at the conclusion that  $J_1^* = \lim_{e \rightarrow 0} J_1 = J_1^*$ . In spite of this fact, the leader does not have a mixed Stackelberg strategy since, for the only candidate  $y^\circ = (4/7, 3/7)$ ,  $R_2(y^\circ) = Z$ , and therefore  $\max_{z \in Z} y^\circ Az = 4/7$  which is higher than  $J_1^*$ .  $\square$  The preceding example thus substantiates the possibility that a mixed Stackelberg strategy might not exist for the leader, but he can still do better than his pure Stackelberg cost  $J_1^*$  by employing some sub-optimal mixed strategy (such as the one  $y = (4/7 - e, 3/7 + e)'$  in Example 3.14, for sufficiently small  $e > 0$ ). In fact, whenever  $J_1^* < J_1^*$ , there will always exist such an approximating mixed strategy for the leader. If  $J_1^* = J_1^*$ , however, it is, of course, reasonable to employ the pure Stackelberg strategy which always exists by Thm. 3.3. The following proposition now verifies that  $J_1^* < J_1^*$  and  $J_1^* = J_1^*$  are the only two possible relations we can have between  $J_1^*$  and  $J_1^*$ ; in other words, the



inequality  $J_1^* > J_1^*$  never holds. Proposition 3.18 For every two-person finite game, we have  $J_u < J_u$ . (3.44) Proof. Let  $Y_q$  denote the subset of  $Y$  consisting of all one-point distributions. Analogously, define  $Z_q$  as comprised of one-point distributions in  $Z$ . Note that  $V_b$  is equivalent to  $T_1$ , and  $Z_S$  is equivalent to  $T_2$ . Then, for each  $y \in V_b$   $\min_{z \in Z} y'Az = \min_{z \in Z_q} y'Az$  since any minimizing solution in  $Z$  can be replaced by an element of  $Z_q$ . This further implies that, for each  $y \in Y_q$ , elements of  $R_2(y)$  are probability distributions on  $R_2(y)$ , where the latter set is defined by (3.42) with  $Z$  replaced by  $Z_q$ . Now, since  $Y_0 \subset Y$ ,  $J_1^* = \min_{y \in Y} \max_{z \in R_2(y)} y'Az < \min_{y \in Y_0} \max_{z \in R_2(y)} y'Az$  and further, because of the cited relation between  $R_2(-)$  and  $R_2(-)$ , the latter quantity is equal to  $\min_{y \in Y_0} \max_{z \in R_2(y)} y'Az$  which, by definition, is  $J_1^*$ , since  $R_2(y)$  is equivalent to the pure-strategy optimal response set of the follower, as defined by (3.36). Hence,  $J_1^* < J_1^*$ . D

**IV-PERSON NONZERO-SUM FINITE GAMES 143** Computation of a mixed-strategy Stackelberg equilibrium (whenever it exists) is not as straightforward as in the case of pure-strategy equilibria, since the spaces  $Y$  and  $Z$  are not finite. The standard technique is first to determine the minimizing solution(s) of  $\min_{z \in Z} y'Az$  as functions of  $y \in Y$ . This will lead to a decomposition of  $Y$  into subsets (regions), on each of which a reaction set for the follower is defined. (Note that in the analysis of Example 3.14,  $Y$  has been decomposed into three regions.) Then, one has to minimize  $y'Az$  over  $y \in Y$ , subject to the constraints imposed by these reaction sets, and under the stipulation that the same quantity is maximized on these reaction sets whenever they are not singletons. This brute-force approach also provides approximating strategies for the leader, whenever a mixed Stackelberg solution does not exist, together with the value of  $J_1^*$ . If the two-person finite game under consideration is a dynamic game in extensive form, then it is more reasonable to restrict attention to behavioral strategies (cf. Def. 2.9). Stackelberg equilibria within the class of behavioral strategies can be introduced as in Defs. 3.31 and 3.32, by replacing the mixed-strategy sets with the behavioral-strategy sets. Hence, using the terminology and notation of Def. 2.9, we have the following counterparts of Defs. 3.31 and 3.32, in behavioral strategies. Definition 3.33 Given a two-person finite dynamic game with behavioral-strategy sets  $(T_1^2)$  and average cost functions  $(J_1, J_2)$  the set  $\&tf) = \{z_2 \in T_2 : J_2(\gamma^2 z_2) < J_2(\gamma^2 z_2), \forall z_2 \in T_2\}$ , (3.45) is the optimal response (rational reaction) set of  $P_2$  in behavioral strategies to the behavioral strategy  $\gamma_1 \in T_1$  of  $P_1$ . Definition 3.34 In a two-person finite

dynamic game with PI acting as the leader, a behavioral strategy  $\gamma_1^* \in T_1$  is called a behavioral Stackelberg equilibrium strategy for the leader if  $\sup J_1(\gamma_1^*, \gamma_2) = \inf \sup J_1(\gamma_1, \gamma_2^*) = J_1^*$ . (3.46)  $\gamma_2^* \in T_2(y^*)$  The quantity  $J_1^*$  is the Stackelberg cost of the leader in behavioral strategies. Remark 3.21 The reason why we use "supremum" in (3.46), instead of "maximum", is because the set  $R^{\wedge 1}$  is not necessarily compact. The behavioral Stackelberg cost for the leader is again a well-defined quantity, but a behavioral Stackelberg strategy for the leader does not necessarily exist.  $\square$  We now conclude our discussion on the notion of behavioral Stackelberg equilibrium by presenting an example of a single-act dynamic game that admits a Stackelberg solution in behavioral strategies.

144 T. BAŞAR AND G. J. OLSDER 2,0 1,1 1,1 2,0 0,2/3 2,1 Figure 3.10: The two-person single-act game of Example 3.15. Example 3.15 Consider the two-person single-act game whose extensive form is depicted in Fig. 3.10. With P2 acting as the leader, it admits two pure-strategy Stackelberg solutions:  $\gamma_2^*(2) = (L \text{ if } u_1 = f_i, 1 \text{ if } u_1 = \bar{f}_i \text{ otherwise, } \gamma_1^* \in T_1$  and  $\gamma_2^*(r_2) = L, IU(r, 1) = R, v$  with the unique Stackelberg outcome in pure strategies being  $(0, \bar{f}_i)$ . Now, let us allow the leader to use behavioral strategies. The bimatrix game at the first (counting from the left) of the leader's information sets is  $P_2 \begin{matrix} L & M \\ 1 & 1 \end{matrix} P_1, B = \begin{matrix} L & M \\ 0 & 1 \end{matrix} P_1, L \ R \ L \ R$  for which  $J_2^* = 1/2$ . The mixed Stackelberg strategy exists for the leader, given as  $2 \text{ w.p. } 1/2, \gamma_1^* \in T_1$  w.p.  $1/2$ , and any mixed strategy is an optimal response strategy for PI (the follower) yielding him an average cost level of  $J_1 = 3/2$ . We thus observe that if the leader can force the follower to play L or M, his average cost can be pushed down to  $1/2$  which is lower than  $J_2^* = 2/3$ , and the follower's cost is then pushed up to  $3/2$  which is higher than the cost level of  $J_1 = 0$  which would have been attained if P2 were allowed to use only pure strategies. There indeed exists such a (behavioral) strategy for P2, which is given as  $\{L \text{ w.p. } 1/2, 1 \text{ if } u_1 = \bar{f}_i \text{ or } M, \dots R \text{ w.p. } 1/2 \}$  ( $\gamma_2^*$ )  $R \text{ w.p. } 1 \text{ if } u_1 = R$ , and the behavioral Stackelberg cost for the leader is  $J_1^* = 1/2$ . ( $\ll$ )

iV-PERSON NONZERO-SUM FINITE GAMES 145 It should be noted that the follower (PI) does not dare to play  $u_1 = R$  since, by announcing (i) ahead of time, the leader threatens him with an undesirable cost level of  $J_1 = 2$ . We now leave it to the reader to verify that the behavioral strategy (i) is indeed a Stackelberg strategy for the leader and the quantity (ii) is his Stackelberg cost

in behavioral strategies. This verification can be accomplished by converting the original game in extensive form into equivalent normal form and then applying Def. 3.34. Another alternative method is to show that (ii) is the mixed-strategy Stackelberg cost for the leader, and since (i) attains that cost level, it is also a behavioral Stackelberg strategy, by the result of Problem 21 (Section 3.8).

a Many-player games Heretofore, we have confined our investigation on the Stackelberg concept in nonzero-sum games to two-person games wherein one of the players is the leader and the other one is the follower. Extension of this investigation to  $N$ -player games could be accomplished in a number of ways; here we discuss the possible extensions to the case when  $N = 3$ , with the further generalizations to  $N > 3$  being conceptually straightforward. For three-person nonzero-sum games, we basically have three different possible modes of play among the players within the framework of noncooperative decision making. (1) There are two levels of hierarchy in decision making—one leader and two followers. The followers react to the leader's announced strategy by playing according to a specific equilibrium concept among themselves (for instance, Nash). (2) There are still two levels of hierarchy in decision making—now two leaders and one follower. The leaders play according to a specific equilibrium concept among themselves, by taking into account possible optimal responses of the follower. (3) There are three levels of hierarchy. First  $P_1$  announces his strategy, then  $P_2$  determines his strategy by also taking into account possible responses of  $P_3$  and enforces this strategy on him, and finally  $P_3$  optimizes his cost function in view of the announced strategies of  $P_1$  and  $P_2$ . Let us now provide precise definitions for the three types of hierarchical equilibria outlined above. To this end, let  $J_1^1, J_2, J_3$  denote the cost function of  $P_i$  corresponding to the strategy triple  $\{s_1 \in T_1, s_2 \in T_2, s_3 \in T_3\}$ . Then, we have the following.

**Definition 3.35** For the three-person finite game with one leader ( $P_1$ ) and two followers ( $P_2$  and  $P_3$ ),  $s_1^* \in T_1$  is a hierarchical equilibrium strategy for the leader if  $\max_{s_1} J_1(s_1, s_2, s_3) = \min_{s_2, s_3} \max_{s_1} J_1(s_1, s_2, s_3)$ , (3.47a)

146 T. BAŞAR AND G. J. OLSDER where  $R_1(s_1)$  is the optimal response set of the followers' group and is defined for each  $s_1 \in T_1$  by  $R_1(s_1) = \{(s_2, s_3) : J_2(s_1, s_2, s_3) < J_2(s_1, s_2, s_3) \text{ and } J_3(s_1, s_2, s_3) < J_3(s_1, s_2, s_3), \forall s_2 \in T_2, s_3 \in T_3\}$ . (3.47b) Any  $(s_2^*, s_3^*) \in R_1(s_1)$  is a corresponding optimal strategy pair for the followers' group.

**Definition 3.36** For the three-person finite game with two leaders ( $P_1$  and  $P_2$ ) and one follower ( $P_3$ ),  $s'' \in T_3$  is a hierarchical

equilibrium strategy for  $P_i$  ( $i = 1, 2$ ) if  $\max_{x_i} J_i(x_i, x_{-i}) = \min_{x_{-i}} \max_{x_i} J_i(x_i, x_{-i})$ , (3.48a) s.p.T\* 2./V\*,72\*,73)= min max, J2(7U,72,73), (3.48b)  
 $73 \in R_3(7^*;72^*)$   $72 \in R_2(71^*;72)$  where  $i^*(7x;72)$  is the optimal response set of the follower and is defined for each  $(7X,72) \in T_1 \times T_2$  by  $R^i(W) = \{x_i \in T_i : J_i(x_i, x_{-i}) \geq J_i(x_i', x_{-i}), \forall x_i' \in T_i\}$ . (3.48c)  $j_4nj/$  strategy  $73^* \in S_3(71^*;72^*)$  is a corresponding optimal strategy for the follower (P3).  
**Definition 3.37** For the three-person finite game with three levels of hierarchy in decision making (with  $P_1$  enforcing his strategy on  $P_2$  and  $P_3$ , and  $P_2$  enforcing his strategy on  $P_3$ ),  $71^* \in T_1$  is a hierarchical equilibrium strategy for  $P_1$  if  $t_1, \dots, 1^* 2, \dots, 3 \max_{x_1} \max_{x_2, x_3} J(x_1, x_2, x_3) = \min_{x_2, x_3} \max_{x_1} J(x_1, x_2, x_3)$ , (3.49a)  $71 \in R_1(72,73)$  where  $S^i(W) = \{x_i \in T_i : J_i(x_i, x_{-i}) \geq J_i(x_i', x_{-i}), \forall x_i' \in T_i\}$ , (3.49b)  $73 \in S_3(71;72)$   $S_3(W) = \{x_3 \in T_3 : J_3(x_3, x_{-3}) \geq J_3(x_3', x_{-3}), \forall x_3' \in T_3\}$ . (3.49c)  $j_4nt/$   $72^* \in S_2(71^*)$  is a corresponding optimal strategy for  $P_2$ , and any  $73 \in S_3(71^*, 72^*)$  is an optimal strategy for  $P_3$  corresponding to the pair  $(71^*, 72^*)$ . Several simplifications in these definitions are possible if the optimal response sets are singletons. In the case of Def. 3.35, for example, we would have the following.

**AT-PERSON NONZERO-SUM FINITE GAMES 147** Proposition 3.19 In the three-person finite game of Def. 3.35, if  $R^i(x_{-i})$  is a singleton for each  $x_{-i} \in T_{-i}$ , then there exist unique mappings  $T_2 : T_1 \rightarrow T_2$  and  $T_3 : T_1 \rightarrow T_3$  such that and  $J_1(71^*, T_2(71^*), T_3(71^*)) = \min_{x_2, x_3} J_1(71, x_2, x_3) < J_2(71, T_2(71), T_3(71))$ ,  $J_3(71, T_2(71), T_3(71)) < J_3(71, T_2(71), T_3(71))$  (3.50a) (3.50b) (3.50c) and for each  $71 \in T_1$ . Analogous results can be obtained in the cases of Defs. 3.36 and 3.37, if the optimal response sets are singletons. The following example now illustrates all these cases with singleton optimal response sets.  
**Example 3.16** Using the notation of Section 3.3, consider a three-person game with possible outcomes given as  $(a_i, i, i) > a_i, i, i$   $(a_1, 2, 1) > a_1, 2, 1$   $(a_2, 1, 1) > a_2, 1, 1$   $(a_2, 2, 1) > a_2, 2, 1$   $(a_1, 1, 2) > a_1, 1, 2$   $(a_1, 2, 2) > a_1, 2, 2$   $(a_2, 1, 2) > a_2, 1, 2$   $(a_2, 2, 2) > a_2, 2, 2$   $(1, 0, 0)$ ,  $(2, 1, -1)$ ,  $(0, 0, 1)$ ,  $(2, -1, 1)$ ,  $(1, -1, 1)$ ,  $(0, 1, 0)$ ,  $(-1, 2, 0)$ ,  $(0, 1, 2)$ , and determine its hierarchical equilibria (of the three different types discussed above). 1. One leader ( $P_1$ ) and two followers ( $P_2$  and  $P_3$ ). For the first alternative of  $P_1$  (i.e.,  $n = 1$ ), the bimatrix game faced by  $P_2$  and  $P_3$  is  $P_2$   $(1, 0, 0) > (1, -1, 1)$   $(2, 1, -1)$   $(0, 1, 0)$   $P_3$  which clearly admits the unique Nash equilibrium solution  $\{n^* = 1, n^* = 1\}$  whose cost to the leader is  $J_1 = 1$ . For  $n = 2$ , the bimatrix game is  $P_2$   $(0, 0, 1) > (-1, 2, 0)$   $(2, -1, 1) > (0, 1, 2)$   $P_3$  with a unique Nash equilibrium

solution  $\{n^* = 2, r_{ij} = 1\}$  whose cost to  $P_1$  is  $J_1 = 2$ . This determines the mappings  $T_2$  and  $T_3$  of Prop. 3.19 uniquely. Now, since  $1 < 2$ , the leader has a unique hierarchical equilibrium strategy that

148 T. BAŞAR AND G. J. OLSDER satisfies (3.50a), which is  $7_1^*$  — "alternative 1", and the corresponding unique optimal strategies for  $P_2$  and  $P_3$  are  $y_2^* =$  "alternative 1",  $y_3^* =$  "alternative 1", respectively. The unique equilibrium outcome is  $(1,0,0)$ .

2. Two leaders ( $P_1$  and  $P_2$ ) and one follower ( $P_3$ ). There are four pairs of possible strategies for the leaders' group, which are listed below together with the corresponding unique optimal responses of the follower and the cost pairs incurred to  $P_1$  and  $P_2$ :  $\{m=1, n_2=1\} : n^* = 1 - (1,0)$ ,  $\{n_1=1, n_2=2\} : n^* = 1 \rightarrow (2,1)$ ,  $\{n_1=2, n_2=1\} : n^* = 2 - (-1,2)$ ,  $\{m=2, n_2=2\} : n_j = 1 \rightarrow (2, -1)$ . The two leaders are faced with the bimatrix game  $P_2$   $(1,0) (-1,2) (2,1) (2,-1)^*$  whose Nash solution is uniquely determined as  $\{n^* = 2, n^* = 2\}$ . Hence, the hierarchical equilibrium strategies of the leaders are unique:  $7_1^* =$  "alternative 2",  $7_2^* =$  "alternative 2". The corresponding unique optimal strategy of  $P_3$  is  $7_3^* =$  "alternative 1", and the unique equilibrium outcome is  $(2, -1, 1)$ .

3. Three levels of hierarchy (first  $P_1$ , then  $P_2$  and finally  $P_3$ ). We first observe that  $S_2^1$  is a singleton and is uniquely determined by the mapping  $P_2 : T_1 \rightarrow T_2$  defined by  $[2$  if  $r_{ii} = 2$ . Hence, if  $P_1$  chooses  $n_i = 1$ , his cost will be  $J_1 = 1$ , and if he chooses  $n_i = 2$ , his cost will be  $J_1 = 2$ . Since  $1 < 2$ , this leads to the conclusion that  $P_1$ 's unique hierarchical equilibrium strategy is  $7_1^* =$  "alternative 1" and the corresponding unique optimal strategies of  $P_2$  and  $P_3$  are  $j_2^* =$  "alternative 1" and  $7_3^* =$  "alternative 1", respectively. The unique equilibrium outcome is  $(1,0,0)$ .  $\square$

### 3.7 Nonzero-Sum Games with Chance Moves

In this section, we briefly discuss an extension of the general analysis of the previous sections to the class of nonzero-sum games which incorporate chance moves, that is, games wherein the final outcome is determined not only by the actions of the players, but also by the outcome of a chance mechanism. A special class of such decision problems was already covered in Section 2.6 within the context of zero-sum games, and the present analysis in a way builds on the material of Section 2.6 and extends that analysis to nonzero-sum games under the Nash and Stackelberg solution concepts. To this end, let us first provide a

N-PERSON NONZERO-SUM FINITE GAMES 149 precise definition of an extensive form description of a nonzero-sum finite game with chance moves.

Definition 3.38 An extensive form of an  $N$ -person nonzero-sum finite game with chance moves is a tree structure with (i) a specific vertex indicating the starting point of the game, (ii)  $N$  cost functions, each one assigning a real number to each terminal vertex of the tree, where the  $i$ th cost function determines the loss incurred to  $P_i$  for each possible set of actions of the player together with the possible choices of nature, (iii) a partition of the nodes of the tree into  $N + 1$  player sets (to be denoted by  $N_i$  for  $P_i$ ,  $i \in N$ , and by  $N^\circ$  for nature), (iv) a probability distribution, defined at each node of  $N^\circ$ , among the immediate branches (alternatives) emanating from that node, (v) a subpartition of each player set  $N_i$  into information sets  $\{I_i^k\}$  such that the same number of immediate branches emanates from every node belonging to the same information set and no node follows another node in the same information set. For such a game, let us again denote the (pure) strategy space of  $P_i$  by  $P_i$ , with a typical element designated as  $s_i$ . For a given  $N$ -tuple of strategies  $\{s_j \in P_j; j \in N\}$  the cost function  $J_i(s_1, \dots, s_N)$  of  $P_i$  is in general a random quantity, and hence the real quantity of interest is the one obtained by taking the average value of  $J_i$  over all relevant actions of nature and under the a priori known probability distribution. Let us denote that quantity by  $L_i(s_1, \dots, s_N)$  which in fact provides us with a normal form description of the original nonzero-sum game with chance moves in extensive form. Apart from a difference in notation, there is no real difference between this description and the normal form description studied heretofore in this chapter, and this observation readily leads to the conclusion that all concepts introduced and general results obtained in the previous sections on the normal form description of nonzero-sum games are also equally valid in the present context. In particular, the concept of Nash equilibrium for  $N$ -person nonzero-sum finite games with chance moves can be introduced through Def. 3.12 by merely replacing  $J_i$  with  $L_i$  ( $i = 1, \dots, N$ ). Analogously, the Stackelberg equilibrium solution in two-person nonzero-sum finite games with chance moves can be introduced by Defs. 3.26-3.29. Rather than list all these apparent analogies, we shall consider, in the remaining part of this section, two examples which illustrate derivation of Nash and Stackelberg equilibria in static and dynamic games that incorporate chance moves. Example 3.17 Consider the two-person static nonzero-sum game depicted in Fig. 3.11. If nature picks the left branch, then the bimatrix game faced by the

sum game that incorporates a chance move. players is P2 L R (4,0) (0,0) (0,4) (-4,4) PI, (i) L R whereas if nature picks the right branch, the relevant bimatrix game is P2 L 1 (8,-4) I (4,-8)~| R I (4,4) 1 (0,-4)1 ri" L R («) Since the probabilities with which nature chooses the left and right branches are 1/4 and 3/4, respectively, the final bimatrix game of average costs is P2 L R (7,-3) (3,3) (3,-5) (-1,-2) PI. (iii) R This game admits a unique Nash equilibrium solution in pure strategies, which is  $.i^* 7 R, Y^* = R$  with an average outcome of  $(-1, -2)$ . To determine the Stackelberg solution with PI as the leader, we again focus our attention on the bimatrix game (iii) and observe that the Stackelberg solution is in fact the same as the unique Nash solution determined above. D Example 3.18 Let  $x_0$  be a discrete random variable taking the values  $\%o = -1$  w.p. 1/4,  $0$  w.p. 1/2,  $1$  w.p. 1/4.

TV-PERSON NONZERO-SUM FINITE GAMES 151 We view  $x_0$  as the random initial state of some system under consideration. PI, whose decision variable (denoted by  $u_1$ ) assumes the values  $-1$  or  $+1$ , acts first, and as a result the system is transferred to another state ( $x_1$ ) according to the equation  $x_1 = x_0 + u_1$ . Then P2, whose decision variable  $u_2$  can take the same two values ( $-1$  and  $+1$ ), acts and transfers the system to some other state ( $x_2$ ) according to  $x_2 = x_1 + u_2$ . It is further assumed that, in this process, PI does not make any observation (i.e., he does not know what the real outcome of the chance mechanism (value of  $x_2$ ) is), whereas P2 observes  $\text{sgn}(x_1)$ , where  $1 \setminus A \text{sgn}(x) =$  This observation provides P2 with partial information concerning the value of  $x_0$  and the action of PI. The cost structure of the problem is such that  $P_i$  wishes to minimize  $J_i, i = 1, 2$ , where  $J_1 = g(x_2) + u_1, J_2 = g(x_2) - u_2$  with  $f$   $-1$ , if  $x = -3, -1, 2, g(x) \wedge 0$ , if  $x = 0, 3, \{ 1$ , if  $x = -2, 1$ . We now seek to obtain Nash equilibria and also Stackelberg equilibria of this single-act game with either player as the leader. To this end, let us first note that this is a finite nonzero-sum game which incorporates a chance move, but it has not been cast in the standard framework. Though, by spelling out all possible values of  $J_1$  and  $J_2$ , and the actions of the players leading to those values, an extensive tree formulation can be obtained for the game, which is the one depicted in Fig. 3.12. This extensive form also clearly displays the information sets and thereby the number of possible strategies for each player. Since PI gains no information, he has two possible strategies:  $7_i = -1, 7_2 = +1$ - P2, on the other hand, has eight possible strategies which can be listed as  $7_f(V) = -i, t_2(V) = +i, 7_3(V) = 1; \wedge \text{ilin}_2) = \{ +1 \text{ if } 2 = 'i' -y_2(n_2) = \{ \sim 1A \text{ in } 7_4 \{T\} ' \setminus -1, \text{ otherwise}$

,  $7sW' \setminus +1$ , otr 2 - n2 otherwise ,

152 T. BAŞAR AND G. J. OLSDER  $uz = -1 -2, 0 -2, -2 0, 0 2, 0 0, 2 -1, -1 1, 1 0, -2 -2, 0 0, 0 2, 2 1, -1$  Figure 3.12: Extensive tree formulation for the single-act two-person game of Example 3.18.  $\sim 2(r \setminus -1 +1, \text{if } ?2 = \wedge 1, 2 2 f -1, \text{if } 7eI'M - \wedge \_1 ?$  otherwise ,  $7?WJ (+1, c 78fo) = ( \_lj0 r?2 = rj32, \text{otherwise } ri2 = n1$  otherwise . Then, the equivalent bimatrix representation in terms of average costs is  $P2 A = B = 1 2 1 2 -1 1 -1 3/4 1 2 1 1 -1 -5/4 -1/2 3/4 3 1 -5/4 -3/2 1 -3/2 1/4 -1/2 3/2 4 5 6 P2 -1 1 -1 -3/4 1 1 -1 1 -1 3/2 7 1 -3/4 -1 1/4 8 - PI PI$ , Since the first row of matrix A strictly dominates the second row, PI clearly has a unique (permanent) optimal strategy  $71^* = -1$  regardless of whether it is a Nash or a Stackelberg game. P2 has four strategies that are in Nash or Stackelberg equilibrium with  $71^*$ , which are  $7!$ ,  $74$ ,  $75$  and  $77$ . Hence, we have four sets of Nash equilibrium strategies, with the equilibrium outcome being either  $(-1, -1)$  or  $(-3/2, -1)$ . If PI is the leader, his Stackelberg strategy is  $71^* = -1$  and the Stackelberg average cost (as defined in Def. 3.27) is  $L_u = -1$ . He may, however, also end up with the more favorable average cost level of  $L_1 = -3/2$ , depending on the actual play of P2. If P2 is the leader, he has four Stackelberg strategies ( $7I$ ,  $74$ ,  $75$  and  $77$ ), with his unique Stackelberg average cost being  $L_2^* = -1$ . To summarize the results in terms of the terminology of the original problem, regardless of the specific mode of play (i.e., whether it is Nash or Stackelberg), PI always chooses  $u_1 = -1$ , and P2 chooses  $u_2 = -1$  if  $x_1 < 0$ , and is indifferent between his two alternatives  $(-1$  and  $+1)$  if  $x_1 > 0$ . D

#### iV-PERSON NONZERO-SUM FINITE GAMES 153 3.8 Problems 1.

Determine the Nash equilibrium solutions of the following bimatrix game, and investigate the admissibility properties of these equilibria:  $A = P2 8 6 12 0 0 -4 1 0 0 4 8 6 0 12 0 PI, B = P2 4 0 0 0 0 0 0 4 4 1 0 4 0 0 0 PI$ . 2. Verify that minimax values of a bimatrix game are not lower (in an ordered way) than the pair of values of any Nash equilibrium outcome. Now, construct a bimatrix game wherein minimax strategies of the players are also their Nash equilibrium strategies, but in which the minimax values are not the same as the ordered values of the Nash equilibrium outcome. 3. Obtain the mixed Nash equilibrium solution of the following bimatrix game:  $P2 P2 A = 10 -1 -2 1 PI, B = -5 1 2 -1 PI$ . What is the mixed Nash equilibrium solution of the game  $(-A, -B)$ ? If a bimatrix game  $\{A, B\}$  admits multiple minimax strategies for a



player, then these can be ordered according to the notion of admissibility (which is different from the admissibility of Nash equilibrium a la Def. 3.3). A minimax strategy for P1, say "row i", is said to be admissible if there exists no other minimax strategy, say "row j", with the property  $a_{ik} > a_{jk} \forall k = 1, \dots, n$  with the inequality being strict for at least one k. An admissible minimax strategy for P2 can be defined analogously. (i) Does every bimatrix game admit admissible minimax strategies for P1 and P2? If they do, are they necessarily unique? (ii) Obtain the set of all admissible minimax strategies for P1 and P2 in the following bimatrix game: P2 P2 1 3 1 2 4 1 3 0 3 P1, B 4 0 0 0 4 0 0 3 4 P1.

5. Let  $(A, B)$  be an  $(m \times n)$  bimatrix game admitting a unique Nash equilibrium solution in mixed strategies, say  $(y^*, z^*)$ . Prove that the number

154 T. BAŞAR AND G. J. OLSDER of nonzero entries of  $y^*$  and  $z^*$  are equal. How would this statement have to be modified if the number of mixed-strategy equilibrium solutions is greater than one? 6. Prove (or disprove) the following conjecture: Every bimatrix game that admits more than one pure Nash equilibrium solution with different equilibrium outcomes necessarily also admits a Nash equilibrium solution in proper mixed strategies (that is, mixed strategies that are not pure). 7. Determine all Nash equilibria of the three-person nonzero-sum game, with  $m_1 = m_2 = m_3 = 2$ , whose possible outcomes are given below:  $(a_{1,1,1}, a_{2,1,1}, a_{2,2,1}, a_{1,1,2}, a_{1,2,2}, a_{2,1,2}, a_{2,2,2}, a_{1,2,1}, a_{2,1,1}, a_{2,2,1}, a_{1,1,2}, a_{1,2,2}, a_{2,1,2}, a_{2,2,2}, a_{1,1,1}, a_{1,2,1}, a_{2,2,1}, a_{1,1,2}, a_{1,2,2}, a_{2,1,2}, a_{2,2,2}) = (0,0,1), (-2,1,2), (0,-3,1), (-1,1,-2), (-2,1,1), (-2,0,1), (-1,-2,2), (1,0,0)$ . 8. Obtain the minimax strategies of the players in the three-person game of Problem 7. What are the security levels of the players? 9. Consider the following bimatrix game, which admits three pure-strategy Nash equilibria: P2 P2 1 1 2 3 2 2 1 1 2 1 2 1 Which of these three Nash equilibria are (i) perfect, (ii) proper? (Note: See the discussion in subsection 3.5.5.) 10. This problem is concerned with a specific TV-person zero-sum game<sup>23</sup> ( $N > 2$ ) which is called Mental Poker after Epstein (1967). Each of the TV players chooses independently an integer out of the set  $\{1, 2, \dots, n\}$ . The player who has chosen the lowest unique integer is declared the winner and he receives  $N - 1$  units from the other players each of whom loses one unit. If there is no player who has chosen a unique integer, then no one wins the game and everybody receives zero units. If for instance  $N = 3, n = 5$  <sup>23</sup>In terms of the notation of Section 3.3, the game is zero-sum if  $\sum_{i=1}^N c_i a_{ij} = 0, \forall j \in M_j, j \in N$ .

TV-PERSON NONZERO-SUM FINITE GAMES 155 and the integers chosen are 2, 2 and 5, then the player who has chosen 5 is the winner. Show that for  $N = 2$ ,  $n$  arbitrary, there exists a unique saddle-point strategy for both players, which is to choose integer 1 with probability 1. Next consider the case  $N = 3$ . A mixed strategy for  $P_i$  will be denoted by  $y^* = \{y_1^*, \dots, y_n^*\}$ ,  $i = 1, 2, 3$ , with  $y_j^* > 0$  and  $\sum_{j=1}^n y_j^* = 1$ . We seek a Nash equilibrium solution  $\{y_1^*, y_2^*, y_3^*\}$  with the property  $y_1^* = y_2^* = y_3^*$ . Show that there exists a unique such Nash equilibrium solution, which is given by  $y_j^* = (1/2)^{j-1}$ ,  $j = 1, \dots, n-1$ ,  $y_n^* = (1/2)^{n-1}$ . Generalize these results to  $N > 3$ .

11. Obtain all pure-strategy Nash equilibria of the following single-act games in extensive form. Which of these equilibria are (i) admissible, (ii) of the delayed commitment type?  $(0, 1, 2, 1, 1, 3, -1, 2, 1, 1, 0, 0, 0, 1, 2, 1, 1, 3, -1, 2, 1, -2, 0, -2)$  (a) (b)

12. Obtain all pure-strategy Nash equilibria of the following three-player single-act game wherein the order in which  $P_2$  and  $P_3$  act is a variant of  $P_1$ 's strategy. (Hint: A recursive procedure still applies here for each fixed strategy of  $P_i$ .)

13. Investigate whether the three-person single-act game in nested extensive form depicted in Fig. 3.13 admits a Nash equilibrium of the delayed commitment type.

156 T. BAŞAR AND G. J. OLSDER Figure 3.13: Single-act game of Problem 13.

14. Determine all multi-act nonzero-sum feedback games which are informationally inferior to the feedback game whose extensive form is depicted in Fig. 3.9. Obtain their delayed commitment type feedback Nash equilibria.

15. Investigate whether the multi-act game of Fig. 3.14 admits pure-strategy Nash equilibria, by transforming it into equivalent normal form. Figure 3.14: Multi-act game of Problem 15.

16. Determine the perfect Nash equilibria (in pure and mixed strategies) of the multi-act game of Problem 15.

17. Determine the Stackelberg strategies for the leader ( $P_1$ ) and also his Stackelberg costs in the single-act games of Problem 11. Solve the same game problems under the Stackelberg mode of play with  $P_2$  acting as the leader.

18. Let  $I$  be a two-person single-act finite game in extensive form in which  $P_2$  is the first-acting player, and denote the Stackelberg cost for the leader ( $P_1$ ) as  $J^*$ . Let  $II$  be another two-person single-act game that is informationally inferior to  $I$ , and whose Stackelberg cost for the leader ( $P_1$ ) is  $J_j^*$ . Prove the inequality  $J^* < J_j^*$ .

N-PERSON NONZERO-SUM FINITE GAMES 157

19. Determine the Stackelberg solution of the two-person dynamic game of Problem 15 with (i)

PI as the leader, (ii) P2 as the leader. 20. Obtain the pure, behavioral and mixed Stackelberg costs with P2 as the leader for the single-act game of Fig. 3.15. Does the game admit a behavioral Stackelberg strategy for the leader (P2)? 10,-5 -2,2 -1,1 1,-1 -1,1 0,2 -2,1 Figure 3.15: Single-act game of Problem 20. 21. Let  $\gamma_1^*$  be a mixed Stackelberg strategy for the leader (PI) in a two-person finite game, with the further property that  $\gamma_1^* \in T_1$ , i.e., it is a behavioral strategy. Prove that  $\gamma_1^*$  also constitutes a behavioral Stackelberg strategy for the leader. 22. Prove that the hierarchical equilibrium strategy  $\gamma_1^*$  introduced in Def. 3.37 always exists. Discuss the reasons why  $\gamma_1^*$  introduced in Def. 3.35, and  $\gamma_1^*$  and  $\gamma_2^*$  introduced in Def. 3.36, do not necessarily exist. 23. Determine hierarchical equilibria in pure or behavioral strategies of the single-act game of Fig. 3.4(b) when (i) P2 is the leader, PI and P3 are followers, (ii) PI and P2 are leaders, P3 is the follower, (iii) there are three levels of hierarchy with the ordering being P2, PI, P3. 24. In the four-person single-act game of Fig. 3.5, let P3 and P4 form the leaders' group, and P1 and P2 form the followers' group. Stipulating that the noncooperative Nash solution concept is adopted within each group and the Stackelberg solution concept between the two groups, determine the corresponding hierarchical equilibrium solution of the game in pure or behavioral strategies. 25. Determine the pure-strategy Nash equilibria of the three-person single-act game in extensive form depicted in Fig. 3.16, which incorporates a chance move. Also determine its (pure-strategy) Stackelberg equilibria with (i) PI as leader, P2 and P3 as followers (i.e., two levels of hierarchy), (ii) P2 as leader, PI as the first follower, P3 as the second follower (i.e., three levels of hierarchy).

158 T. BAŞAR AND G. J. OLSDER, ©. Figure 3.16: Three-person single-act game of Problem 25. 26. Let  $x_q$  be a discrete random variable taking the values  $x_0 = 1$  w.p.  $1/3$ ,  $0$  w.p.  $1/3$ ,  $1$  w.p.  $1/3$ . Consider the two-act game whose evolution is described by  $X_1 = X_0 + U_q + u_1$ ,  $x_2 = x_1 + u_2 + u_3$ , where  $U_q$  (respectively,  $u_i$ ) denotes the decision (action) variable of  $P_i$  at level 1 (respectively, level 2),  $i = 1, 2$ , and they take values out of the set  $\{-1, 0, 1\}$ . At level 1 none of the players make any observation (i.e., they do not acquire any information concerning the actual outcome of the chance mechanism ( $x_0$ )). At level 2, however, P2 still makes no observation, while P1 observes the sign of  $x_1$ . At this level, both players also recall their own past actions. Finally, let  $j_i = (x_1)^{+1} \bmod (X_2) + \text{sgn}(u_j) + \text{sgn}(u_i)$ ;  $i = 1, 2$ ; denote

the cost function of  $P_i$  which he desires to minimize, where  $\text{sgn}(x) = \begin{cases} 1, & \text{if } x > 0 \\ -1, & \text{if } x < 0 \\ 0, & \text{if } x = 0 \end{cases}$ . (i) Formulate this dynamic multi-act game in the standard extensive form and obtain its pure-strategy Nash equilibrium. (ii) Determine its pure-strategy Stackelberg equilibria with PI as leader.

TV-PERSON NONZERO-SUM FINITE GAMES 159 3.9 Notes Sections 3.2 and 3.3. The first formulation of nonzero-sum finite games dates back to the pioneering work of Von Neumann and Morgenstern (1947). But the non-cooperative equilibrium solution discussed in these two sections within the context of JV-person finite games in normal form was first introduced by Nash (1951) who also gave a proof of Thm. 3.2. For interesting discussions on the reasons and implications of occurrence of nonunique Nash equilibria, see Howard (1971). A nontechnical treatment of the "prisoner's dilemma" phenomenon can be found in the recent book by Poundstone (1992). Some other selected texts on nonzero-sum (as well as zero-sum) games are Binmore (1992), Fudenberg and Tirole (1991), Kreps (1990), Myerson (1991) and Szep and Forgo (1985). Section 3.4. For further discussion on the relationship between bimatrix games and nonlinear programming, see Section 7.4 of Parthasarathy and Raghavan (1971). Section 7.3 of the same reference also presents the Lemke-Howson algorithm which is further exemplified in Section 7.6 of that reference. Sections 3.5 and 3.7. A precise formulation for finite nonzero-sum dynamic games in extensive form was first given by Kuhn (1953), who also proved that every TV-person nonzero-sum game with perfect information (i.e., with singleton information sets) admits a pure-strategy Nash equilibrium solution (note that this is a special case of Prop. 3.11 given in subsection 3.5.4). He also showed that in every game in extensive form with perfect recall, every behavioral-strategy mixture (and hence a mixed strategy) for a player can be replaced by a realization-equivalent behavioral strategy, and hence equilibrium can be sought in the class of behavioral strategies. The notion of ladder-nested extensive forms and the recursive derivation of the Nash equilibrium solution of the delayed commitment type are introduced here for the first time. The "informational nonuniqueness" property of the Nash equilibrium solution in finite dynamic games is akin to a similar feature in infinite nonzero-sum games observed by Basar (1974, 1976a, 1977b); see also Section 6.3. A special version of this feature was observed earlier by Starr and Ho (1969a,b) who showed that, in nonzero-sum finite feedback games, the

feedback Nash equilibrium outcome could be different (for all players) from the open-loop Nash equilibrium outcome. The result of Prop. 3.10 was also proved in Dubey and Shubik (1981), and later included in the text by Shubik (1983). This latter reference calls the static (open-loop) version of an extensive form the most coarsened form, and its (open-loop) Nash equilibria information-insensitive equilibria, as they also constitute equilibria for all refined versions. The material of subsection 3.5.5 is based on Selten (1975) and Myerson (1978). Further refinements on Nash equilibrium can be found in Kreps and Wilson (1982), Kohlberg and Mertens (1986) and van Damme (1987, 1989). A detailed study of JV-player nonzero-sum games on finite graphs with cycles, viewed as "nonzero-sum extensions" of the zero-sum theory treated in subsection 2.7.1, can be found in Alpern (1991). The cost  $jV$ -tuples corresponding to an infinite path on the graph in such dynamic games are the long term averages of the local cost JV- tuples. Results have been obtained for the existence of Nash equilibria confined to certain classes of pure strategies. If a static game (such as the game of the prisoners' dilemma) is played repeatedly, one speaks of repeated games. In such games one gets to know about the past behavior of the other player(s) which may therefore influence the current decision to be made.

160 T. BAŞAR AND G. J. OLSDER With this role of information about the past actions of the players, the repeated game must be viewed as a dynamic one. It is possible to derive conditions for the existence of a Nash equilibrium for such a dynamic game, which, if viewed at the individual time steps, converges to a Pareto solution of the underlying static ("one shot") game as time evolves. The reason why a "sequence of Pareto-like solutions" can form a (stable) Nash equilibrium is the ability to perform threats, which, in its turn, is possible by applying strategies which include history. Many results exist in this direction. One can for instance consult Radner (1981, 1985), Mertens (1985, 1991) and the references therein. Sections 3.6 and 3.7. The Stackelberg solution in nonzero-sum static games was first introduced by H. Von Stackelberg (1934) within the context of economic competition and it was later extended to dynamic games primarily in the works of Chen and Cruz, Jr (1972) and Simaan and Cruz, Jr (1973a,b) who also introduced the concept of a feedback Stackelberg solution, but their formulation is restricted by the assumption that the follower's response is unique for every strategy of the leader. The present analysis is free from such a restriction. The feature of the

mixed-strategy Stackelberg solution exhibited in (Counter-) Example 3.14 is reported here for the first time. Its counterpart in the context of infinite games was first observed in Başar and Olsder (1980a); see also Section 4.4. For a discussion on a unified solution concept, from which Nash, Stackelberg, minimax solutions (and also cooperative solutions) follow as special cases, see Blaquiere (1976).

Chapter 4 Static Noncooperative Infinite Games 4.1 Introduction This chapter deals with static zero-sum and nonzero-sum infinite games, i.e., static noncooperative games wherein at least one of the players has at his disposal an infinite number of alternatives to choose from. In Section 4.2, the concept of equilibrium solutions is introduced, and several of its features are discussed. In Section 4.3, continuous-kernel games defined on closed and bounded subsets of finite dimensional spaces are treated, and general questions of the existence and uniqueness of Nash and saddle-point equilibrium solutions in such games are addressed; furthermore, some iterative algorithms are developed for their computation. Section 4.4 is devoted to the Stackelberg solution of continuous-kernel games, in both pure and mixed strategies; it presents some existence and uniqueness results, together with illustrative examples. Section 4.5 focuses on the consistent conjectural variations (CCV) equilibrium in continuous-kernel games, and introduces different levels of approximation to the CCV solution. Finally, Section 4.6 deals with quadratic games and economic applications. Here, the theory developed in the previous sections is specialized to games whose kernels are quadratic functions of the decision variables, and in that context the equilibrium strategies are obtained in closed form. These results are then applied to different price and quantity models of oligopoly. Section 4.7 deals with the Braess paradox, which shows that the Nash equilibrium solution concept could lead to rather counter-intuitive results in certain situations. 161

162 T. BAŞAR AND G. J. OLSDER 4.2 e Equilibrium Solutions Abiding by the notation and terminology of the previous chapters, we have  $N$  players,  $P_1, \dots, P_N$ , where  $P_i$  has the cost function  $J_i$  whose value depends not only on his action but also on the actions of some or all of the other players.  $P_i$ 's action is denoted by  $u_i$ , which belongs to his action space  $U_i$ .<sup>24</sup> For infinite games, the action space of at least one of the players has infinitely many elements. In Chapters 2 and 3, we have only dealt with finite games in normal and extensive

form, and in that context we have proven that noncooperative equilibrium solutions always exist in the extended space of mixed strategies (cf. Thm. 3.2). The following example now shows that this property does not necessarily hold for infinite games in normal form; it will then motivate the introduction of a weaker equilibrium solution concept — the so-called  $\epsilon$  equilibrium solution.

**Example 4.1** Consider the zero-sum matrix game  $A$  represented by

$$A = \begin{pmatrix} 2 & 0 & 1/2 \\ 1/2 & 1/3 & 2/3 \\ 1/4 & 3/4 & 1/5 \\ 4/5 & & \end{pmatrix}$$

Here,  $P_1$  has two pure strategies, whereas  $P_2$  has countably many alternatives. As in finite zero-sum games, let us denote the mixed strategies of  $P_1$  and  $P_2$  by  $H_i$ ,  $i = 1, 2$ ; and  $Z_j$ ,  $j = 1, 2, \dots$ , respectively; that is,  $y_i$  denotes the probability with which  $P_1$  chooses row  $i$ , and  $Z_j$  denotes the probability with which  $P_2$  chooses column  $j$ . Obviously we should have  $y_i > 0$ ,  $Z_j > 0$ ,  $y_1 + y_2 = 1$ . If we employ a natural extension of the graphical solution method described in Section 2.3, to determine the mixed security strategy of  $P_1$ , we arrive at Fig. 4.1. The striking feature here is that the upper envelope is not well defined. Suppose that the matrix  $A$  is truncated to  $A_k$ , where  $A_k$  comprises the first  $k$  columns of  $A$ . The mixed security strategy for  $P_1$  is then  $y^k = (k - 1)/(3k - 2)$ ;  $y^k = (2k - 1)/(3k - 2)$ , and the average security level of  $P_1$  is  $V_m(A_k) = 2(k - 1)/(3k - 2)$ . If we let  $k \rightarrow \infty$ , we obtain  $y^k \rightarrow 1/2$ ,  $y^k \rightarrow 1/2$ ,  $V_m(A_k) \rightarrow 1/2$ . Thus it is easily seen that  $(y^k = 1/2, y^k = 1/2, V_m(A_k) = 1/2)$  is a mixed security strategy for  $P_1$  in the matrix game  $A$ . The mixed security strategy for  $P_2$  in the truncated game  $A_k$  is  $(z^k = (k - 2)/(3k - 2), z^k = 2k/(3k - 2))$  and the average security level of  $P_2$  is  $V_m(A_k) = 2(k - 1)/(3k - 2)$ . In the limit as  $k \rightarrow \infty$ , we obtain  $(z^k = 1/3, z^k = 2/3)$  and  $V_m(A_k) = 1/2$ . But,  $P_2$  cannot guarantee for himself the average lower value  $V_m(A)$ , since he cannot choose column " $\infty$ ". We observe, however, that if he chooses column  $k$  with probability  $|$ ,  $k$  being sufficiently large, then he can secure for himself an average lower value arbitrarily close to  $V_m(A) = 1/2$ . In accordance with this observation, if  $P_2$  chooses  $k$  such that  $24$

In this chapter, we deal with static problems only, and in that context we do not distinguish between strategy space and action space, since they are identical.

STATIC NONCOOPERATIVE INFINITE GAMES 163

Figure 4.1: Graphical solution to the  $(2 \times \infty)$  matrix game of Example 4.1.  $V_m(A_k) > | - c$ , then the corresponding mixed security strategy is called an  $c$ -mixed security strategy, which is also an  $\epsilon$ -mixed saddle-point strategy for  $P_2$ . The reason for the observed disparity between finite and infinite games lies in the fact that the set of mixed strategies of  $P_2$  is not compact in the latter case. Consequently, an  $(m$

x oo) matrix game might not admit a saddle point in mixed strategies, in spite of the fact that the outcome is continuous with respect to mixed strategies of P2. D The preceding example has displayed the non-existence of a saddle point within the class of mixed strategies. It is even possible to construct examples of zero-sum (semi-infinite) matrix games which admit e saddle-point strategies within the class of pure strategies. This happens, for instance, in the degenerate matrix game A, where  $A = [0 \ || \ | \dots]$ . Similar features are also exhibited by nonzero-sum matrix games under the Nash solution concept. We now provide below a precise definition of an e equilibrium solution in iV- person games within the context of pure strategies. Extension to the class of mixed strategies is then immediate, and the last part of this section is devoted to a discussion on that extension, as well as to some results on existence of mixed e equilibrium solutions.

**Definition 4.1** For a given  $\epsilon > 0$ , an N-tuple  $\{u^1, \dots, u^N\}$ , with  $u^i \in U^i$ ,  $i \in N$ , is called a (pure)  $\epsilon$  Nash equilibrium solution for an N-person nonzero-sum infinite game if  $J^*(\langle f^1, \dots, \langle f^N \rangle) < \inf_{\langle f^1, \dots, \langle f^N \rangle} J_i(\langle f^1, \dots, \langle f^N \rangle) + \epsilon$ ,  $\forall i \in N$ . For  $\epsilon = 0$ , one simply speaks of "equilibrium" instead of " $\epsilon$  equilibrium" solution, in which case we denote the equilibrium strategy of  $P_i$  by  $u^i$ . For zero-sum two-person games the terminology is somewhat different. For  $N = 2$  and  $J_1 = -J_2 = J$ , we have the following definition.

164 T. BAŞAR AND G. J. OLSDER **Definition 4.2** For a given  $\epsilon > 0$ , the pair  $\{u^1, u^2\} \in U^1 \times U^2$  is called an  $\epsilon$  saddle point if  $J(u^1, u^2) - \epsilon < J(u^1, u^2) < J(u^1, u^2) + \epsilon$  for all  $\{v^1, v^2\} \in U^1 \times U^2$ . For  $\epsilon = 0$  one simply speaks of a "saddle point". By direct analogy with the treatment given in Section 2.2 for finite games, the lower value of a zero-sum two-person infinite game is defined by  $V_- = \sup \inf J(u^1, u^2)$ , which is also the security level of P2 (the maximizer). Furthermore, the upper value is defined by  $V_+ = \inf \sup J(u^1, u^2)$ ,  $U^1 \times U^2$  which is also the security level of P1. Since the strategy spaces are fixed, and in particular the structure of  $U^i$  does not depend on  $u^j$ ,  $i \neq j$ ,  $i, j = 1, 2$ , it can easily be shown, as in the case of Thm. 2.1 (hi), that  $V_- < V_+$ . If  $V_- = V_+$ , then  $V = V_- = V_+$  is called the value of the game. We should note that existence of the value of a game does not necessarily imply existence of (pure) equilibrium strategies; it, however, implies existence of an  $\epsilon$  saddle point—a property that is verified in Thm. 4.1 below, which proves, in addition, that the converse statement is also true.

**Theorem 4.1** A two-person zero-sum (infinite) game has a finite value if and only if, for every  $\epsilon > 0$ , an  $\epsilon$  saddle point exists.



Proof. First, suppose that the game has a finite value ( $V = V_- = V$ ). Then, given an  $\epsilon > 0$ , one can find  $u^1 \in U_1$  and  $u^2 \in U_2$  such that  $J(u^1, u^2) > V - \epsilon$ ,  $\forall u^1 \in U_1$ , (i)  $J(u^1, u^2) < V + \epsilon$ ,  $\forall u^2 \in U_2$ , (ii) which follow directly from the definitions of  $V_-$  and  $V$ , respectively. Now, since  $V_- = V = V$ , by adding  $\epsilon$  to both sides of the first inequality we obtain  $J(u^1, u^2) + \epsilon > V + \epsilon > J(u^1, u^2)$ ,  $\forall u^1 \in U_1$ , (iii) where the latter inequality follows from (ii) by letting  $u^2 \rightarrow u^2$ . Analogously, if we now add  $-\epsilon$  to both sides of (ii), and also make use of (i) with  $u^1 = u^1$ , we obtain  $J(u^1, u^2) - \epsilon < V - \epsilon < J(u^1, u^2)$ ,  $\forall u^2 \in U_2$ . (iv)

STATIC NONCOOPERATIVE INFINITE GAMES 165 If (iii) and (iv) are collected together, the result is the set of inequalities  $J(u^1, u^2) - \epsilon < J(u^1, u^2) < J(u^1, u^2) + \epsilon$ , (v) for all  $u^1 \in U_1$ ,  $u^2 \in U_2$ , which verifies the sufficiency part of the theorem, in view of Def. 4.2. Second, suppose that for every  $\epsilon > 0$ , an  $\epsilon$  saddle point exists, that is, a pair  $\{u^1 \in U_1, u^2 \in U_2\}$  can be found satisfying (v) for all  $u^1 \in U_1$ ,  $u^2 \in U_2$ . Let the middle term be denoted as  $J_\epsilon$ . We now show that the sequence  $\{J_{\epsilon_k}, J_{\epsilon_{k+1}}, \dots\}$ , with  $\epsilon_k > \epsilon_{k+1} > \dots > 0$  and  $\lim_{k \rightarrow \infty} \epsilon_k = 0$ , is Cauchy. Toward this end, let us first take  $\epsilon = \epsilon^*$  and  $\epsilon = \epsilon_j$ ,  $j > k$ , in subsequent order in (v) and add the resulting two inequalities to obtain  $J(u^1, u^2) - J(u^1, u^j) + J(u^1, u^2) - J(u^1, u^k) < 2(\epsilon_k + \epsilon_j) < 4\epsilon_k$ . Now, substituting first  $\{u^1 = u^1, u^2 = u^2\}$  and then  $\{u^1 = u^k, u^2 = u^2\}$  in the preceding inequality, we get  $-4\epsilon_k < J_k - J_j < 4\epsilon_k$  for any finite  $k$  and  $j$ , with  $j > k$ , which proves that  $\{J_k\}$  is indeed a Cauchy sequence. Hence, it has a limit in  $\mathbb{R}$ , which is the value of the game. It should be noted that, although the sequence  $\{J_{\epsilon_k}\}$  converges, the sequences  $\{u^1_k\}$  and  $\{u^2_k\}$  need not have limits. D Extension to mixed strategies All definitions, as well as the result of Thm. 4.1, of this section have been presented for pure strategies. We now outline possible extensions to mixed strategies. We define a mixed strategy for  $P_i$  as a probability distribution  $\mu_i$  on  $U_i$ , and denote the class of all such probability distributions by  $M_i$ . For the case in which  $U_i$  is a finite set, we have already seen in Chapters 2 and 3 that  $M_i$  is a simplex of an appropriate dimension. For an infinite  $U_i$  however, its description will be different. Let us consider the simplest case in which  $U_i$  is the unit interval  $[0, 1]$ . Then, elements of  $M_i$  are defined as mappings  $f_i : [0, 1] \rightarrow [0, 1]$  with the properties  $f_i(0) = 1 - M_i = 0$ ,  $f_i(u) > f_i(u')$  whenever  $u > u'$ , (4.1)  $f_i(u) = f_i(u')$  if  $u \sim u'$ , that is,  $M_i$  is the class of all probability distribution functions defined on  $[0, 1]$ . More generally, if  $U_i$  is a closed and bounded subset of a finite dimensional space, a mixed strategy for  $P_i$  can likewise be defined since every such  $U_i$  has the

cardinality of the continuum and can therefore be mapped into the unit For a definition of a Cauchy sequence, the reader is referred to Appendix A.

166 T. BAŞAR AND G. J. OLSDER interval in a one-to-one manner. However, such a definition of a mixed strategy- leads to difficulties in the construction of the average cost function, and therefore it is convenient to define  $f_x$  as a probability measure (see Ash (1972)) on  $[0, 1]$ , which is valid even if  $U^*$  is not closed and bounded. The average cost function of  $P_i$  is then defined as a mapping  $J^* : M_1 \times \dots \times M_N \rightarrow \mathbb{R}$  by  $J^*(u_1, \dots, u_N) = \int \sum_{i=1}^N J_i(u_1, \dots, u_N) d\mu_i(u_i)$  (4.2) whenever this integral exists, which has to be interpreted in the Lebesgue- Stieltjes sense (see Appendix B.3). With the average cost functions defined as in (4.2) the mixed  $\epsilon$  Nash equilibrium solution concept can be introduced as in Def. 4.1, by merely replacing  $J_1$  by  $J$ ,  $U_1$  by  $M^0$ ,  $u^0$  by  $f^0$  and  $u_f$  by  $u^*$ . Analogously, Def. 4.2 can be extended to mixed strategies, and Thm. 4.1 finds a direct counterpart in mixed strategies, with  $V_-$  and  $V$  replaced by  $Z_m = \sup \inf J$  and  $V_m = \inf \sup J$ , respectively, and with  $V$  replaced by  $V_m$ . We devote the remainder of this section to the mixed  $\epsilon$  Nash equilibrium solution concept within the context of semi-infinite bimatrix games. The two matrices, whose entries characterize the possible costs to be incurred to  $P_1$  and  $P_2$ , will be denoted by  $A = \{a_{ij}\}$  and  $B = \{b_{ij}\}$ , respectively, and both will have size  $(m \times \infty)$ . If  $m < \infty$ , then we speak of a semi-infinite bimatrix game, and if  $m = \infty$ , we simply have an infinite bimatrix game. The case when the matrices both have the size  $(\infty \times n)$  follows the same lines and will not be treated separately. A mixed strategy for  $P_1$ , in this case, is an  $m$ -tuple  $\{t_i, j/2, \dots, V_m\}$  satisfying  $\sum_{i=1}^m t_i = 1$  and  $t_i \geq 0$ . The class of such  $m$ -tuples is denoted by  $Y$ . The class  $Z$  is defined, similarly, as the class of all sequences  $\{z_i, Z_i, \dots\}$  satisfying  $\sum_{i=1}^{\infty} z_i = 1$  and  $Z_i \geq 0$ . Each element of  $Z$  defines a mixed strategy for  $P_2$ . Now, for each pair  $\{y \in Y, z \in Z\}$ , the average costs incurred to  $P_1$  and  $P_2$  are defined, respectively, by  $J_1(y, z) = \sum_{i=1}^m y_i \sum_{j=1}^{\infty} z_j a_{ij}$  and  $J_2(y, z) = \sum_{i=1}^{\infty} z_i \sum_{j=1}^m y_j b_{ij}$  assuming that the infinite series corresponding to these sums are absolutely convergent. If the entries of  $A$  and  $B$  are bounded, for example, then such a requirement is fulfilled for every  $y \in Y, z \in Z$ . The following theorem now verifies the existence of mixed  $\epsilon$  equilibrium strategies for the case  $m < \infty$ .

STATIC NONCOOPERATIVE INFINITE GAMES 167 Theorem 4.2 For each  $\epsilon > 0$ , the semi-infinite bimatrix game  $(A, B)$ , with  $A = \{a_{ij}\}_{i=1, j=1}^m, B =$

$\{b_{ij}\}^{i=1, \dots, m}$  and the entries  $a^i$  and  $b_{ij}$  being bounded, admits an equilibrium solution in mixed strategies. Proof. Let  $N$  denote the class of all positive integers. For each  $j \in N$ , define  $C_j$  as the  $j$ th column of  $B$ , i.e.,  $C_j = B_{e_j}$ . The set  $C = \{c_j, j \in N\}$  is a bounded subset of  $R^m$ . Let  $\epsilon > 0$  be given. Then, a finite subset  $C \subset C$  exists such that to each  $c \in C$  a  $c' \in C$  corresponds with the property  $\|c - c'\|_\infty < \epsilon$ , which means that the magnitude of each component of  $(c - c')$  is less than  $\epsilon$ . Without loss of generality, we assume that  $C = \{c^1, c^2, \dots, c^n\}$  with  $n < \infty$ . If necessary, the pure strategies of  $P_2$  can be rearranged so that  $C$  consists of the first  $n$  columns of  $B$ . Define  $A_n = \{a_{ii}\}_{i=1, \dots, n}$ ,  $B_n = \{b_{ij}\}_{i=1, \dots, n, j=1, \dots, n}$ . By Thm. 3.1, the "truncated" bimatrix game  $(A_n, B_n)$  has at least one equilibrium point in mixed strategies, to be denoted by  $(y, z)$ . We now prove that  $(y, z)$  is a mixed  $\epsilon$  equilibrium solution for  $(A, B)$ , where  $z$  is defined as  $(z_1, 0, 0, \dots)'$ . For all  $y \in Y$  we have  $y'Az = y'A_n z > y'A_n z = y'Az$ . For each  $r \in N$ , we choose a  $b(r) \in \{1, \dots, n\}$  such that  $\|c_r - c_{b(r)}\|_\infty < \epsilon$ . For each  $j \in \{1, \dots, n\}$ , let  $R(j) = \{r \in N : b(r) = j\}$ . Next, for each  $z \in Z$ , we define  $z' = (z'_1, \dots, z'_n)$  by  $z'_j = \sum_{r \in R(j)} z_r$ . Then  $z'_j > 0$  and  $\sum_{j=1}^n z'_j = 1$ . Subsequently,  $\|z - z'\|_\infty \leq \sum_{r=1}^m z_r \|c_r - c_{b(r)}\|_\infty < \epsilon$ . Hence  $y'Bz > y'B_n z - \epsilon > y'B_n z - \epsilon = y'Bz - \epsilon$  for all  $z \in Z$ . Inequalities (i) and (ii) verify that  $(y, z)$  constitutes a mixed  $\epsilon$  equilibrium solution for  $(A, B)$ .  $\square$  Corollary 4.1 Every bounded semi-infinite zero-sum matrix game has a value in the class of mixed strategies. Proof. This result is a direct consequence of Thm. 4.2, and constitutes a natural extension of Thm. 4.1 to mixed strategies.  $\square$  For  $(\infty \times \infty)$  bimatrix games such powerful results do not exist as shown in the following (counter-) example. Example 4.2 Each one of two players chooses a natural number, independently of the other. The player who has chosen the highest number wins and receives

168 T. BAŞAR AND G. J. OLSDER one unit from the other player. If both players choose the same number, then the outcome is a draw. The matrix corresponding to this zero-sum game is  $\begin{pmatrix} 0 & 1 & 1 & \dots \\ -1 & 0 & 1 & \dots \\ 1 & -1 & 0 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$  whose size is  $(\infty \times \infty)$ . For this game,  $V_- = -1$  and  $V_+ = +1$  and consequently the average value  $V_m$  does not exist, in spite of the fact that the matrix entries are bounded.  $\square$  4.3 Continuous-Kernel Games: Reaction Curves, and Existence and Uniqueness of Nash and Saddle-Point Equilibria This section deals with static games in which the number of alternatives available to each player is a continuum and the cost functionals are continuous. In particular, we shall consider the class of games for which a pure strategy of

each player can be represented as an element of a finite-dimensional space, and hence take  $U_i \subset \mathbb{R}^{m_i}$ , where  $m_i$  is an integer denoting the dimension of the decision vector of  $P_i$ . We will sometimes speak of "games on the square", which will refer to two-player games where  $U_i = [0,1]$ ,  $i = 1,2$ . We first discuss the role of reaction curves in the construction of pure-strategy Nash equilibria in continuous-kernel games, and present some classification of Nash equilibria based on the topological classifications of the reaction curves. Then, we state and prove some results on the existence of Nash equilibria, and as a special case the existence of saddle-point solutions in continuous-kernel games, in both pure- and mixed-strategy spaces. Finally, we discuss some computational algorithms motivated by the notion of stability of Nash equilibria. Reaction curves and classification of Nash equilibria Pure-strategy Nash equilibrium solution in infinite static games can be obtained as the common intersection point of the reaction curves of the players. The concept of a reaction set was already introduced in Chapter 3 (Def. 3.26) within the context of finite games, and its counterpart in infinite games describes a curve, or a family of curves, with some properties like continuity and differentiability, depending on the structure of the action sets and the cost functionals. We now make this notion precise for TV-person games.

STATIC NONCOOPERATIVE INFINITE GAMES 169 Definition 4.3 In an  $N$ -person nonzero-sum game, let the minimum of the cost function of  $P_i$ ,  $J_i(u_1, \dots, u_N)$ , with respect to  $u_i \in U_i$  be attained for each  $u_i \in U_i$ , where  $u_i = \{u_1, \dots, u_N\}$  and  $C_i = U_1 \times \dots \times U_N$ . Then, the set  $R_i(u_{-i}) \subset U_i$  defined by is called the optimal response or rational reaction set of  $P_i$ . If  $R_i(u_{-i})$  is a singleton for every  $u_{-i} \in U_{-i}$ , then it is called the reaction curve or reaction function of  $P_i$ , and is denoted by  $r_i(u_{-i})$ . The reaction sets and curves of  $P_i$ ,  $i = 2, \dots, N$  are similarly defined (simply by replacing the index 1 by  $i$ ).  $J = \text{constant} / J = \text{constant} - J$  Figure 4.2: Constant level curves for  $J_1$  and  $J_2$ , and the corresponding reaction curves  $\{r_1$  and  $r_2\}$  of  $P_1$  and  $P_2$ , respectively. To illustrate the role of reaction curves in the derivation of Nash equilibria, we have drawn in Fig. 4.2, the "constant level" or iso-cost curves corresponding to two cost functions  $J_1$  and  $J_2$  for a specific two-person game with  $U_1 = U_2 = \mathbb{R}$ . For fixed  $u_1$ , say  $u_1 = u_1$ , the best  $P_2$  can do is to minimize  $J_2$  along the line  $u_1 = u_1$ . Assuming that this minimization problem admits a unique solution, the said optimal response of  $P_2$  is determined, in the figure, as the point where the line  $u_1 = u_1$  is tangent to iso-cost curve  $J_2 = \text{constant}$ . For each different  $u_1$ , a

possibly different unique optimal response can thus be found for P2, and the collection of all these points forms the reaction curve of P2, indicated by  $\lambda$  in the figure. The reaction curve of P1 is similarly constructed: it is the collection of all points  $(u_1, u_2)$  where horizontal lines are tangent to the iso-cost curves of J1, and it is indicated by  $\mu$  in the figure. By definition, the Nash solution must lie on both reaction curves, and therefore, if these curves have only one point of intersection, as in the figure, the Nash solution exists and is unique. The configuration of the reaction curves is not always as depicted in Fig. 4.2. In Fig. 4.3, some other possibilities have been displayed. In all these cases it has been assumed that  $U_1 = U_2 = R$ . In Fig. 4.3a,  $\mu$  and  $\lambda$  are parallel straight lines and neither a Nash nor an  $\epsilon$  Nash solution exists. In Fig. 4.3b,  $\mu$  and  $\lambda$

170 T. BAŞAR AND G. J. OLSDER J ■ // '• // " // / / » ■ (a)  $u_1$   $u_2$  (0) (d) Figure 4.3: Possible configurations of the reaction curves. partly coincide and there exists a continuum of Nash solutions. In Fig. 4.3c, there is a finite number of Nash solutions. In Fig. 4.3d, one of the reaction curves is not connected and neither a Nash solution nor an  $\epsilon$  Nash solution exists. The nonzero-sum games associated with Figs. 4.3b and 4.3c both exhibit multiple Nash equilibria, in which case, as discussed earlier in Section 3.2, the Nash concept is weak as an equilibrium solution concept. Unless other criteria are imposed, there is no reason why players should prefer one particular equilibrium solution over the other(s). In fact, since the players make their decisions independently, it could so happen that the outcome of their joint (but noncooperative) choices is not an equilibrium point at all. Such nonequilibrium outcomes of unilateral equilibrium strategies are shown in Figs. 4.3b and 4.3c, where P1 picks a Nash strategy  $u_1$  and P2 chooses a Nash strategy  $u_2$ . In the case of nonunique Nash equilibrium solutions, we can classify them in a number of ways. One such classification is provided by the notion of "robustness", which is introduced below for (two-person) games on the square, which, however, easily extends to general TV-person games covered by Def. 4.3. Definition 4.4 Given two connected curves  $v^? — h_1(u_1)$  and  $u_1 — h_2(u_2)$  on the square, denote their weak  $\delta$ -neighborhoods by  $N_\delta$  and  $N_\delta$ , respectively.<sup>26</sup> Then, a point P of intersection of these two curves is said to be robust if, given  $\epsilon > 0$ , there exists a  $\delta > 0$  so that every ordered pair selected from  $N_{2\delta}$  and  $N_\delta$  has an intersection in an  $\epsilon$ -neighborhood of P. As specific examples, consider the two-person nonzero-sum games for which the reaction functions are depicted in Figs. 4.4a and 4.4b. In Fig. 4.4a, the <sup>26A</sup> weak  $\delta$ -neighborhood  $N_\delta$ , of:  $[0,1] \times [0,1]$  such that  $|(x,y) -$

$f(x) < 6, \forall x \in [0,1]. [0,1]$  is defined as the set of all maps  $f: [0,1] \rightarrow$

STATIC NONCOOPERATIVE INFINITE GAMES 171

Figure 4.4: Games on the square, illustrating robust Nash solutions. point  $P_i$ , constitutes a robust Nash solution, while point  $P^*$  is not robust. In Fig. 4.4b, however, all the Nash equilibrium solutions are robust. It should be noted that the number of robust intersections of two connected curves is actually odd, provided that possible points of intersections at the corners of the square are excluded. Hence, we can say that nonzero-sum two-person static games on the square, with well-defined reaction functions for both players, essentially admit an odd number of Nash equilibrium solutions. An exception to this rule is the case when the two reaction curves intersect only at the corners of the square. Yet another classification within the class of multiple Nash equilibrium solutions of a two-person nonzero-sum static game is provided by the notion of "stability" of the solution(s) of the fixed point equation. Given a Nash equilibrium solution, consider the following sequence of moves: (i) One of the players (say  $P_1$ ) deviates from his corresponding equilibrium strategy, (ii)  $P_2$  observes this and minimizes his cost function in view of the new strategy of  $P_1$ , (iii)  $P_1$  now optimally reacts to that (by minimizing his cost function), (iv)  $P_2$  optimally reacts to that optimum reaction, etc. Now, if this infinite sequence of moves converges to the original Nash equilibrium solution, and this being so regardless of the nature of the initial deviation of  $P_1$ , we say that the Nash equilibrium solution is stable. If convergence is valid only under small initial deviations, then we say that the Nash equilibrium solution is locally stable. Otherwise, the Nash solution is said to be unstable. A nonzero-sum game can of course admit more than one locally stable equilibrium solution, but a stable Nash equilibrium solution has to be unique. The reaction functions of two different nonzero-sum games on the square are depicted in Fig. 4.5. In the first one (Fig. 4.5a) the equilibrium solution is stable, whereas in the second one it is unstable. The notion of stability, as introduced above for two-person games, brings in a refinement to the concept of Nash equilibrium, which finds natural extensions to the  $N$ -player case. Essentially, we have to require that the equilibrium be "restorable" under any rational readjustment scheme when there is a deviation from it by any player. For  $N > 2$  this will depend on the specific scheme adopted, which brings us to the following formal definition of a stable Nash equilibrium.

172 T. BAŞAR AND G. J. OLSDER (a) (b) Figure 4.5: Illustration of (a) a stable and (b) an unstable Nash solution. Definition 4.5 A Nash equilibrium  $u_i, i \in N$ , is (globally,) stable with respect to an adjustment scheme  $S$  if it can be obtained as the limit of the iteration:  $u^r = \lim_{k \rightarrow \infty} u_i^k$  (4.3)  $u_i^{k+1} = \arg \min_{u_i} f_i(u_i, u_{-i}^k)$ ,  $u_i(0) \in IP$ ,  $t \in N$ , (4.4) where the superscript  $S^k$  indicates that the precise choice of  $u_{-i}$  depends on the readjustment scheme selected. One possibility for the scheme above is  $(S_k) = (f_i)$  (4.5) which corresponds to the situation where the players update (readjust) their actions simultaneously, in response to the most recently determined actions of the other players. Yet another possibility is  $(S_{fc}) = (u_i^{fc+1}, \dots, u_{-i}^{fc+1}, \hat{u}_i^{fc}, \dots, \hat{u}_{-i}^{fc})$  (4.6) where the players do the update in a predetermined (in this case numerical) order. A third possibility is  $(S_{\langle m \rangle}) = (u_i^{\langle m \rangle}, \dots, u_{-i}^{\langle m \rangle}, \hat{u}_i^{mU_k}, \dots, \hat{u}_{-i}^{mU_k})$ , (4.7) where  $m^k$  is an integer-valued random variable, satisfying the bounds  $\max(0, fc-d) < m^k < k + 1$ ,  $j^i, j \in N, i \in N$ , which corresponds to a situation where  $P_i$  receives action update information from  $P_j$  at random times, with the delay not exceeding  $d$  time units. Clearly, if the iteration of Def. 4.5 converges under any one of the readjustment schemes above (or any other readjustment scheme where a player receives update information from every other player infinitely often), then the Nash

STATIC NONCOOPERATIVE INFINITE GAMES 173 equilibrium solution is unique. Every unique Nash equilibrium, however, is not necessarily stable, nor is a Nash equilibrium that is stable with respect to a particular readjustment scheme necessarily stable with respect to some other scheme. Hence stability is generally given with some qualification (such as "stable with respect to scheme  $\langle S \rangle$ " or "with respect to a given class of schemes"), except when the number of players is two, in which case all schemes (with at most a finite delay in the transmission of update information) lead to the same condition of stability, as one then has the simplified recursions  $u_i^{k+\Delta} = \lim_{i \rightarrow \infty} (u_i(rki))$ ,  $A_i = 0, 1, \dots$ ;  $i = 1, 2$ , where  $r^{\Delta}, r_{2,i}, r_{\$,i}, \dots$  denote the time instants when  $P_i$  receives new action update information from  $P_j$ ,  $j \neq i$ ,  $i, j = 1, 2$ . Existence of Nash equilibria In view of the discussion in the previous subsection, the existence of pure-strategy Nash equilibria in  $N$ -person continuous-kernel games can be established by proving existence of well-defined reaction functions with a common point of intersection. The following theorem does precisely that, and provides a set of sufficient conditions under which  $N$ -person nonzero-sum games admit pure-strategy Nash equilibria. Theorem 4.3

For each  $i \in N$ , let  $U_i$  be a closed, bounded and convex subset of a finite-dimensional Euclidean space, and the cost functional  $J^* : U_1 \times \dots \times U_N \rightarrow \mathbb{R}$  be jointly continuous in all its arguments and strictly convex in  $u_i$  for every  $u_j \in U_j$ ,  $j \in N$ ,  $j \neq i$ . Then, the associated  $N$ -person nonzero-sum game admits a Nash equilibrium in pure strategies. Proof. Let us take  $i = 1$ . By strict convexity, there exists a unique mapping  $\lambda : U_{-1} \rightarrow U_1$  such that  $u_1 = \lambda(u_2, \dots, u_N)$  uniquely minimizes  $J^*(u_1, u_2, \dots, u_N)$  for any given  $(N-1)$ -tuple  $\{u_2, \dots, u_N\}$ . The mapping  $\lambda$  is actually the reaction function of  $P_1$  in this  $(N-1)$ -person game. Similarly, reaction functions  $\lambda_i$ ,  $i = 2, \dots, N$ , can be defined as unique mappings from  $U_{-i}$  into  $U_i$ . Using vector notation, these relations can be written in compact form as  $u = L(u)$ , where  $u = (u_1, \dots, u_N) \in U = U_1 \times \dots \times U_N$ , and  $L = (\lambda_1, \dots, \lambda_N)$ . It will be shown in the sequel that the individual reaction functions  $\lambda_i$  are continuous in their arguments, and hence  $L$  is a continuous mapping. Since  $L$  maps a closed and bounded subset  $U$  of a finite-dimensional space into the same subset, this then implies, by utilization of Brouwer's fixed point theorem (see Appendix C), that there exists  $u^*$  such that  $u^* = L(u^*)$  (that is,  $u^*$  is a fixed point of  $L$ ). Obviously, the individual components of  $u^*$  constitute a Nash equilibrium solution. To complete the proof of the theorem, what remains to be shown is the continuity of  $J^*$ . Let us take  $i = 1$  and assume that, to the contrary,  $\lambda_1$  is discontinuous at  $(u_2, \dots, u_N)$ . Further, let  $\lambda_1(u_2, \dots, u_N) = u_1$ . Then there exists

174 T. BAŞAR AND G. J. OLSDER a sequence of vectors  $\{u_j = (u_2, \dots, u_N)^T; j = 1, 2, \dots\}$  such that  $(u_2, \dots, u_N)^T$  is the limit of this sequence but  $u_j$  is not the limit of  $\lambda_1(u_2, \dots, u_N)$  as  $j \rightarrow \infty$ . Because of compactness of the action spaces, there exists a subsequence of  $\{u_j\}$ , say  $\{u_{j_k}\}$ , such that  $h(u_{j_k})$  converges to a limit  $u_1^* \neq u_1$ , and simultaneously the following inequality holds:  $J^*(u_{j_k}, u_{j_k}) < J^*(u_1, u_{j_k})$ . Now, by taking limits with respect to the subsequence of indices  $\{j_k\}$ , one obtains the inequality which, together with  $u_1^* \neq u_1$ , forms a contradiction to the initial hypothesis that  $u_1$  is the unique  $u_1$  which minimizes  $J^*(u_1, u_2, \dots, u_N)$ . Hence,  $\lambda_1$  is continuous. The continuity of  $\lambda_i$ ,  $i > 1$ , can analogously be proven.  $\square$  One of the conditions of Thm. 4.3 was the compactness of the action spaces  $U_i$ , which was needed in the proof of the theorem to ensure that for each  $u_{-i} \in U_{-i}$  there exists a  $u_i \in U_i$  that minimizes  $J^*$ , and hence that the reaction curve of each player is well defined. If the  $U_i$  are unbounded (such as  $U_i = \mathbb{R}^m$ ), then the same end result can be obtained provided that the cost function of the corresponding player becomes arbitrarily



large (positive) as its action variable becomes arbitrarily large in norm. Under such a condition, the player's action variable can be restricted to a closed and bounded set, without any loss of generality, and hence the theorem again applies. This reasoning then leads to the following corollary to Thm. 4.3.

**Corollary 4.2** For each  $i \in N$ , let  $U_i = R^{m_i}$ , the cost functional  $J^* : U_1 \times \dots \times U_N \rightarrow R$  be jointly continuous in all its arguments and strictly convex in  $u_i$  for every  $u^j \in [P, j \in N, j \neq i]$ . Furthermore, let  $J(u, u_i) \rightarrow \infty$  as  $\|u_i\| \rightarrow \infty$ ,  $\forall u_i \in U_i, i \in N$ . Then, the associated  $N$ -person nonzero-sum game admits a Nash equilibrium in pure strategies. The notion of stability introduced in the previous subsection (cf. Def. 4.5) can be used to sharpen the result above for two-person games, so as to establish unicity of equilibria under some additional conditions. Toward that end, let  $J_1$  and  $J_2$  be twice continuously differentiable on  $R^{m_1} \times R^{m_2}$ , with  $J_1$  strictly convex on  $R^{m_1}$  for each  $u_2 \in R^{m_2}$ , and  $J_2$  strictly convex on  $R^{m_2}$  for each  $u_1 \in R^{m_1}$ . Consider the parallel readjustment scheme whereby each player responds (optimally) to the previously selected action of the other player, which can be written as  $u_i = l_i(u_{-i})$ ;  $u = h(u)$ , where  $l_i$  and  $h$  are continuous (reaction) functions, uniquely solvable from  $\forall u_1 J_1(l_1(u_2), u_2) = 0, \forall u_2 J_2(u_1, l_2(u_1)) = 0$ .

**STATIC NONCOOPERATIVE INFINITE GAMES** 175 Hence, if  $(u_1^*, u_2^*)$  is a stable Nash equilibrium, then  $l_i^* = l_i(h^k(u))$ ,  $u = h^k(u)$ ,  $k \rightarrow \infty$  where it is generated from  $l(M^{-1})$ ,  $h$ ,  $h^k(u) = (h \circ l^k)(u)$ . Now introduce the matrix  $T(u, w, v) = [V^j l_i^j - V^j J_1^j, l_i] \cdot [V^i J_2(w, v) - V^i J_2^i(u, v)]$ , (4.8) where  $w$  is an arbitrary  $m_i$ -dimensional vector,  $v = h(w)$ , and  $u = l(v)$ . If the operator norm of  $T$  (equivalently, its spectral radius) is bounded above by  $a$ ,  $0 < a < 1$ , for all  $w \in R^{m_i}$ , it can be shown (see Li and Basar, 1987) that for some  $a$ ,  $a < a < 1$ ,  $\|l_i(h^k(u)) - l_i(h^{k+1}(u))\| < a \|u - h(u)\|$ ,  $k = 0, 1, 2, \dots$  which implies (by a contraction mapping argument) that  $u \rightarrow u$ , and hence  $u_2 = h(u) \rightarrow u_2^*$ —which proves existence of a stable Nash equilibrium, which by definition is also unique. Hence, we have the following. **Proposition 4.1** For a two-person nonzero-sum game, in addition to the hypotheses of Corollary 4-2, assume that  $J^*$  is twice continuously differentiable in  $u_i$  for each  $u^j \in R^{m_j}, j \neq i, i, j = 1, 2$ . Further assume that the matrix function (4-8) has operator norm strictly less than 1. Then the game admits a unique Nash equilibrium, which is also stable. **Remark 4.1** One class of games for which the last condition of Prop. 4.1 (that is, the one involving the operator norm of (4.8)) is satisfied is that with weakly coupled players, that is one with cost functions:  $J_z(u, u_2; e) =$

$J_i(u^j + \epsilon_j v^j, u^i)$ ,  $j \neq i$ ,  $i, j = 1, 2$ , where  $\epsilon$  is a sufficiently small scalar. More precisely, there exists an  $\epsilon_i > 0$  such that for all  $\epsilon \in (0, \epsilon_i)$ , this class of convex-kernel games admits a unique stable Nash equilibrium, which can be obtained as the limit point of the parallel update scheme introduced earlier (see Srikant and Başar, 1992).  $\square$  One important class of games not covered by Thm. 4.3 are those where the action spaces (or the constraint sets) of the players are not rectangular. As an example of such a scenario consider the two-person scalar game on the unit circle (instead of the unit square), where the scalar decision variables of the two players are coupled through the inequality  $(u_1)^2 + (u_2)^2 < 1$ . Note that in such a case,

176 T. BAŞAR AND G. J. OLSDER there does not exist a  $U_1$  and a  $U_2$  from which the actions of the players can be chosen independently, but rather a single set  $U = \{(u_1, u_2) : (u_1)^2 + (u_2)^2 < 1\}$ , to which the pair  $(u_1, u_2)$  belongs. Extending this to the  $N$ -person case, we have the coupled constraint set  $U \subset \mathbb{R}^m$ , where the  $N$ -tuple  $u = (u_1, u_2, \dots, u_N)$  belongs, with  $m = \sum_{j=1}^N n_j$ . The concept of Nash equilibrium is still well defined in this general framework, with the obvious modification that  $u \in U$  is a Nash equilibrium if  $J_i(u) \leq J_i(u, v^*)$ ,  $\forall v^* \in U_i$ ,  $i \in N$ , (4.9) where  $U_i(u_{-i})$  is a subset of  $\mathbb{R}^{n_i}$ , obtained by the projection  $U_i(u_{-i}) = \{v \in \mathbb{R}^{n_i} : (u_{-i}, v) \in U\}$ . Note that (4.9) can equivalently be written as  $J(u; u) \leq J(u; v)$ ,  $\forall v \in U$ , (4.10) where  $N J(u; v) = \sum_{i=1}^N J_i(u; v_i)$  a form that we will have occasion to use in the sequel. The following theorem now generalizes the result of Thm. 4.3 to such games.

**Theorem 4.4** Let  $U$  be a closed, bounded and convex subset of  $\mathbb{R}^m$ , and for each  $i \in N$  the cost functional  $J_i : U \rightarrow \mathbb{R}$  be continuous on  $U$  and convex in  $u_i$  for every  $u_{-i} \in W$ ,  $j \in N$ ,  $j \neq i$ . Then, the associated  $N$ -person nonzero-sum game admits a Nash equilibrium in pure strategies. *Proof* First note that under the hypotheses of the theorem the function  $J(u; v)$  defined prior to the statement of the theorem is continuous in  $u$  and  $v$  and is convex in  $v$  for every fixed  $u$ , with  $(u, v) \in U \times U$ . Introduce the reaction set for the game:  $T_u = \{v \in U : J(u; v) \leq J(u; w) \forall w \in U\}$ , which, by the continuity and convexity property of  $J$  cited above, is an upper semicontinuous mapping that maps each point  $u$  in the convex and compact set  $U$  into a closed convex subset of  $U$ . Then, by the Kakutani fixed point theorem (see Appendix C), there exists a point  $u^*$  such that  $u^* \in T_{u^*}$ , or equivalently such that it minimizes  $J(u^*; v)$  over  $v \in U$ . Such a point indeed  $\square$  This theorem is of course also valid for the special case in which the action constraint sets are rectangular, and for this class it relaxes

the assumption of strict convexity of the cost functions to simple convexity.

STATIC NONCOOPERATIVE INFINITE GAMES 177 constitutes a Nash equilibrium, because if it does not, then this would imply that for some  $i \in N$  there would be a  $u_1 \in T'(u_{-j})$  such that  $J_i(u^*_i, u_1) < J_i(u^*_i, u^*_i)$  which would in turn imply (by adding  $J_j(u^*_j)$  to both sides and summing over  $j \in N, j \neq i$ ) the strict inequality  $J(u^*; u) < J(u^*; u^*)$ ,  $u = (u^*_i, u_1)$ , contradicting the initial hypothesis that  $u^*$  minimizes  $J(u^*; v)$  over  $v \in U$ . D Since two-person zero-sum games are special types of (two-person) nonzero-sum games, the statements of Thms. 4.3 and 4.4 are equally valid for saddle-point solutions. The first one can even be sharpened to some extent. We first recall the ordered interchangeability property of multiple saddle points (viewed as a direct extension of Corollary 2.1), which we will need in the sequel. Property 4.1 For a given two-person zero-sum game with rectangular action sets  $U_1$  and  $U_2$  for the players, multiple saddle points satisfy the ordered interchangeability property; that is, if  $\{u^*_1, u^*_2\}$  and  $\{v_1, u_2\}$  are two saddle-point equilibrium solutions, so are the pairs  $\{u^*_1, u_2\}$  and  $\{v_1, u^*_2\}$ . Theorem 4.5 Consider a two-person zero-sum game on convex finite-dimensional action sets  $U_1 \times U_2$ , defined by the continuous kernel  $J(u_1, u_2)$ . Suppose that  $J(u_1, u_2)$  is strictly convex in  $u_1$  for each  $u_2 \in U_2$  and strictly concave in  $u_2$  for each  $u_1 \in U_1$ . Suppose that either (i)  $U_1$  and  $U_2$  are closed and bounded, or (ii)  $U_1 = \mathbb{R}^m, m = 1, 2$ , and  $J(u_1, u_2) \rightarrow \infty$  as  $\|u_1\| \rightarrow \infty$ , and  $J(u_1, u_2) \rightarrow -\infty$  as  $\|u_2\| \rightarrow \infty$ . Then, the game admits a unique pure-strategy saddle-point equilibrium. Proof. Existence of a saddle-point equilibrium is a direct consequence of Thm. 4.3 and Corollary 4.2. Furthermore, by strict convexity and concavity, there can be no saddle-point solutions outside the class of pure strategies. Hence only uniqueness within the class of pure strategies remains to be proven, which, however, follows readily from the ordered interchangeability property of multiple saddle points (cf Property 4.1). D The next theorem, which follows as a special case of Thm. 4.4, provides a generalization of the preceding theorem to coupled constraint sets, but in the absence of uniqueness of equilibria. Theorem 4.6 Let  $J(u_1, u_2)$  be a functional defined on the (not necessarily rectangular) convex and compact action set  $U$ . If  $J$  is continuous on  $U$ , concave in  $u_2$  for each  $u_1 \in U_1(u_2)$  and convex in  $u_1$  for each  $u_2 \in U_2(v_1)$ , then a saddle point exists in pure strategies, but it is not necessarily unique.

178 T. BAŞAR AND G. J. OLSDER We now focus attention on nonzero-sum

games whose cost functionals are continuous but not necessarily convex. For such games one cannot, in general, hope to obtain pure-strategy Nash equilibria; however, in the enlarged class of mixed strategies, the Nash equilibrium solution exists as it is stated in the following theorem. Theorem 4.7 An  $N$ -person nonzero-sum game in which the finite-dimensional action spaces  $U_i$  ( $i \in N$ ) are compact and the cost functionals  $J_i$  ( $i \in N$ ) are continuous on  $U_1 \times \dots \times U_N$  admits a Nash equilibrium in mixed strategies. Proof. A proof of this theorem can be found in Owen (1974). The underlying idea is to make the kernels  $J_i$  discrete so as to obtain an  $N$ -person finite matrix game that suitably approximates the original game in the sense that a mixed-strategy solution of the latter (which always exists by Thm. 3.1) is arbitrarily close to a mixed equilibrium solution of the former. Compactness of the action spaces ensures that a limit to the sequence of solutions obtained for approximating finite matrix games exists. For yet another proof of this result, see Glicksberg (1950). As a special case of Thm. 4.7 we now have the following. Corollary 4.3 Every continuous-kernel two-person zero-sum game with compact action (strategy) spaces admits a saddle point in mixed strategies. The question of the weakest conditions under which a mixed-strategy saddle point exists naturally arises. Corollary 4.3 says that continuity is a sufficient condition. But it is by no means necessary, and, in fact, it can be replaced by semi-continuity conditions (see Glicksberg, 1950). However, it is not possible to relax the semi-continuity conditions any further and still retain existence of a mixed-strategy saddle point within a general enough framework (see Sion and Wolfe, 1957). We conclude this section with an example of a zero-sum game whose cost functional is continuous but not convex-concave, and which admits a mixed saddle-point equilibrium. Figure 4.6: Reaction curves for the zero-sum game with kernel  $J = (u_1 - u_2)^2$ . Example 4.3 Consider the two-person zero-sum game on the square  $[0,2]^2$  characterized by the kernel  $J = (u_1 - u_2)^2$ . In Fig. 4.6, the two reaction curves

STATIC NONCOOPERATIVE INFINITE GAMES 179  $u_1$  and  $u_2$  have been drawn, which clearly do not intersect. It can readily be verified that the upper value  $V = 1$  and the lower value  $V_- = 0$ , i.e., a pure-strategy saddle point does not exist. In the extended class of mixed strategies, however, a candidate saddle-point solution directly follows from Fig. 4.6, which is  $u_1 = 1$  w.p.  $1/2$ ,  $u_2 = 2$  w.p.  $1/2$ . It can readily be verified that this pair of strategies indeed provides a mixed saddle-point

solution. □ 4.4 Stackelberg Solution of Continuous-Kernel Games This section is devoted to the Stackelberg solution of static nonzero-sum game problems in which the number of alternatives available to each player is not a finite set and the cost functions are described by continuous kernels. For the sake of simplicity and clarity in exposition, we shall deal primarily with two-person static games. A variety of possible extensions of the Stackelberg solution concept to TV-person static games with different levels of hierarchy will be briefly mentioned towards the end of the section, with the details left to the interested reader as exercises. Let  $u_1 \in U_1$  denote the action variable of  $P_1$ , where his action set  $U_1$  is assumed to be a subset of an appropriate metric space and, the cost function  $J_1$  of  $P_1$  be defined as a continuous function on the product space  $U_1 \times U_2$ . Then we can give the following general definition of a Stackelberg equilibrium solution, which is the counterpart of Def. 3.27 for infinite games. Definition 4.6 In a two-person game, with  $P_1$  as the leader, a strategy  $u_1 \in U_1$  is called a Stackelberg equilibrium strategy for the leader if  $J_1(u_1) = \sup_{u_2 \in R_2(u_1)} J_1(u_1, u_2) < \sup_{u_1 \in U_1} J_1(u_1, u_2)$  (4.11) for all  $u_1 \in U_1$ . Here,  $R_2(u_1)$  is the rational reaction set of the follower as introduced in Def. 4-3. Remark 4.2 If  $R_2(u_1)$  is a singleton for each  $u_1 \in U_1$ , in other words, if it is described completely by a reaction curve  $R_2 : U_1 \rightarrow U_2$ , then inequality (4.11) in the above definition can be replaced by  $J_1(u_1) = J_1(u_1, R_2(u_1)) < \sup_{u_1 \in U_1} J_1(u_1, R_2(u_1))$  (4-12) for all  $u_1 \in U_1$ . □

180 T. BAŞAR AND G. J. OLSDER If a Stackelberg equilibrium strategy exists for the leader, then the LHS of inequality (4.11) is known as the leader's Stackelberg cost, and is denoted by  $J_1^*$ . A more general definition for  $J_1^*$  is, in fact,  $J_1^* = \inf_{u_1 \in U_1} \sup_{u_2 \in R_2(u_1)} J_1(u_1, u_2)$  (4.13) which also covers the case when a Stackelberg equilibrium strategy does not exist. It follows from this definition that the Stackelberg cost of the leader is a well-defined quantity, and that there will always exist a sequence of strategies for the leader which will ensure him a cost value arbitrarily close to  $J_1^*$ . This observation brings us to the following definition of  $\epsilon$  Stackelberg strategies. Definition 4.7 Let  $\epsilon > 0$  be a given number. Then, a strategy  $u_1 \in U_1$  is called an  $\epsilon$  Stackelberg strategy for the leader ( $P_1$ ) if  $\sup_{u_2 \in R_2(u_1)} J_1(u_1, u_2) < J_1^* + \epsilon$ . The following two properties of  $\epsilon$  Stackelberg strategies now readily follow. Property 4.2 In a two-person game, let  $J_1^*$  be a finite number. Then, given an arbitrary  $\epsilon > 0$ , an  $\epsilon$  Stackelberg strategy for the leader necessarily exists. Property 4.3 Let  $\{u_1^i\}$  be a given sequence of  $\epsilon$  Stackelberg strategies in  $U_1$ , and with  $\epsilon_i > \epsilon_j$  for  $i < j$

and  $\lim_{t \rightarrow \infty} t_j = 0$ . Then, if there exists a convergent subsequence  $\{u_{i k}\}$  in  $U_1$  with its limit denoted as  $u_1$ , and further if  $\sup_{u_2 \in R_2(u_1)} J_1(u_1, u_2)$  is a continuous function of  $u_1$  in an open neighborhood of  $u_1 \in U_1$ ,  $u_1$  is a Stackelberg strategy for PI. The equilibrium strategy for the follower, in a Stackelberg game, would be any strategy that constitutes an optimal response to the one adopted (and announced) by the leader. Mathematically speaking, if  $u_1$  (respectively,  $u_1^*$ ) is adopted by the leader, then any  $u_2 \in R_2(u_1)$  (respectively,  $u_2 \in R_2(u_1^*)$ ) will be referred to as an optimal strategy for the follower that is in equilibrium with the Stackelberg (respectively, a Stackelberg) strategy of the leader. This pair is referred to as a Stackelberg (respectively, a Stackelberg) solution of the two-person game with PI as the leader (see Def. 3.28). The following theorem now provides a set of sufficient conditions for two-person nonzero-sum games to admit a Stackelberg equilibrium solution.

**Theorem 4.8** Let  $U_1$  and  $U_2$  be compact metric spaces, and  $J_i$  be continuous on  $U_1 \times U_2$ ,  $i = 1, 2$ . Further let there exist a finite family of continuous mappings  $f_i : U_1 \rightarrow U_2$ , indexed by a parameter  $i \in I = \{1, \dots, M\}$ , so that  $R_2(u_1) = \{u_2 \in U_2 : u_2 = f_i(u_1), i \in I\}$ . Then, the two-person nonzero-sum static game admits a Stackelberg equilibrium solution.

**STATIC NONCOOPERATIVE INFINITE GAMES 181** Proof. It follows from the hypothesis of the theorem that  $R_2(u_1)$ , as defined by (4.13), is finite. Hence, by Property 4.2, a sequence of Stackelberg strategies exists for the leader, and it admits a convergent subsequence whose limit lies in  $U_1$ , due to compactness of  $U_1$ . Now, since  $J_1(u_1, u_2)$  can be constructed from a finite family of continuous mappings (by hypothesis),  $\sup_{u_2 \in R_2(u_1)} J_1(u_1, u_2) = \max_{i \in I} J_1(u_1, f_i(u_1))$ , and the latter function is continuous on  $U_1$ . Then, the result follows from Property 4.3.

**Remark 4.3** The assumption of Thm. 4.8, concerning the structure of  $R_2(u_1)$ , imposes some severe restrictions on  $J_2$ ; but such an assumption is inevitable as the following example demonstrates. Take  $U_1 = U_2 = [0, 1]$ ,  $J_1 = -u_1 u_2$  and  $J_2 = (u_1 - u_2)u_2$ . Here,  $R_2(u_1)$  is determined by a mapping  $f(u_1)$  which is continuous on the half-open intervals  $[0, 1)$ ,  $(0, 1]$ , but is multivalued at  $u_1 = 1$ . The Stackelberg cost of the leader is clearly  $J_1 = -1/4$ , but a Stackelberg strategy does not exist because of the "infinitely multivalued" nature of  $f(u_1)$ .  $\square$

If  $R_2(u_1)$  is a singleton for every  $u_1 \in U_1$ , the hypothesis of Thm. 4.8 can definitely be made less restrictive. One such set of conditions is provided in the following corollary to Thm. 4.8 under which there exists a unique  $J_1$  which is continuous (by an argument similar to the one used in the proof of Thm. 4.3).

Corollary 4.4 Every two-person nonzero-sum continuous-kernel game on the square, for which  $J_2(u_1, \bullet)$  is strictly convex for all  $u_1 \in U_1$  and  $P_1$  acts as the leader, admits a Stackelberg equilibrium solution. It should be noted that the Stackelberg equilibrium solution for a two-person game exists under a set of sufficiency conditions which are much weaker than those required for existence of Nash equilibria (cf. Thm. 4.3). It should further be noted, however, that the statement of Thm. 4.8 does not also rule out the existence of a mixed-strategy Stackelberg solution which might provide the leader with a lower average cost. We have already observed occurrence of such a phenomenon within the context of matrix games, in Section 3.6, and we now investigate to what extent such a result could remain valid in continuous-kernel games. If mixed strategies are also allowed, then permissible strategies for  $P_i$  will be probability measures  $\mu_i$  on the space  $U_i$ . Let us denote the collection of all such probability measures for  $P_i$  by  $M_i$ . Then, the quantity replacing  $J_1$  will be the average cost function  $J_1(\mu_1, \mu_2) = \int \int J(u_1, u_2) d\mu_1(u_1) d\mu_2(u_2)$ , (4.14)  $J_m$   $\mu_1$  and the reaction set  $R_2$  will be replaced by  $J^*(\mu_1) = \{ \mu_2 \in M_2 : J(\mu_1, \mu_2) < J_2(\mu_1, \mu_2), \forall \mu_2 \in M_2 \}$ . (4.15)

182 T. BAŞAR AND G. J. OLSDER Hence, we have the following. Definition 4.8 In a two-person game with  $P_1$  as the leader, a mixed strategy  $\mu_1 \in M_1$  is called a mixed Stackelberg equilibrium strategy for the leader if  $J(\mu_1, \mu_2) < \sup_{\mu_2 \in M_2} J(\mu_1, \mu_2)$  for all  $\mu_1 \in M_1$ , where  $J_1^*$  is known as the average Stackelberg cost of the leader in mixed strategies. Proposition 4.2  $J_1^* < J_1^*$ . (4.16) Proof. Since  $M_1$  also includes all one-point measures, we have (by an abuse of notation)  $U_1 \subset M_1$ . Then, for each  $u_1 \in U_1$ , considered as an element of  $M_1$ ,  $\inf_{\mu_2 \in M_2} J(u_1, \mu_2) = \inf_{\mu_2 \in M_2} \int J(u_1, u_2) d\mu_2 = \inf_{u_2 \in U_2} J(u_1, u_2)$ , where the last equality follows since any infimizing sequence in  $M_2$  can be replaced by a subsequence of one-point measures. This implies that, for one point measures in  $M_1$ ,  $R_2(\mu_1)$  coincides with the set of all probability measures defined on  $U_2$ . Now, since  $M_1 \supset U_1$ ,  $J_1^* = \inf_{\mu_1 \in M_1} \sup_{\mu_2 \in M_2} J(\mu_1, \mu_2) < \inf_{\mu_1 \in U_1} \sup_{\mu_2 \in M_2} J(\mu_1, \mu_2)$ , and because of the cited relation between  $R_2^*(\mu_1)$  and  $R_2(\mu_1)$ , the last expression can be written as  $\inf_{\mu_1 \in U_1} \sup_{\mu_2 \in M_2} J(\mu_1, \mu_2) = J^*$ . We now show, by a counter-example, that, even under the hypothesis of Thm. 4.8, it is possible to have strict inequality in (4.16). Example 4.4 Consider a two-person continuous-kernel game with  $U_1 = U_2 = [0, 1]$ , and with cost functions  $J(u_1, u_2) = u_1^2 + u_2^2 - \epsilon u_1 u_2$ ;  $J_2 = (u_2 - u_1)^2$ , where  $\epsilon > 0$  is a sufficiently small parameter. The unique Stackelberg solution of this game, in pure strategies, is  $u_1^* = 0$ ,  $u_2^* = (w_1)^2$ , and the

Stackelberg cost for the leader is  $J_1^* = 0$ . We now show that the leader can actually do better by employing a mixed strategy.

STATIC NONCOOPERATIVE INFINITE GAMES 183 First note that the follower's unique reaction to a mixed strategy of the leader is  $u_2 = \sqrt{(u_1)^2}$  which, when substituted into  $J_1$ , yields the expression  $J_1 = eEKu_1)^2} + EtfWiEWm-EKu_1)^2]$ . Now, if the leader uses the uniform probability distribution on  $[0,1]$ , his average cost becomes  $\int_0^1 e-i i A" J= \dots + 2V3$  which clearly indicates that, for  $e$  sufficiently small,  $J_1' < 0 = J_1'$ .  $\square$  The preceding example has displayed the fact that even Stackelberg games with strictly convex cost functionals may fail to admit only pure-strategy solutions, and the mixed Stackelberg solution may in fact be preferable.<sup>28</sup> However, if we further restrict the cost structure to be quadratic, it can then be shown that only pure-strategy Stackelberg equilibria exist. Proposition 4.3 Consider the two-person nonzero-sum game with  $U_1 = Rm_i$ ,  $C/2=Rm^2$ , and where  $R \setminus t > 0$ ,  $R_{lu}$ ,  $R \setminus \wedge$ ,  $R_1 \wedge$  are matrices of appropriate dimensions, and  $r \setminus$ ,  $r^*$ - are vectors of appropriate dimensions. This "quadratic" game can only admit a pure-strategy Stackelberg solution, with either  $P_1$  or  $P_2$  as the leader. Proof. Without any loss of generality take  $P_1$  as the leader, and assume, to the contrary, that the game admits a mixed-strategy Stackelberg solution, and denote the leader's optimal mixed strategy by  $\{j,1$ . Furthermore, denote the expectation operation under  $\{i,1$  by  $E[-]$ . If the leader announces this mixed strategy, then the follower's reaction is unique and is given by  $v_f \wedge -R_1 \wedge R_1.Elu \wedge -R_1 \wedge r_l$ . By substituting this in  $J_1^* = \wedge [J_1]$ , we obtain  $J_1^* = \wedge [ix_1 \wedge ^1] + E \wedge YKElu_1] + Elu \wedge ^k + c$ , <sup>28</sup>In retrospect, this should not be surprising since for the special case of zero-sum games (without pure-strategy saddle points) we have already seen that the minimizer could further decrease his guaranteed expected cost by playing a mixed strategy; here, however, it holds even if  $J_1 \wedge = J_2$ .

184 T. BAŞAR AND G. J. OLSDER where  $K = \dots$   $\begin{matrix} "21 & "22 & "22-\wedge 22 & "21 \\ "12" * 22 & "21 & > k = 7 * 1 \\ "21 & "22 & r_2 \\ "12" & "22 & r_2 + "21 \\ -\wedge 22 & "22" & "22 & r_2 i \\ A & 1 & 9' t > 2 \\ - & 1 & n_1 & 7 - > 2 \\ - & 1 & 2 & 2' 7 - > 2 \\ - & 1 & 1 & c = 2r_2 \wedge 22 \\ \wedge 22 \wedge 22 & r \setminus & - & r \% Rfc \\ r \setminus \end{matrix}$ . Now, applying the Cauchy-Schwarz inequality (see Appendix B.4) on the first term of  $J_1^*$ , we further obtain the bound  $J_1^* > 1 - E \setminus uxyR \setminus xE \setminus ul \setminus + Elu \wedge ^k Eiu_1 \setminus + E[v \setminus ]^k + c$ , (i) which depends only on the mean value of  $u_1$ . Hence,  $J_1^* = \inf \{ -u \wedge R \wedge u_1 + ul'Kul + v \setminus \}^k + c, \setminus < J_1^*$ . This implies that enlargement of the strategy space of the leader, so as to include mixed strategies as well, does not



yield him any better performance. In fact, since  $E[u_1] > E[u_1]'$ , whenever the probability distribution is not one-point, it follows that the inequality in (i) is actually strict for the case of a proper mixed strategy. This then implies that, outside the class of pure strategies, there can be no Stackelberg solution. D Graphical display of Stackelberg solutions and the notion of relative leadership If  $U_1$  and  $U_2$  are one-dimensional spaces, and if, in addition,  $R_2\{u_1\}$  is singleton for each  $u_1 \in U_1$ , then the Stackelberg equilibrium solution can easily be visualized on a graph. In Fig. 4.7, iso-cost curves for  $J_1$  and  $J_2$  have been drawn, together with the reaction curves  $l_1$  and  $l_2$ . With  $PI$  as the leader, the Stackelberg solution will be situated on  $l_2$ , at the point where it is tangent to the Figure 4.7: Graphical construction of the Stackelberg solution in a two-person game—a "stalemate" solution.

STATIC NONCOOPERATIVE INFINITE GAMES 185 appropriate iso-cost curve of  $J_1$ . This point is designated as  $S_i$  in Fig. 4.7. The coordinates of  $S_i$  now correspond to the Stackelberg solution of this game with  $PI$  as the leader. The point of intersection of  $l_1$  and  $l_2$  (which is denoted by  $N$  in the figure) denotes the unique Nash equilibrium solution of this game. It should be noted that the Nash costs of both players are higher than their corresponding equilibrium costs in the Stackelberg game with  $PI$  as the leader. This is, though, only a specific example, since it would be possible to come up with situations in which the follower is worse off in the Stackelberg game than under the associated "Nash game". However, when the reaction set  $R_2(u_1)$  is a singleton for every  $u_1 \in U_1$ , the leader cannot do worse in the "Stackelberg game" than his best performance in the associated "Nash game" since he can, at the worst, play the strategy corresponding to the most favorable (from the leader's point of view) Nash equilibrium solution pair, as discussed earlier in Section 3.6 (see, in particular, Prop. 3.14). To summarize, we have the following. Proposition 4.4 For a two-person nonzero-sum game, let  $V^*$  denote the infimum of all the Nash equilibrium costs of  $PI$ . Then, if  $R_2\{u_1\}$  is a singleton for every  $u_1 \in U_1$ ,  $V^* > J_1^*$ . Hence, in two-person nonzero-sum games with unique follower responses, the leader never prefers to play the "Nash game" instead of the "Stackelberg game", whereas the follower could prefer to play the "Nash game", if such an option is open to him. In some games, even the option of who should be the leader and who should follow might be open to the players, and in such situations there is the question of whether it is profitable for either player to act as the leader rather than be the follower. For

the two-person game depicted in Fig. 4.7, for example, both players would prefer to act as the follower (in the figure, point  $5^*2$  characterizes the Stackelberg solution with P2 as the leader). There are other cases, however, in which both players prefer the leadership of only one of the players (see Fig. 4.8a, where leadership of P2 is more advantageous to both players) or in which each player wants to be the leader himself (Fig. 4.8b). (a) (b) Figure 4.8: Different types of Stackelberg equilibrium solutions, (a) Concurrent solution, (b) Nonconcurrent solution.

186 T. BAŞAR AND G. J. OLSDER When the players mutually benefit from the leadership of one of them, then this constitutes a stable situation since there is no reason for either player to deviate from the corresponding Stackelberg solution which was computed under mutual agreement—such a Stackelberg equilibrium solution is called a "concurrent" solution. If each player prefers to be the leader himself, then the Stackelberg solution is called "nonconcurrent". In such a situation each player will try to dictate to the other player his own Stackelberg strategy and thus force him to play the "Stackelberg game" under his leadership. In this case, the one who can process his data faster will certainly be the leader and announce his strategy first. However, if the slower player does not actually know that the other player can process his data faster than he does, and/or if there is a delay in the information exchange between the players, then he might tend to announce a Stackelberg strategy under his own leadership quite unaware of the announcement of the other player—which certainly results in a nonequilibrium situation. Finally, for the case in which neither player wants to be the leader (as in Fig. 4.7), the "Stackelberg game" is said to be in a "stalemate", since each player will wait for the other one to announce his strategy first.

4.5 Consistent Conjectural Variations Equilibrium

A third noncooperative solution concept for nonzero-sum games is the consistent conjectural variations (CCV) equilibrium, which we introduce here for the two- player case. This solution concept is in a sense a "double-sided Stackelberg equilibrium", where an equilibrium is sought in the class of reaction functions (or, more generally, reaction sets). For a precise mathematical formulation, let  $\Gamma_1 \times \Gamma_2$  be the class of all mappings  $(T_1, T_2), T^* : W \rightarrow X \setminus \{j\}^i, i, j = 1, 2$ , with the property that the composite maps  $T_1 \circ T_2$  and  $T_2 \circ T_1$  have unique fixed points. For a given pair  $(r_1, T_2) \in \Gamma_1 \times \Gamma_2$ , let  $\{v^*, u^r\} \in U_1 \times U_2$  denote the unique fixed points, satisfying  $u_1^* = T_1 \circ T_2(v^*)$ ,  $u_2^* = T_2 \circ T_1(u_1^*)$ . Introduce the unique pair  $\{u_1, u_2\} \in U_1 \times U_2$ , defined by  $u_x = \arg \min$

$(u_1, u_2)$ ;  $u_2 = \arg \min_{u_2 \in C} J_2(Tf(u_2), u_2)$  Let  $A_i; V = W - TtWC$ ,  
 $3 * i, i, J = 1, 2; g^?TV) = \arg \min. J_u \setminus u > + Alf(u^0/j ? i, i, j = 1, 2$ . Note that  
 $AT^u?)$  is the "differential" reaction of  $P_i$  under  $Tf$  to a deviation  $AT^?$  from  $u]C$   
 to  $u_3$ . Furthermore,  $g_i$  can be interpreted as a reaction function of (4.17)  
 (4.18) (4.19a) (4.19b)

STATIC NONCOOPERATIVE INFINITE GAMES 187  $P_i$  when  $P_j$  uses a  
 policy that is additively composed of a predetermined policy ( $u_j$ ) and a policy  
 determined by the policy choice of  $P_i$  under the reaction function  $ATf$ . Clearly,  
 $it_1$  introduced by (4.18) can also be obtained from  $V = gfhujc$ . (4.20) This  
 now brings us to the following definition. Definition 4.9 A pair of reaction  
 functions  $(T^1, T^2)$  &  $T_1 \times T_2$ ; along with their unique fixed point  $(u_1, u_2) \in U_1$   
 $\times U_2$ , constitute a consistent conjectural variations equilibrium (CCVE) if we  
 have the consistency of reaction functions:  $g_i^* = T_i$ ,  $g_j^* = T_j$  (4.21) and  
 consistency of policies  $u_1 = u_1 \setminus u_2 = u_2 \setminus$  (4.22) Remark 4.4 In the most general  
 case, it may be impossible to validate the consistency of the reaction functions,  
 though this may be possible to any "order" under appropriate smoothness  
 conditions on  $J_1$  and  $J_2$  (this will be further discussed in the sequel).

Consistency of policies, on the other hand, is more readily testable.  $\square$  To gain  
 more insight into CCVE, let us take  $U_1 = U_2 = R$ , and  $J_1$  and  $J_2$  to be  
 continuously differentiable and (jointly) strictly convex in their arguments.  
 Furthermore, let  $T_i \times T_j$  be chosen such that  $J_1(-, T_j(-))$  and  $J_2(T_i(-), -)$  are  
 strictly convex in their arguments, for every pair  $(T_i, T_j) \in T_1 \times T_2$ . To simplify  
 the notation, let  $u_1 = u$ ,  $u_2 = v$ . Then, two reaction functions  $(Tf, T2) \in T_1 \times T_2$   
 are in CCV equilibrium if (and only if)  $dJ_1(u, v) + dJ_2(u, v) \cdot g_j^* = T_i$ ,  
 $\{1\} \text{ ou } dv \text{ ou } 8^{\wedge} + 8^{\wedge}$  for  $\{ \} \text{ ov } \text{ ou } \text{ ov}$  Note that these are two coupled partial  
 differential equations which are, in general, difficult to solve. To gain some  
 further insight, we expand these around the CCV solution  $(u_c, v_c)$  and perform a  
 local analysis. First, to third order in  $v$ . and likewise, to third order in  $u$ :  
 $dT(u_c, v_c) + d^2T(u_c, v_c) \cdot dv + d^3T(u_c, v_c) \cdot dv^2 + \dots$

188 T. BAŞAR AND G. J. OLSDER Now, rewriting (i):  $dJ(T(v), v) +$   
 $dJ(Tf(v), v) + dT^*(v)) \in R \text{ du } dv \text{ du}$  and using the above expansion for  $Tf(u)$   
 around  $u = v_c$ , we arrive at the following: zeroth order  $dJ(u_c, v_c) + dJ(u_c, v_c)$   
 $dT(u_c) \text{ du } dv \text{ du} (0)$  first order (arguments at  $u = u_c, v = v_c$ )  $d^2J + dTf + d^2J$   
 $du^2 \text{ dv } dudv + dT \{ dv \text{ du } dn + d^2J + dT^{\wedge} + dJ_1 \text{ dv}^2 \text{ du } dv + d^i \cdot dJ_{\pm} = o(i) \text{ du}^2 \text{ dv}^2$   
 $\{ \text{ ' second order (arguments at } u = u_c, v = v_c) \} + d^3J \text{ du } dv^2 + \dots$



two-player games is that the Nash equilibrium  $(u^*, v^*)$ , where each player takes the policy of the other player as given (and fixed), is a zeroth-order CCVE, with  $Tf(v) = u^*$ ,  $Tg(u) = v^*$ . It is also a first-order CCVE under the set of restrictive conditions  $d^2J(u^*, v^*) \succeq d^2J(u, v)$  For  $(u^*, v^*)$  to be a second-order CCVE, also the set of conditions  $d^2J(u^*, v^*) \succeq d^2J(u, v)$  has to be satisfied. Since the latter two conditions are overly restrictive, we can say that generically the Nash equilibrium solution is a zeroth-order CCVE. Clearly this holds not only for the scalar game, and hence we have the following. Proposition 4.5 If a two-player nonzero-sum game admits a unique Nash equilibrium, say  $(u_1, u_2)$ , then this pair is a zeroth-order CCVE, where  $u_1$  and  $u_2$  are the constant reaction functions. The Nash equilibrium is not necessarily a higher-order CCVE.

190 T. BAŞAR AND G. J. OLSDER Remark 4.5 The notion of a CCVE can be extended from two- to many- player games, at least at the conceptual level. Computationally, however, the difficulties would be compounded here, as a player's response now depends on the policies or actions of more than one player (cf. Def. 4.3). □ 4.6 Quadratic Games with Applications in Microeconomics In this section, we obtain explicit expressions for the Nash, Stackelberg and consistent conjectural variations equilibrium solutions of static nonzero-sum games in which the cost functions of the players are quadratic in the decision variables—the so-called quadratic games. The action (strategy) spaces will be taken as Euclidean spaces of appropriate dimensions, but the results are also equally valid (under the right interpretation) if the strategy spaces are taken as infinite-dimensional Hilbert spaces. In that case, the Euclidean inner products will have to be replaced by the inner product of the underlying Hilbert space, and the positive-definiteness requirements on some of the matrices will have to be replaced by strong positive definiteness of the corresponding self-adjoint operators. This section will also include some discussion on iterative algorithms for the computation of Nash equilibria, and an illustration of the quadratic model and its Nash and Stackelberg solutions within the context of noncooperative economic equilibrium behavior of firms in oligopolistic markets. A general quadratic cost function for  $P_i$ , which is strictly convex in his action variable, can be written  $J_i = \frac{1}{2} u_i' R u_i + u_i' r + c_i$  (4-23)  $j=1, \dots, n$  where  $u_i \in U^i = R^{m_j}$  is the  $m_j$ -dimensional decision variable of  $P_j$ ,  $R^k$  is an  $(m_j \times m_j)$ -dimensional matrix with  $R^k \succ 0$ ,  $r_j$  is an  $m_j$ -dimensional vector and  $c_i$  is a constant. Without loss of generality,

we may assume that, for  $j \neq k$ ,  $R_{jk} = 0$ , since if this were not the case, the corresponding two quadratic terms could be written as  $u_j^T R_{jk} u_k + u_k^T R_{kj} u_j = u_j^T (R_{jk} + R_{kj}) u_k + u_k^T (R_{jk} + R_{kj}) u_j$  (4.24) and redefining  $R_{l,k}$  as  $(R_{lk} + R_{kl})/2$ , a symmetric matrix could be obtained. By an analogous argument, we may take  $R_{jj}$  to be symmetric, without any loss of generality. Quadratic cost functions are of particular interest in game theory, firstly because they constitute second-order approximation to other types of nonlinear cost functions, and secondly because they are analytically tractable, admitting, in general, closed-form equilibrium solutions which provide insight into the properties and features of the equilibrium solution concept under consideration.

191 To determine the Nash equilibrium solution in strictly convex quadratic games, we differentiate  $J_i$  with respect to  $u_i$  ( $i \in N$ ), set the resulting expressions equal to zero, and solve the set of equations thus obtained. This set of equations, which also provides a sufficient condition because of strict convexity, is  $R^* y + J_2^* = 0$  ( $i \in N$ ), which can be written in compact form as  $Ru = -r$ , (4.25) (4.26a) where  $A \in \mathbb{R}^{n \times n}$ ,  $r \in \mathbb{R}^n$ ,  $R_2 = \begin{bmatrix} R_{11} & \dots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{n1} & \dots & R_{nn} \end{bmatrix}$ ,  $r = (r_1, \dots, r_n)^T$ . (4.26b) (4.26c) (4.26d) This then leads to the following proposition. Proposition 4.6 The quadratic  $N$ -player nonzero-sum static game defined by the cost functions (4.23) and with  $R_{ii} > 0$ , admits a Nash equilibrium solution if, and only if, (4.26a) admits a solution, say  $u^*$ ; this Nash solution is unique if the matrix  $R$  defined by (4.26b) is invertible, in which case it is given by  $-R^{-1}r$ . (4.27) Remark 4.6 Since each player's cost function is strictly convex and continuous in his action variable, quadratic nonzero-sum games of the type discussed above cannot admit a Nash equilibrium solution in mixed strategies. Hence, in strictly convex quadratic games, the equilibrium analysis can be confined to the class of pure strategies. D We now investigate the stability properties of the unique Nash solution of quadratic games, where the notion of stability was introduced in Section 4.3. Assuming  $N = 2$ , and directly specializing Prop. 4.1 to the quadratic case, we arrive at the following iteration:  $u_i^{(k+1)} = C_i u_i^{(k)} + d_i + u_i^{(k+1)}$ ,  $i = 1, 2$ ,  $k = 0, 1, \dots$  (4.28)

192 T. BAŞAR AND G. J. OLSDER with an arbitrary starting choice  $u^0$ , where This iteration corresponds to the sequential (Gauss-Seidel) update

scheme where PI responds to the most recent past action of P2, whereas P2 responds to the current action of PI. The alternative to this is the parallel (Jacobi) update scheme where (4.28) is replaced by  $u^{k+1} = C^{-1}(W + d)$ ,  $u^{k+1} = C^{-1}(u^k + d)$ ,  $f_c = 0, 1, \dots$  (4.29) starting with arbitrary initial choices  $(u^0, u^0)$ . Then, the question of stability of the Nash solution (4.27), with  $N = 2$ , reduces to the question of stability of the fixed point of either (4.28) or (4.29). Note that, apart from a relabeling of indices, stability of these two iterations is equivalent to the stability of the single iteration:  $u^{k+1} = C^{-1}(C^{-1}u^k + d)$ . Since this is a linear difference equation, a necessary and sufficient condition for it to converge (to the actual Nash strategy of PI) is that the eigenvalues of the matrix  $C^{-1}C^{-1}$ , or equivalently those of  $C^{-1}C$ , should be in the unit circle, i.e.,  $\rho(C^{-1}C) = \rho(C^{-1}C) < 1$ , (4.30) where  $\rho(A)$  is the spectral radius of the matrix  $A$ . This spectral radius condition is precisely the one given in Prop. 4.1, which shows that the condition of Prop. 4.1 is tight for the quadratic case. Note that the condition of stability is considerably more stringent than the condition of existence of a unique Nash equilibrium, which is  $\det(J - C^{-1}C) \neq 0$ . (4.31) The question we address now is whether, in the framework of Gauss-Seidel or Jacobi iterations, this gap between (4.30) and (4.31) could be shrunk or even totally eliminated, by allowing players to incorporate memory into the iterations. While doing this, it would be desirable for the players to need to know as little as possible regarding the reaction functions of each other (note that no such information is necessary in the Gauss-Seidel or Jacobi iterations given above). To study this issue, consider the Gauss-Seidel iteration (4.28), but with a one-step memory for (only) PI. Then, the "relaxed" algorithm will be (using the simpler notation  $u^k = u^k$ ,  $u = u^k$ ):  $u^{k+1} = C^{-1}(C^{-1}u^k + d)$ ,  $u^{k+1} = C^{-1}(C^{-1}u^k + d)$ ,  $J'$  where  $A$  is a gain matrix, yet to be chosen. Substituting the second (for  $v^k$ ) into the first, we obtain the single iteration  $u^{k+1} = [C + A(I - C)]^{-1}u^k + (I - A)^{-1}C^{-1}d$ ,

STATIC NONCOOPERATIVE INFINITE GAMES 193 where  $c = c^{-1}C^{-1}$ . By choosing  $A = -C(I - C)^{-1}$ , (4.33) where the required inverse exists because of (4.31), we obtain an immediate convergence, assuming that the true value of  $C^{-1}$  is known to PI. If the true value of  $C^{-1}$  is not known, but a nominal value is given in a neighborhood of which the true value lies, the scheme (4.32) along with the choice (4.33) and using the nominal value, still leads to convergence (but not in a finite number of steps) provided that the neighborhood is sufficiently small (see Başar, 1987). Now, if the original scheme is instead the

parallel (Jacobi) scheme, then a one-step memory for PI will not be sufficient to obtain a finite-step convergence result as above. In this case we replace (4.32) by  $u_{k+1} = C_1 v_k + d_1 + B(u_{k-1} - C_1 v_k - d_1)$ ,  $v_{k+1} = C_2 u_{k+1} + d_2$ , where  $B$  is another gain matrix. Note that here PI uses, in the computation  $u_{k+1}$ , not  $v_k$  but rather  $u_{k-1}$ . Now, substituting for  $v_k$  from the second into the first equation of (4.34), we arrive at the iteration  $u_{k+1} = [C + B(J-C)]u_k + (J-B)[d_1 + C_1 d_2]$ , which again shows immediate convergence, with  $B$  chosen as  $B = -C(I-C)^{-1}$ . (4.35) Again, there is a certain neighborhood of nominal  $C_2$  or equivalently of the nominal  $C$ , where the iteration (4.34) is convergent. In general, however, the precise scheme according to which P2 responds to P1's policy choices may not be common information, and hence one would like to develop relaxation-type algorithms for PI which would converge to the true equilibrium solution regardless of what particular scheme P2 adopts (for example, Gauss-Seidel or Jacobi). Consider, for example, the scheme where P2's responses for different  $k$  are modeled by (see also (4.7))  $v_{k+1} = C_2 u_{k+1-k} + d_2$ , (4.36) where  $k > 0$  is an integer denoting the delay in the receipt of current policy information by P2 from PI. The choice  $k = 0$  for all  $k$ , would correspond to the Gauss-Seidel iteration, and the choice  $k = 1$  for all  $k$ , to the Jacobi iteration — assuming that  $u_{k+1}$  is still determined according to (4.28). An extreme case would be the totally asynchronous communication where  $\{k\}_{k \geq 0}$  could be any sequence of positive integers. Under the assumptions that PI communicates (4.34)

194 T. BAŞAR AND G. J. OLSDER new policy choices to P2 infinitely often, and he uses the simple ("nonrelaxed") iteration  $u_{k+1} = C_1 v_k + d_1$  (4.37) it is known from the work of Chazan and Miranker (1969) that such a scheme converges if, and only if,  $\rho(C) < 1$  (4.38) where  $C$  is the matrix derived from  $C_1$  by multiplying all its negative entries by -1. This condition can be improved upon, however, by incorporating relaxation terms in (4.37), as well as in (4.36), such as  $u_{k+1} = a C_1 v_k + (I - a) C_1 v_{k-1} + d_1$  . . .  $v_{k+1} = (3 C_2 u_{k+1} + (1 - 3) C_2 u_k + d_2$ , where  $a$  and  $3$  are some scalars. A sufficient condition for convergence of any asynchronously implemented version of (4.39) is (see Başar (1987) for details)  $\rho(A(a)) < 1$ , (4.40) where  $A(a) = \begin{pmatrix} C_1 & 0 \\ 0 & 3 C_2 \end{pmatrix} + \begin{pmatrix} -a & a \\ 0 & 0 \end{pmatrix}$ . Clearly, there are values of  $a$  for which (4.40) requires a less stringent condition (on  $C_1$  and  $C_2$ ) than (4.38). For example, if  $C_1$  and  $C_2$  are positive scalars, and  $a = 3$ , inequality (4.40) dictates  $C_1 C_2 < 4$  while (4.38) requires that  $C_1 C_2 < 1$ . From



a game-theoretic point of view, each of the iteration schemes discussed above corresponds to a game with a sufficiently large number of stages and with a particular mode of play among the players. Moreover, the objective of each player is to minimize a kind of an average long horizon cost, with costs at each stage contributing to this average cost. Viewing this problem overall as a multi-act nonzero-sum game, we observe that the behavior of each player at each stage of the game is rather "myopic", since at each stage the players minimize their cost functions only under past information, and quite in ignorance of the possibility of any future moves. If the possibility of future moves is also taken into account, then the rational behavior of each player at a particular stage could be quite different. For an illustration of this possibility the reader is referred to Example 4.5 which is given later in this section. Such myopic decision making could make sense, however, if the players have absolutely no idea as to how many stages the game comprises, in which case there is the possibility that at any stage a particular player could be the last one to act in the game. In such a situation, risk-aversing players would definitely adopt "myopic" behavior, minimizing their current cost functions under only the past information, whenever given the opportunity to act.

STATIC NONCOOPERATIVE INFINITE GAMES 195 Two-person zero-sum games Since zero-sum games are special types of two-person nonzero-sum games with  $J_1 = -J_2$  ( $P_1$  minimizing and  $P_2$  maximizing), in which case the Nash equilibrium solution concept coincides with the concept of saddle-point equilibrium, a special version of Prop. 4.6 will be valid for quadratic zero-sum games. To this end, we first note that the relation  $J_1 = -J_2$  imposes in (4.23) the restrictions  $P_1' = -P_2$ ,  $P_2 = -P_1$ ,  $P_1 = -P_2$ ,  $r_2 = -r_1$ ,  $r_1 = -r_2$ ,  $c_1 = -c_2$ . The matrix  $R$  defined by (4.26b) can be written as  $R = (R_1 \quad R_2) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & R_2 \\ -R_1 & 0 \end{pmatrix}$  which has to be nonsingular for existence of a saddle point. Since  $R$  can also be written as the sum of two matrices  $R = (R_1 \quad 0) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & R_2 \\ -R_1 & 0 \end{pmatrix}$  the first one being positive definite and the second one skew-symmetric, and since eigenvalues of the latter are always imaginary, it readily follows that  $R$  is a nonsingular matrix. Hence we arrive at the conclusion that every quadratic strictly convex-concave zero-sum game admits a unique saddle-point equilibrium in pure strategies—a result that also follows as a special case of Thm. 4.5. Corollary 4.5 The strictly convex-concave quadratic zero-sum game with cost function  $J = \frac{1}{2} u^T R u + u^T v + c$ ,  $R \succ 0$

$0, R_{22} > 0$ , admits a unique saddle-point equilibrium in pure strategies, which is given by  $u^1 = -[R_{11}^{-1}(R_{12} - R_{11}^{-1}R_{12}')]^{-1}h_1 + [I + R_{11}^{-1}(R_{12} - R_{11}^{-1}R_{12}')]^{-1}R_{11}^{-1}R_{12}'(R_{22} - R_{11}^{-1}R_{12} - R_{11}^{-1}R_{12}')^{-1}r$ . Remark 4.7 The positive-definiteness requirements on  $R_{11}$  and  $R_{22}$  in Corollary 4.5 are necessary and sufficient for the game kernel to be strictly convex- strictly concave, but this structure is clearly not necessary for the game to admit a saddle point. If the game is simply convex-concave (that is, if the matrices above are nonnegative definite, with a possibility of zero eigenvalues), then in view of Thm. 4.6 a saddle point will still exist provided that the upper and lower values are bounded.<sup>29</sup> If the quadratic game is not convex-concave, however, then either the upper or the lower value (or both) will be unbounded, implying that a saddle point will not exist.  $\square$  <sup>29</sup>For a convex-concave quadratic game, the upper value will not be bounded if, and only if, there exists a  $v \in R^{n_2}$  such that  $v'R_{22}v = 0$  while  $v'r > 0$ . A similar result applies to the lower value.

196 T. BAŞAR AND G. J. OLSDER Team problems Yet another special class of nonzero-sum games is the team problems in which the players (or equivalently, members of the team) share a common objective. Within the general framework, this corresponds to the case  $J_1 = J_2 = \dots = J_N = J$ , and the objective is to minimize this cost function over all  $u_i \in U_i, i = 1, \dots, N$ . The resulting solution  $N$ -tuple  $(u_1, u_2, \dots, u_N)$  is known as the team-optimal solution. The Nash solution, however, corresponds to a weaker solution concept in team problems, the so-called person-by-person (pbp) optimality. In a two-member team problem, for example, a pbp optimal solution  $(u_1, u_2')$  dictates satisfaction of the pair of inequalities  $J(u_1, u_2') < J(u_1, u_2'')$ ,  $\forall u_1 \in U_1, J(u_1, u_2') < J(u_1, u_2)$ ,  $\forall u_2 \in U_2$ , whereas a team-optimal solution  $(u_1, u_2)$  requires satisfaction of a single inequality  $J(u_1, u_2) < J(u_1, u_2'')$ ,  $\forall u_1 \in U_1, u_2 \in U_2$ . A team-optimal solution always implies pbp optimality, but not vice versa. Of course, if  $J$  is quadratic and strictly convex on the product space  $U_1 \times \dots \times U_N$ , then a unique pbp optimal solution exists, and it is also team-optimal.<sup>30</sup> However, for a cost function that is strictly convex only on individual spaces  $U_i$ , but not on the product space, this latter property may not be true. Consider, for example, the quadratic cost function  $J = (u_1)^2 + (u_2)^2 + |u_1|^2 + 2u_1u_2 + 3u_2^2$  which is strictly convex in  $u_1$  and  $u_2$ , separately. The matrix corresponding to  $R$  defined by (4.26b) is  $U_1 \times U_2$  which is clearly nonsingular. Hence a unique pbp optimal solution will exist. However, a team-optimal solution does not exist since the said matrix (which is also the Hessian of  $J$ )

has one positive and one negative eigenvalue. By cooperating along the direction of the eigenvector corresponding to the negative eigenvalue, the members of the team can make the value of  $J$  as small as possible. In particular, taking  $u_2 = -\Gamma^{-1} \alpha u_1$  and letting  $u_1 \rightarrow +\infty$ , drives  $J$  to  $-\infty$ . The Stackelberg solution We now elaborate on the Stackelberg solutions of quadratic games of type (4.23) but with  $N = 2$ , and  $P_I$  acting as the leader. We first note that since the

This result may fail to hold true for team problems with strictly convex but nondifferentiable kernels (see Problem 10, Section 4.8).

STATIC NONCOOPERATIVE INFINITE GAMES 197 quadratic cost function  $J_1$  is strictly convex in  $u_1$ , by Prop. 4.3 we can confine our investigation of an equilibrium solution to the class of pure strategies. Then, to every announced strategy  $u_1$  of  $P_I$ , the follower's unique response will be as given by (4.25) with  $N = 2$ ,  $i = 2$ :  $u_2 = -\{R_{12}\}^{-1}[B^1 u_1 + r_1]$ . (4.42) Now, to determine the Stackelberg strategy of the leader, we have to minimize  $J_1$  over  $U_1$  and subject to the constraint imposed by the reaction of the follower. Since the reaction curve gives  $u_2$  uniquely in terms of  $u_1$ , this constraint can best be handled by substitution of (4.42) in  $J_1$  and by minimization of the resulting functional (to be denoted by  $J_1$ ) over  $U_1$ . To this end, we first determine  $J_1: J_1(u_1) = \frac{1}{2} u_1^T R_{11} u_1 + [R_{12}^T u_1 + r_1]^T u_2 + \frac{1}{2} u_2^T R_{22} u_2 + u_2^T r_2 + C_1$ . For the minimum of  $J_1$  over  $U_1$  to be unique, we have to impose a strict convexity condition on  $J_1$ . Because of the quadratic structure of  $J_1$ , this condition amounts to having the coefficient matrix of the quadratic term in  $u_1$  positive definite, which is  $\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} - R_{21} R_{22}^{-1} R_{12}$ . Under this condition, the unique minimizing solution can be obtained by setting the gradient of  $J_1$  equal to zero, which yields  $W_1 = -[R_{11} + R_{21} R_{22}^{-1} R_{12}]^{-1} [R_{12} R_{22}^{-1} r_2 + r_1]$ . (4.44) Proposition 4.7 Under condition (4.43), the two-person version of the quadratic game (4.23) admits a unique Stackelberg strategy for the leader, which is given by (4.44). The follower's unique response is then given by (4.42). Remark 4.8 The reader can easily verify that a sufficient condition for condition (4.43) is strict convexity of  $J_1$  on the product space  $U_1 \times U_2$ .  $\square$  The consistent conjectural variations equilibrium Again working with the two-player version of the quadratic game defined by the cost functions (4.23), but with  $\Gamma^1 = \Gamma^2 = I$  for simplicity in notation (and without any loss of

generality), let us "conjecture" linear reaction functions:  $u_j = K_{m^{**}i} u_i + r_j, i, j = 1, 2, (4.45)$

198 T. BAŞAR AND G. J. OLSDER where  $K_t$  is an  $m, \times m, j$  dimensional matrix, and  $f_c, G \in \mathbb{R}^m$ . Substituting this into (4.21) we arrive at  $u_1 + R^2 u_2 + r^1 + K^2 \{R^2 u_1 + R^2 u_2 + r^1\} = 0$ , for  $u_1 = K^1 u_2 + f^1$  ( $u_2 + R^1 u_1 + r^1 + K^1 \{R^1 u_2 + R^2 u_1 + r^2\} + r^2 = 0$ , for  $u_2 = K^2 u_1 + k^2$ . These subsequently lead to the pair of equations  $\{K^1 + K^1 R^1 K^1 + R^1 + K^2 R^2\} u_2 + k^1 + r^1 + K^1 r^1 + K^1 h = 0$ ,  $(K^1 + K^1 R^1 K^1 + R^1 + K^1 R^1 h + r^1 + K^1 r^1 + K^1 R^1 k^2 = 0$ , which are required to hold for all  $\{u_1, u_2\} \in \mathbb{R}^m \times \mathbb{R}^m$ . A sufficient condition would be the existence of  $K^1, K^2, f^1, k^2$  that satisfy  $R^2, (4.46) R^2, (4.47) r^1 - K^2 r^1 (4.48) r^2 - K^1 r^2. (4.49)$  Note that what we have here is a pair of quadratic equations for  $K^1$  and  $K^2$ , and after  $K^1$  and  $K^2$  are determined the next two linear equations yield the corresponding choices for  $f^1$  and  $k^2$ . This shows that a CCVE equilibrium in the form of linear reaction functions is plausible in the quadratic case, but existence is by no means guaranteed. For some numerical results on the scalar ( $\mathbb{R}$ ) version of this problem we refer the reader to Başar, Turnovsky and D'Orey (1986), and Turnovsky, Başar and D'Orey (1988). Applications in microeconomics We now consider specific applications of the quadratic nonzero-sum game theory in microeconomics, in establishing economic equilibria in certain types of markets. The markets we have in mind are defined with respect to a specific good or a collection of goods, and an individual participant (player) in these markets is either a buyer or a seller. The traded goods are produced by the seller(s) and are consumed by the buyer(s). An important characteristic of such markets is the number of participants involved. Markets in which there are a large number of both buyers and sellers, and in which each participant is involved with only a negligible fraction of the transactions, are called (bilaterally) competitive. A key characteristic of such competitive markets is that no single participant can, by his own actions, have a noticeable effect on the prices at which the transactions take place. There is a market price, and no seller needs to sell for less, and no buyer needs to pay more. Our interest in the sequel will be on the so-called oligopoly markets, which comprise a few sellers and many competitive buyers. In the special case of  $(I + K^1 + K^1 R^1 + K^1 R^2) u_2 + k^1 + r^1 + K^1 r^1 + K^1 R^1 k^2 = (I + K^1 R^1 h = \{I + K^1 R^1\} k^2 =$

STATIC NONCOOPERATIVE INFINITE GAMES 199 two sellers this market

is known as a duopoly. In an oligopoly, the buyers cannot influence the price or quantities offered, and it is assumed that the collective behavior of the buyers is fixed and known. The sellers, however, have an influence via the price they ask and the output (production) they realize. Example 4.5 This example refers to what is known as a Cournot duopoly situation. There are two firms with identical products, which operate in a market in which the market demand function is known. The demand function relates the unit price of the product to the total quantity offered by the firms. It is then assumed that the whole production is sold at the prevailing price dictated by the demand function. Let  $q_i$  denote the production level of firm  $i$  ( $P_i$ ),  $i = 1, 2$ , and  $p$  denote the price of the commodity. Then, assuming a linear structure for the market demand curve, we have the relation  $p = a - \beta(q_1 + q_2)$ , where  $a$  and  $\beta$  are positive constants known to both firms. Furthermore, if we assign quadratic production costs to both firms, the profit functions are given by  $P_1 = q_1(p - h_1 q_1) - k_1(q_1)^2$ ,  $P_2 = q_2(p - h_2 q_2) - k_2(q_2)^2$ , for  $P_1$  and  $P_2$ , respectively, where  $h_1, h_2$  are nonnegative constants.  $P_1$  wants to maximize  $P_1$  and  $P_2$  wants to maximize  $P_2$ . Under the stipulation that the firms act rationally and that there is no explicit collusion, the Nash equilibrium concept seems to fit rather well within this framework. Formulated as a nonzero-sum game, the Nash equilibrium solution of this decision problem is  $q_1 = \frac{a - \beta h_2}{\beta(3 + 2k_2)} - \frac{a - \beta h_1}{\beta(3 + 2k_1)}$ ,  $q_2 = \frac{a - \beta h_1}{\beta(3 + 2k_2)} - \frac{a - \beta h_2}{\beta(3 + 2k_1)}$ ;  $Q = q_1 + q_2 = \frac{2(a - \beta(h_1 + h_2))}{\beta(3 + 2k_1 + 2k_2)}$ ;  $P = \frac{2(a - \beta(h_1 + h_2))^2}{\beta(3 + 2k_1 + 2k_2)}$  which has been obtained by solving equation (4.25). The corresponding (Nash equilibrium) level of price is  $p = \frac{a - \beta(3 + 2/3(k_1 + k_2))}{\beta}$ . To obtain the Stackelberg solution with, for instance,  $P_1$  as the leader, we first compute the reaction curve of  $P_2$ ,  $q_2 = \frac{a - \beta q_1}{\beta(2 + k_2)}$ , and then substitute this into  $P_1$  and maximize the resulting expression over  $q_1$ . The result is the following quantity level for the leader:  $q_1 = \frac{a - \beta h_1}{\beta(2 + k_1)}$ .

200 T. BAŞAR AND G. J. OLSDER The corresponding quantity level for the follower, and the price level are, respectively,  $q_2 = \frac{a - \beta h_2}{\beta(2 + k_2)}$ ,  $p = \frac{a - \beta(h_1 + h_2)}{\beta}$ . The CCV solution can be obtained by making direct use of (4.46)-(4.49). For simplicity in computation, let us consider the symmetric case where  $k_1 = k_2 = k$ , and let  $\beta = f_3/k$ . Then, (4.46)-(4.47) admit two pairs of solutions:  $\pm y/1 + 2P - \{1 + 0\}$   $K_1 = K_2 = -$  with the corresponding values of  $f_{c1}$  and  $k_2$  from (4.48)-(4.49) being a  $1 + P^y J_1 + 20$ . These two determine completely the optimum reaction functions (4.45) (in the

sense of CCVE). Now letting  $q_1 = K_1 q_2 + k_1$ ,  $q_2 = K_2 q_1 + k_2$  and solving for  $q_1$  and  $q_2$ , we obtain the following unique quantity levels for this symmetric duopoly under the CCVE:  $q_1 = q_2 = \frac{a(1+\beta)(1+2\beta T_v/1+2\beta)}{2/3(2/3 + 1)}$  □ Example 4.6 This is a somewhat more sophisticated, and extended version of Example 4.5. There will be  $N$  firms or players in the oligopoly problem to be defined, and it will be possible for the price of the product to vary from one firm to another. Let us first introduce the following notation:  $P_i$  — the profit for the  $i$ th firm ( $P_i$ ),  $p^*$  — the price per unit product charged by  $P_i$ ,  $p$  = average price =  $(\sum_{i=1}^N P_i) / N$ ,  $f_3$  = price sensitivity of overall demand,  $V$  = the price at which demand is zero,  $c$  = fixed costs per unit product of  $P_i$ ,  $d_i$  = demand for the product of the  $i$ th firm. Suppose that demand and price of the product of  $P_i$  are related through the relation  $d_i = [V - J(p - c)]$ , (4.50)

STATIC NONCOOPERATIVE INFINITE GAMES 201 where  $J > 0$  is a constant, or equivalently through  $p_i = (V - J(d_i - c))$  where we have replaced  $d_i$  with  $q_i$  since the quantity produced by each firm is assumed to be equal to the demand for the product of that firm under the price dictated by this relation. The profit of  $P_i$  can then be written as  $P_i = (p^* - c)d_i$  where  $d_i$  is to be interpreted as being equal to  $q_i$  if (4.51) is used. Note that in this example all firms enter symmetrically into the model. One could incorporate an asymmetry by including weighting factors in the determination of the average price  $p$  or by making  $c$  firm-dependent. This modified model would definitely be more realistic; however, the structural properties of equilibria are not that sensitive to such changes, and therefore we have avoided such a more complicated model for the sake of simplicity in exposition. For this example, we again seek the Nash equilibrium solution; but, from the model, it is not clear at the outset which set of variables should be taken as the decision (action) variables: the prices or the quantities? We shall elaborate here on both possibilities. In the so-called "price game" the Nash solution must satisfy the set of equations  $dP_i = 0, i \in N$ , (4.52) and in the so-called "quantity game" the first-order conditions for Nash equilibrium are  $dP_i = 0, i \in N$ . (4.53) In the price game it is assumed that the demands are determined via the prices according to (4.50), whereas in the quantity game it is assumed that the prices satisfy (4.51). In both cases the profit function of each firm is quadratic in the decision variables, and strictly concave in his own decision variable, and therefore it suffices to consider only the first order conditions. For the price game, the set of equations (4.52) becomes  $d \ln \{ \prod_{i=1}^N p_i \} + (p^* - c) \sum_{i=1}^N \{ p_i - p \} = 0, i \in N$ , and since

the  $p_i$ s appear symmetrically in this set of equations, we may substitute  $p = p_1 = p_2 = \dots = p_N$  and solve for  $p$ , to obtain  $p = \frac{2 + 1}{V-7} \dots$  (4.4)

202 T. BAŞAR AND G. J. OLSDER If this solution is substituted into (4.50), the following equilibrium values for the demands are obtained:  $d_i = \dots = d^* A d \sim \dots$  (4.55) Now, for the quantity game, the set of equations (4.53) can be rewritten as  $p_i(q_1, \dots, q_N) - c + q_i d p_i \{q_1, \dots, q_N\} = 0, i \in N$ . (4.56) In order to express  $p_i$  as a function of  $q_1, \dots, q_N$  only, we must solve (4.51), which as it stands is an implicit equation in the variables  $p_i$ . Because of symmetry, the solution should have the form  $p_i = a_1 + a_2 d_i + a_3 Y^j \wedge^t i$  where the coefficients  $a_j, k = 1, 2, 3$ , do not depend on  $i$ . Some analysis then leads to  $p = p_1 = \dots = p_N$  and  $q = q_1 = \dots = q_N$  and hence  $P = p^* = V^- \wedge, * \in N$ . (4.58) Now evaluating the quantity  $d p_i / d q_i$  from (4.57), we obtain  $d p^* N N + 7 d < t P N (1 + 7), i \in N$ . (4.59) Substitution of (4.58) and (4.59) into (4.56) yields the following unique equilibrium value for  $q_i$ : If this value is further substituted into (4.58), we obtain  $, \wedge^{m+7} + \wedge^{(1+7)} i \in N$  (4.61)  $P \sim \sim P 2 A r + 7(1 + A T) ' i$   $\wedge \wedge b l$ ) A comparison of (4.54) with (4.61) readily yields the conclusion that the Nash equilibria of price and quantity games are not the same. For the trivial game with only one firm, i.e.,  $N = 1$ , however, we obtain in both cases the same result, which is

STATIC NONCOOPERATIVE INFINITE GAMES 203 Figure 4.9: Graphical illustration of the effect of a coordinate transformation on Nash equilibria. Also, in the limit as  $N \rightarrow \infty$ , the corresponding limiting equilibrium values of  $p$  and  $q(d)$  in both games are again the same  $2 + 7$  For  $1 < N < \infty$ , however, the Nash equilibria differ, in spite of the fact that the economic background in both models is the same. Therefore, it is important to know a priori in this market as to which variable is going to be adopted as the decision variable. The price game with its equilibrium is sometimes named after Edgeworth (1925), whereas the quantity game is named after Cournot who first suggested it in the nineteenth century (Cournot, 1838). We now present a graphical explanation for occurrence of different Nash equilibria when different variables are taken as decision variables in the same model. Suppose that the iso-cost curves for a nonzero-sum game are as depicted in Fig. 4.2, which have been redrawn in Fig. 4.9, where the decision variables are denoted by  $p_1$  and  $p_2$ , and the reaction curves corresponding to these decision variables are

shown as  $I_1$  and  $I_2$ , respectively. If a transformation to a new set of decision variables  $(p_1, p_2)$  is made, then it should be apparent from the figure that we obtain different reaction curves  $m_j$  and  $m_i$  and hence a different Nash solution. As a specific example, in Fig. 4.9 we have chosen  $d_1 = p_1$ ,  $d_2 = p_1 + p_2$ , in which case  $m_i = U$ ; but since  $m^h$  Nash equilibrium becomes a variant of the coordinate transformation. It should be noted, however, that if the transformation is of the form  $d_1(p_1)$ ,  $d_2(p_2)$ , then the Nash equilibrium remains invariant.  $\square$

#### 4.7 Braess Paradox

In this section we present a nonzero-sum game for which the Nash equilibrium solution leads to a surprising phenomenon, called the "Braess paradox", after Dietrich Braess who was the first to publish about it (Braess, 1968). Braess' field of application concerned traffic behavior, but the Braess paradox has since

204 T. BAŞAR AND G. J. OLSDER been observed in other fields of application as well (Cohen and Horowitz, 1991). We will more or less follow here Braess' original paper.

**Figure 4.10: The routes of the Braess paradox.** Consider a network of roads as given in Fig. 4.10. All these roads allow only one-way traffic, as indicated by the arrows. The drivers on this network all want to go from point B to point E. They have to choose one of the three possible routes:

- route 1 consists of two segments: from B via C to E;
- route 2 consists of two segments: from B via D to E;
- route 3 consists of three segments: from B via C and D to E.

There are many cars and this leads to congestion on the roads. The time needed to drive along a segment depends on the intensity  $i$ , i.e., the number of cars per time unit that choose this segment. Measurements have shown that the time duration  $t$  needed to traverse each segment is as follows:

- along segment BC:  $t = 10 + i$ ;
- along segment BD:  $t = 50 + i$ ;
- along segment CE:  $t = 10 + i$ ;
- along segment CD:  $t = 10 + i$ ;
- along segment DE:  $t = 10 + i$ .

Suppose that  $X_j$  drivers choose route  $j$ ,  $j = 1, 2, 3$ . The total time needed to go from B to E is then

- along route 1:  $10(X_1 + X_3) + 50X_1 + 10(X_1 + X_3)$ ;
- along route 2:  $50X_2 + 10(X_2 + X_3)$ ;
- along route 3:  $10(X_1 + X_3) + 10 + 2:3 + 10(X_2 + 2:3)$ .

**STATIC NONCOOPERATIVE INFINITE GAMES** 205 Each driver will individually (and independently) choose that route for which the total driving time is the smallest. At equilibrium, this will lead to all three total driving times being equal. If we pick an arbitrary normalization, say  $X_1 + X_2 + X_3 = 6$  (this normalization leads automatically to integer-valued solutions), then these



driving times being equal leads to the unique Nash solution  $x_1 = x_2 = x_3 = 2$ , and the total driving time along each of the three routes equals 92 time units. Now reconsider the problem described above with one change: segment CD is not available (due to roadworks) and the drivers must now choose between route 1 and route 2 only. The unique Nash solution in this new game turns out to be  $x_1 = x_2 = 3$  ( $x_3 = 0$ ) and the total driving time along each of the two routes equals 83 time units. Apparently, reducing the number of possibilities has led to a uniformly better result (the driving time for each driver being now 83 rather than 92). Or, if we reason in the other direction, increasing capacity leads to a worse equilibrium—a rather counter-intuitive result (Cohen, 1988). If in the original problem, with three routes available, all drivers together would agree upon minimizing the maximum of the three total driving times, i.e., minimizing the maximum of the three expressions above for the total driving time, then the result turns out to be (the elementary calculation is left to the reader)  $x_1 = x_2 = 3$ ,  $x_3 = 0$ . Hence in this group optimum, equivalently Pareto optimum (see Chapter 1), route 3 is not used at all even though it is available. However, this solution is cheating-prone as each individual driver would be tempted to use route 3, since if he were the only one doing so, his total driving time would be reduced to approximately 70. This, of course, is not an equilibrium situation.

4.8 Problems

1. Solve the following zero-sum semi-infinite matrix games:  $P_2 \begin{pmatrix} 0 & 0 & 1 & -3/2 & 1 & -4/3 & 1 & -5/4 \end{pmatrix} P_2 \begin{pmatrix} 3/4 & 8/9 & 1 & 0 & 15/16 & 0 & 24/25 & 0 \end{pmatrix}$ . The kernel of a zero-sum game on the unit square is given by  $J(u_1, u_2) = (u_1)^3 - 2u_1u_2 + (u_2)^3$ . Determine the pure or mixed saddle-point strategies. (Hint: Note that  $J$  is strictly convex in  $u_1$ , and that for any  $u_1$  the maximum value of  $J$  with respect to  $u_2$  is attained either at  $u_2 = 0$  or at  $u_2 = 1$ .)

3. Consider the zero-sum continuous-kernel game on the unit square, defined by  $J(u_1, u_2) = (u_1 - u_2)^2 - a(u_2)^2$ , where  $a$  is a scalar parameter. Determine

206 T. BAŞAR AND G. J. OLSDER its pure or mixed saddle-point solution when (i)  $1 < a < 2$  and (ii)  $0 < a < 1$ . Consider the following zero-sum game, known as a duel. Two men, starting at  $t = 0$ , walk toward each other at such a speed that they meet each other at  $t = 1$  if nothing intervenes. Each one has a pistol with exactly one bullet, and may fire it at any time  $t \in [0, 1]$ . If one of them hits the other, the duel is immediately over, and the one who has fired successfully is declared the winner. If neither one fires successfully or if both fire simultaneously and successfully, then the duel becomes a stand-off. The probability of a hit after firing is inversely proportional with the distance

between the two duelists, and assuming a uniform distribution we can let this probability (of hitting the other player) to be  $t$  at time  $t \in [0,1]$ . This is a continuous-time version of the duel described in Section 2.8. Consider now the following two versions of the duel: (i) Silent duel. Assuming that the players do not know whether their opponents have fired (unless, of course, a particular player is hit), first show that an appropriate zero-sum game modeling this (silent) duel is one with the kernel  $-v\{u_2 - u_1 + u_2, u_1 < u_2, J(u_1, u_2) = \{ 0, u_1 = u_2, +u_1 u_2 - u_1 + u_2, u_1 > u_2, \text{ where } u_1 \text{ and } u_2 \text{ are the instants at which the pistols are fired by P1 and P2, respectively. Then, noting that this is a skew-symmetric game (a la Problem 9 of Section 2.9), show that its value in mixed strategies } (V_m) \text{ is zero, and the corresponding mixed saddle-point policies are: } (\sigma^*) = !^* (\sigma^*) = (i(9_{-}(1)2)j i < u < i \bullet$  (ii) Noisy duel. Now consider the case where the pistols are noisy, i.e., a player knows whether his opponent has already fired or not. Show that this version of the duel game can be described by the kernel  $1 - 2u \setminus u_1 < u_2, J(u_1, u_2) = \{ 0, u_1 = u_2, 2u_2 - 1, u_1 > u_2, \text{ and that it admits a saddle point in pure strategies, given by } \{u'_1 = \setminus, u'_2 = \setminus\}$ . 5. Consider the two-person nonzero-sum game with cost functions  $J_1(u, v) = [u(i - u - vai)T01] - \setminus J_2(u, v) = [v(x - v - ua2)T02] - \setminus$

STATIC NONCOOPERATIVE INFINITE GAMES 207 where  $u$  and  $v$  are the action variables of P1 and P2, respectively, belonging to the coupled constraint set  $U = \{u > 0.01, v > 0.01, u + vai < x, v + ua2 < x\}$ , and  $x, 01, 02$  are positive constants, and  $r, /3i, ^2$  are each positive and smaller than 1. (i) Using Thm. 4.4, show that the game admits a pure-strategy Nash equilibrium. (ii) Using Prop. 4.1, obtain a set of sufficient conditions on the six parameters characterizing the game, under which the pure-strategy Nash equilibrium is unique. Show that this condition is satisfied for the set of values:  $x = 1.259, a_1 = 1.1, a_2 = 1.2, ft = 0.8, /32 = 0.48, r = 0.2852$ . (iii) For the same set of parameter values given above, obtain the unique Nash equilibrium solution. Is it (globally) stable? If yes, can you obtain the solution using an iterative procedure? Partial answer for part (iii): The Nash equilibrium solution is  $(u^* = 0.3, v^* = 0.9)$ . 6. Prove the statement of Remark 4.1, and find the range of values of the scalar parameter  $e$  for which the Nash equilibrium solution of the following game (with weakly coupled players) can be obtained using the Gauss-Seidel readjustment scheme  $J_1(u_1, u_2) = u_1^2 + u_1 r + e \{u_1 u_1 + u_1 u_2\}, t_2 / 1 \setminus 2' / - i \setminus 2, 2' i \setminus 2' \setminus 2, l' \setminus 2 \setminus J(u, u) = u \setminus Qu + u \setminus r - e \{u u + u u\}$ , where 7. Consider the two-person Stackelberg game defined by the cost functions  $J_1 = a^2(1 + u_1 + u_2)^2 + (u_1)^2, t = 1, 2, \text{ where}$

$a^*$ ,  $i = 1, 2$ , are scalar parameters; the decision variables  $u_1$  and  $u_2$  are also scalars. For what values of  $(0, 1)$  is the game (i) concurrent, (ii) nonconcurrent, or (iii) stalemate? 8. Consider the two-person Stackelberg game, with PI as the leader, defined on the unit square, and with the cost functions  $J_1 = (1 - u_1)^2 + u_1$ ,  $J_2 = (u_1 - u_2)^2$ ,

208 T. BAŞAR AND G. J. OLSDER where  $\theta$  and  $k$  are scalars. For what values of  $\theta$  and  $k$  does a mixed Stackelberg solution for the leader yield a lower average cost for him than the pure Stackelberg solution? 9. In a three-person game with three levels of hierarchy, first PI announces his strategy and dictates it on both P2 and P3; subsequently P2 announces his strategy and dictates it on P3. The cost functions are given by  $J_1 = (x - 2)^2 + (u_1)^2$ ,  $J_2 = (x - 1)^2 + (u_2)^2$ ,  $J_3 = (x - 3)^2 + (u_3)^2$ , where  $x = u_1 + u_2 + u_3$  and  $u_i \in \mathbb{R}$ ,  $i = 1, 2, 3$ . Determine the hierarchical equilibrium solution of this game. Compare this with the Nash equilibrium solution of a nonzero-sum game having the same cost structure. In which one (the three-level hierarchy or the Nash game) does PI attain a better performance? 10. Consider a two-person game with  $J_1 = J_2 = J$ . Show that the cost function  $J : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined by  $J(u_1, u_2) = (1 - u_1)^2 + (u_2 - u_1)^2$  is strictly convex on  $\mathbb{R}^2$ . What are the reaction sets  $J_2(u_1)$  and  $J_1(u_2)$ ? Prove that infinitely many person-by-person (pbp) optimal solutions exist, but that the team solution is unique. Now prove that if  $J$  is quadratic and strictly convex, the reaction curves can have at most one point in common, and therefore the pbp optimal solution and the team solution are unique and identical. 11. The purpose of this problem is to show that the notion of stability of Nash equilibrium, as introduced in Def. 4.5, indeed depends on the particular scheme used when the number of players is three or more. Consider the scalar three-person game with the cost functions  $J_1 = (u_1)^2 + 2e u_1 u_2 + 3e u_1 u_3 - 2u_1$ ,  $J_2 = e u_2 u_1 + (u_2)^2 + e u_2 u_3 - 2u_2$ ,  $J_3 = e u_3 u_1 + e u_3 u_2 + (u_3)^2 - 2u_3$ , where  $e$  is a scalar parameter. The following two numerical schemes for the determination of Nash equilibrium correspond to (4.5) and (4.6), respectively: Scheme 1 :  $u_{j+1} = \arg \min_{u_j} J_1(u, u_j, u_j)$ ,  $u_{i+1} = \arg \min_{u_i} J_2(u, u_i, u_i)$ ,  $u_{k+1} = \arg \min_{u_k} J_3(u, u_k, u_k)$ ,  $U$

STATIC NONCOOPERATIVE INFINITE GAMES 209 Scheme 2 :  $u_{i+1} = \arg \min_{u_i} J_1(u, u^i, u^i)$ ,  $u_{k+1} = \arg \min_{u_k} J_2(u^{i+1}, u, u^k)$ ,  $u_{j+1} = \arg \min_{u_j} J_3(u^{i+1}, u^{k+1}, u^j)$  where, by a slight abuse of notation,  $u^k$  denotes the value of  $u_k$  at the  $k$ th step of iteration. (i) Prove that, for  $e = 0.36$ , if  $u^k$  is calculated according to scheme

1 (which is the Cournot-Jacobi iteration), the sequences  $\{u_k\}$ ,  $i = 1, 2, 3$ , do not converge, whereas they do converge if calculated according to scheme 2 (which is a particular Gauss-Seidel iteration). (ii) Obtain the complete range of values of  $\epsilon$  for which (a) scheme 1 is convergent, (b) scheme 2 is convergent. (iii) Repeat (ii) for the following variation of scheme 2: Scheme 3 :  $u_{i+1} = \arg \min_u J_1(u, u_k, u_k)$ ,  $u_{i+1} = \arg \min_u J_2(u, u, u)$ ,  $u_{i+1} = \arg \min_u J_3(u, u, u)$ . 12. Consider the following "relaxed" version of scheme 1 in Problem 11, where  $\alpha, \beta, \gamma$  are the relaxation parameters (à la Section 4.6) yet to be chosen: Relaxed Scheme 1 :  $u_{i+1} = \alpha u_i + (1 - \alpha) \arg \min_u J_1(u, u_i, u_i)$ ,  $u_{i+1} = \beta u_i + (1 - \beta) \arg \min_u J_2(u, u_i, u_i)$ ,  $u_{i+1} = \gamma u_i + (1 - \gamma) \arg \min_u J_3(u, u_i, u_i)$ . Let  $\epsilon = 0.36$ , and show that there exist nonzero values of  $\alpha, \beta, \gamma$  such that the sequences generated by the scheme above converge. 13. Relaxation algorithms of the type used in Problem 12 above can also be used in the computation of saddle-point solutions in zero-sum games, not only to assure convergence of a particular scheme, but also to improve the speed of convergence to equilibrium. To explore this possibility, consider the scalar zero-sum two-person game with reaction functions  $f_1, x, v, f_2, \sqrt{1 - 0.608m^2 u^2 + 0.1} w, \sqrt{1 - 0.608m^2 v^2}$  (i) First consider the Gauss-Seidel update scheme, written as:  $u_{k+1} = \arg \min_u J_1(u, v_k, v_k)$ ;  $v_{k+1} = \arg \min_v J_2(u_{k+1}, v, v)$ ;  $v_k$  arbitrarily chosen,  $k = 0, 1, \dots$

210 T. BAŞAR AND G. J. OLSDER Verify that with the starting choice picked as  $v_0 = 0$ , the sequence generated by this algorithm converges (to the nearest six figures) to the unique saddle-point solution ( $u^* = 0.9547778$ ,  $v^* = 0.94110625$ ) in approximately 120 iterations. (ii) Now consider the "relaxed" version:  $u_{k+1} = \alpha u_k + (1 - \alpha) \arg \min_u J_1(u, v_k, v_k)$ ,  $v_{k+1} = \beta v_k + (1 - \beta) \arg \min_v J_2(u_{k+1}, v, v)$ ;  $v_k$  arbitrarily chosen,  $k = 0, 1, \dots$  and show by numerical experimentation that by choosing the values of  $\alpha$  and  $\beta$  appropriately the speed of convergence can be improved significantly. Consider, for example, the values of  $\alpha = 1/3 = 0.67$ , for which the improvement is by a magnitude of 10. 14. Consider the "price" and "quantity" models of oligopoly as introduced in Example 4.6, but with the roles of the players such that P1 is the leader and P2, ..., PN are the followers. Further assume that the followers play according to the Nash solution concept among themselves. Obtain the solution of this "leader-followers" game, and study its asymptotic properties as the number of followers goes to infinity. 15. Consider a two-person nonzero-sum game with the cost functions defined by  $J_1 = (u_1)^2 + (u_2)^2$ ;  $J_2 = (u_1 - 1)^2 + (u_2 - 1)^2$ . (i) Determine the Nash solution, the consistent conjectural variations solution and the Stackelberg solution with P1

as the leader. (ii) Now assume that P1 has access to the action variable  $u_2$  of P2, and can announce a strategy in the form  $u_1 = au_2 + 0$ , where  $a$  and  $0$  are constants yet to be determined. Assuming that P2 is a rational player seeking to minimize his own cost function  $J_2$ , how should P1 choose the values for  $a$  and  $0$ ? Show that the resulting value of  $J_1$  is its global minimum, i.e., zero. Provide a graphical interpretation for this result.

4.9 Notes Section 4.2. The concept of equilibrium solution is as old as the concept of equilibrium solution itself. The results presented in this section are standard, and our proof of Thm. 4.2 follows closely the one given in Burger (1966) where additional results on equilibrium solutions can be found. Corollary 4.1, which follows from Thm. 4.2 in our case, was originally proven by Wald (1945) in a different way. More recently, the concept of equilibrium solution has received attention in the works of Tijs (1975) and Rupp (1977).

STATIC NONCOOPERATIVE INFINITE GAMES 211 Section 4.3. The notion of a reaction curve (also called reaction function) was first introduced in an economics context and goes as far back as 1838, to the original work of Cournot. There are several results in the literature, known as minimax theorems, which establish existence of the (saddle-point) value in pure strategies under different conditions on  $U_1$ ,  $U_2$  and  $J$ . Theorem 4.5 provides only one set of conditions which are sufficient for existence of a value. Some other minimax theorems have been given (in more general spaces) by Sion (1958), Fan (1953), Blackwell and Girshick (1954), Balakrishnan (1976) and Bensoussan (1971). The result on the existence of Nash equilibria in pure strategies (i.e., Thm. 4.3) is due to Nikaido and Isoda (1955), and extensions to mixed strategies were first considered by Ville (1938). The extension of Thm. 4.3 to coupled constraint sets (i.e., Thm. 4.4) is due to Rosen (1965), who also introduces a dynamic (differential equation-based) model for the computation of unique stable equilibrium. For extensions of these existence results to more general topological spaces, with applications to economics, see Aubin (1980). The discretization procedure alluded to in the proof of Thm. 4.7, in the text, suggests a possible method for obtaining mixed Nash equilibria of continuous-kernel games in an approximate way, if an explicit derivation is not possible. In this case, one has to solve for mixed Nash equilibria of only finite matrix games, with the size of these matrices determined by the degree of approximation desired.

Section 4.4. The Stackelberg solution concept was first introduced in economics through the work of H. Von Stackelberg in the early

1900s (cf. Von Stackelberg, 1934). But the initial formulation involved a definition in terms of reaction curves, with the generalization to nonunique reaction curves (or the so-called reaction sets) being relatively recent (cf. Başar and Selbuz, 1979a; Leitmann, 1978). A different version of Thm. 4.8, which accounts only for unique reaction curves, was given by Simaan and Cruz, Jr (1973b), but the more general version given here is due to Başar and Olsder (1980a). The latter reference is also the first source for extension to mixed strategies. A classification of different static Stackelberg solutions as given here is from Başar (1973), but the concept of "relative leadership" goes as far back as the original work of Von Stackelberg (1934). See also Ho and Olsder (1981). Section 4.5. The material in this section is based on Başar (1986a). Other, more restrictive definitions of CCVE can be found in Bresnahan (1981), and Kamien and Schwartz (1983). See also Olsder (1987) for another perspective on CCVE. Section 4.6. Quadratic games have attracted considerable attention in the literature partly because of analytic tractability (reaction functions are always affine). For further discussion on asynchronous implementation of Gauss-Seidel and Jacobi algorithms, and their generalized versions, for two as well as three-player games we refer the reader to Başar (1987). Some other selective references which deal with asynchronous algorithms in other contexts are Bertsekas and Tsitsiklis (1989a,b, 1991) and Tsitsiklis (1987, 1989). Two good sources for economic applications of (quadratic) games are Friedman (1977) and Case (1979). Example 4.5 is the so-called Cournot's quantity model and can be found in the original work of Cournot (1838). Several later publications have used this model in one way or another (see e.g., Intriligator (1971) and Başar and Ho (1974)). Example 4.6 is a simplified version of a model presented in Levitan and Shubik (1971), where the authors also provide a simulation study on some data obtained from the automobile industry in the USA.

212 T. BAŞAR AND G. J. OLSDER Section 4.7. The Braess paradox is named after Braess who apparently was the first to write about a paradoxal phenomenon in traffic control (Braess, 1968). This phenomenon can also occur in other fields of application such as electric networks and mechanical constructions (Cohen and Horowitz, 1991). Section 4.8. Problem 4 can be found in Karlin (1959). Extensions of the game of timing involve the cases of P1 having  $m$  bullets and P2 having  $n$  bullets, where  $m$  and  $n$  are positive integers. The general silent  $(m, n)$  case was solved by Restrepo (1957) and the

general noisy duel with  $m$  and  $n$  bullets was solved by Fox and Kimeldorf (1969). The case of P1 having  $m$  noisy bullets and P2 having  $n$  silent bullets remains unsolved for  $m > 1$ . Problem 5 is taken from Li and Basar (1987), which the reader can consult with motivation and more theoretical as well as numerical results. A theoretical development for weakly coupled games, in which context Problems 11 and 12 arise, can be found in Srikant and Başar (1992). Problem 15 (ii) exhibits a dynamic information structure and can therefore also be viewed as a dynamic game; for more details on such Stackelberg games the reader is referred to Chapter 7 of this book, and in particular to Section 7.4.

## Part II

This page intentionally left blank

Chapter 5 General Formulation of Infinite Dynamic Games 5.1 Introduction  
This chapter provides a general formulation and some background material for the class of infinite dynamic games to be studied in the remaining chapters of the book. In these games, the action sets of the players comprise an infinite number (in fact a continuum) of elements (alternatives), and the players gain some dynamic information throughout the decision process. Moreover, such games are defined either in discrete time, in which case there exists a finite (or at most a countable) number of levels of play, or in continuous time, which corresponds to a continuum of levels of play. Quite analogous to finite dynamic games, infinite dynamic games can also be formulated in extensive form which, however, does not lead to a (finite) tree structure, firstly because the action sets of the players are not finite, and secondly because of existence of a continuum of levels of play if the game is defined in continuous time. Instead, the extensive form of an infinite dynamic game involves a difference (in discrete time) or a differential (in continuous time) equation which describes the evolution of the underlying decision process. In other words, possible paths of action are not determined by following the branches of a tree structure (as in the finite case) but by the solutions of these functional equations which we also call "state equations". Furthermore, the information sets, in such an extensive formulation, are defined as subsets of some infinite sets on which some further structure is imposed (such as Borel subsets). A precise formulation which takes into account all these extensions as well as the

possibility of chance moves is provided in Section 5.2 for discrete-time problems and in Section 5.3 for game problems defined in continuous time. In a discrete-time dynamic game, every player acts only at discrete instants of time, whereas in the continuous-time formulation the players act throughout a time interval which is either fixed a priori or determined by the rules of

216 T. BAŞAR AND G. J. OLSDER the game and the actions of the players. For finite games, the case in which the duration of the game is finite, but not fixed a priori, has already been included in the extensive tree formulation of Chapters 2 and 3. These chapters, in fact, included examples of multi-act games wherein the number of times a player acts is explicitly dependent on the actions of some other player(s)— and these examples have indicated that such decision problems do not require a separate treatment, and they can be handled by making use of the theory developed for multi-act games with an a priori fixed number of levels. In infinite dynamic games, however, an a priori fixed upper bound on the duration may sometimes be lacking, in which case "termination" of the play becomes a more delicate issue, as already shown in Section 2.7. Therefore, we devote a portion of Section 5.3 to a discussion on this topic, which will be useful later in Chapter 8. Even though we shall not be dealing, in this book, with mixed and behavioral strategies in infinite dynamic games, we devote Section 5.4 to a brief discussion on this topic, mainly to familiarize the reader with the difficulties encountered in extending these notions from the finite to the infinite case, and to point out the direction along which such an extension would be possible. Section 5.5 deals with certain standard techniques of one-person single-goal optimization and, more specifically, with optimal control problems. Our objective in including such a section is twofold; first, to introduce the reader to our notation and terminology for the remaining portions of the book, and second, to summarize some of the tools for one-person optimization problems which will be heavily employed in subsequent chapters in the derivation of noncooperative equilibria of infinite dynamic games. Section 5.6 deals with the notion of "representations of strategies on trajectories", and the issue of "time consistency", both of which are of prime importance in infinite dynamic games, as it will become apparent in Chapters 6 and 7. Finally, Section 5.7 deals with so-called viscosity solutions. An important standard technique for continuous-time problems, as described in Section 5.5, is the solution of a partial differential equation. Among all possible solutions of this equation, the viscosity solution is a



particular one; it is quite often unique and has a clear interpretation in terms of a generalized optimal control problem to which some stochastic perturbations have been added.

### 5.2 Discrete-Time Infinite Dynamic Games

In order to motivate the formulation of a discrete-time infinite dynamic game in extensive form, we now first introduce an alternative to the tree model of finite dynamic games—the so-called loop model. To this end, consider an  $N$ -person finite dynamic game with strategy sets  $\{T_i; i \in N\}$  and with decision (action) vectors  $\{V_i \in U_i; i \in N\}$ , where the  $f$ th block component ( $u_{f,k}$ ) of  $u_f$  designates  $P_i$ 's action at the  $f$ th level (stage) of the game and no chance moves are allowed. Now, if such a finite dynamic game is defined in extensive form (cf. Def. 3.10), then for any given  $N$ -tuple of strategies  $\{\gamma_i^* \in T_i; i \in N\}$  the actions of the

GENERAL FORMULATION 217 players are completely determined by the relations  $u_i = \gamma_i(u_{-i})$ ,  $i \in N$ , (5.1) where  $r_i$  denotes the information set of  $P_i$ .

Moreover, since each  $r_i$  is completely determined by the actions of the players, there exists a point-to-set mapping  $g_i : U_1 \times U_2 \times \dots \times U_N \rightarrow N_i$  (i.e.  $N_i$ ) such that  $r_i = g_i(u_{-i})$  (i.e.  $N_i$ ) and when substituted into (5.1), yields the "loop" relation  $u_i = \gamma_i(g_i(u_{-i}))$ ,  $i \in N$ . (5.2) This clearly indicates that, for a given strategy  $N$ -tuple, the actions of the players are completely determined through the solution of a set of simultaneous equations which admits a unique solution under the requirements of Kuhn's extensive form (Kuhn, 1953). If the cost function of  $P_i$  is defined on the action spaces, as  $L_i : U_1 \times U_2 \times \dots \times U_N \rightarrow R$ , then substitution of the solution of (5.2) into  $V_i$  yields  $L_i(u_1, \dots, u_N)$  as the cost incurred to  $P_i$  under the strategy  $N$ -tuple  $\{\gamma_i^* \in T_i; i \in N\}$ . If the cost function is instead defined on the strategy spaces, as  $J_i^* : T_1 \times T_2 \times \dots \times T_N \rightarrow R$ , then we clearly have the relation  $J_i^*(\gamma_1^*, \dots, \gamma_N^*) = L_i(\gamma_1^*, \dots, \gamma_N^*)$  (i.e.  $N_i$ ). Now, if a finite dynamic game is described through a set of equations of the form (5.2), with  $\{\gamma_i; i \in N\}$  belonging to an appropriately specified class, together with an  $N$ -tuple of cost functions  $\{L_i; i \in N\}$ , then we call this a loop model for the dynamic game (Witsenhausen, 1971a, 1975b). The above discussion has already displayed the steps involved in going from a tree model to a loop model in finite deterministic dynamic games. Conversely, it is also possible to start with a loop model and derive a tree model out of that, provided that the functions  $\gamma_i : U_1 \times U_2 \times \dots \times U_N \rightarrow U_i$  (i.e.  $N_i$ ) are restricted to a suitable class satisfying conditions like causality, unique solvability of the loop equation (5.2), etc. In other words, if the sets  $P_i$  (i.e.  $N_i$ ) are chosen properly, then the set of equations  $u_i = \gamma_i(u_{-i})$ ,  $\gamma_i \in P_i$

$i \in \mathbb{N}$ , (5.3) together with the cost structure  $\{L_i; i \in \mathbb{N}\}$  leads to a tree structure that satisfies the requirements of Def. 3.10. If the finite dynamic game under consideration also incorporates a chance move, with possible alternatives for nature being  $u \in Q$ , then the loop equation (5.3) will accordingly have to be replaced by  $u_i = p_i(u_1, \dots, u_{A_i}, w)$ ,  $p_i \in P_i$ ,  $w \in W$ ,  $i \in \mathbb{N}$ , (5.4) where, by an abuse of notation, we again let  $P_i$  be the class of all permissible mappings  $p_i : U_1 \times U_2 \times \dots \times U_N \times W \rightarrow U_i$ . The solution of (5.4) will now be a function of  $w$ ; and when this is substituted into the cost function  $L_i : U_1 \times \dots \times U_N \times W \rightarrow \mathbb{R}$  and expectation is taken over the statistics of  $u_j$ , the resulting quantity determines the corresponding expected loss to  $P_i$ . Here  $i \in V^*$  denotes the set of all  $i$  as in Def. 3.11.

218 T. BAŞAR AND G. J. OLSDER Now, for infinite dynamic games defined in discrete time, the tree model of Def. 3.10 is not suitable since it cannot accommodate infinite action sets. However, the loop model defined by relations such as (5.3) does not impose such a restriction on the action sets; in other words, in the loop model we can take each  $U_i$ , to which the action variable  $u_i$  belongs, to be an infinite set. Hence, let us now start with a set of equations of the form (5.3) with  $\mathbb{N}$  taken as infinite sets. Furthermore let us decompose  $u_i$  into  $K$  blocks such that, considered as a column vector,  $v_i = [u_i^1, u_i^2, \dots, u_i^K]^T$ ,  $i \in \mathbb{N}$ , (5.5) where  $K$  is an integer denoting the maximum number of stages (levels) in the game,<sup>32</sup> and  $u_i^k$  denotes the action (decision or control) variable of  $P_i$  corresponding to his move during the  $k$ th stage of the game. In accordance with this decomposition, let us decompose each  $p_i \in P_i$  in its range space, so that (5.3) can equivalently be written as  $u_i = P_i(u_1, \dots, u_N)$ ,  $p_i \in P_i$ ,  $i \in \mathbb{N}$ ,  $k \in K$ . (5.6) Under the causality assumption which requires  $u_i^k$  to be a function of only the past actions of the players, (5.6) can equivalently be written as (by an abuse of notation)  $u_i^k = P_i^k(u_1^{1..k-1}, \dots, u_N^{1..k-1})$ ,  $p_i^k \in P_i^k$ ,  $i \in \mathbb{N}$ ,  $k \in K$ . (5.7) By following an analysis parallel to the one used in system theory in going from input-output relations for systems to state space models (Zadeh, 1969), we now assume that  $P_i^k$  is structured in such a way that there exist sets  $T_i^k$ ,  $Y_i^k$ ,  $X_i^k$  ( $i \in \mathbb{N}$ ,  $k \in K$ ), with the latter two being finite dimensional, and functions  $f_k : X_i^k \times U_i^1 \times \dots \times U_i^{k-1} \rightarrow X_i^k$ ,  $h_i^k : X_i^k \rightarrow Y_i^k$  ( $i \in \mathbb{N}$ ,  $k \in K$ ), and for each  $p_i^k \in P_i^k$  ( $i \in \mathbb{N}$ ,  $k \in K$ ) there exists a  $f_i^k : Y_i^k \rightarrow U_i^k$  ( $i \in \mathbb{N}$ ,  $k \in K$ ), such that (5.7) can equivalently be written as  $u_i^k = f_i^k(y_i^k, v_i^1, \dots, v_i^{k-1}; v_i^k = u_i^k)$ ,  $i \in \mathbb{N}$ ,  $k \in K$ , (5.8) where  $f_i^k$  is a sub collection of  $\{y_i^k, v_i^1, \dots, v_i^{k-1}; v_i^k = u_i^k\}$ , and  $z_i^k = (z_i^k, u_i^k)$ ,  $z_i^k \in Z_i^k$ , (5.9a)  $z_i^{k+1} =$

$f_k(x_k, u_1, \dots, u_n)$ ,  $x_{k+1} \in X$  (5.9b) for some  $X \in X$ .<sup>32</sup> In other words, no player can make more than  $K$  moves in the game, regardless of the strategies picked.

GENERAL FORMULATION 219 Adopting the system theory terminology, we call  $x_k$  the state of the game,  $y_k$  the (deterministic) observation of  $P_i$  and  $r_k$  the information available to  $P_i$ , all at stage  $k$ ; and we shall take relations (5.8), (5.9a)-(5.9b) as the starting point in our formulation of discrete-time infinite dynamic games. More precisely, we define an  $n$ -person discrete-time infinite dynamic game of prespecified fixed duration as follows. Definition 5.1 An  $n$ -person discrete-time deterministic infinite dynamic game<sup>33</sup> of prespecified fixed duration involves (i) An index set  $N = \{1, \dots, n\}$  called the players' set. (ii) An index set  $K = \{1, \dots, K\}$  denoting the stages of the game, where  $K$  is the maximum possible number of moves a player is allowed to make in the game. (iii) An infinite set  $X$  with some topological structure, called the state set (space) of the game, to which the state of the game ( $x_k$ ) belongs for all  $k \in K$ . (iv) An infinite set  $U_k$  with some topological structure, defined for each  $k \in K$  and  $i \in N$ , which is called the action (control) set of  $P_i$  at stage  $k$ . Its elements are the permissible actions  $u_i$  of  $P_i$  at stage  $k$ . (v) A function  $f_k: X \times U_1 \times \dots \times U_n \rightarrow X$ , defined for each  $k \in K$ , so that  $x_{k+1} = f_k(x_k, u_1, \dots, u_n)$ ,  $k \in K$  (i) for some  $x_1 \in X$  which is called the initial state of the game. This difference equation is called the state equation of the dynamic game, describing the evolution of the underlying decision process. (vi) A set  $Y_k$  with some topological structure, defined for each  $k \in K$  and  $i \in N$ , and called the observation set of  $P_i$  at stage  $k$ , to which the observation  $y_k$  of  $P_i$  belongs at stage  $k$ . (vii) A function  $h_k: X \rightarrow Y_k$ , defined for each  $k \in K$  and  $i \in N$ , so that  $y_k = h_k(x_k)$ ,  $k \in K$ ,  $i \in N$ , which is the state-measurement (-observation) equation of  $P_i$  concerning the value of  $x_k$ . (viii) A finite set  $r_k$ , defined for each  $k \in K$  and  $i \in N$  as a subset of  $\{y_k, v_1, \dots, v_n; V_1, \dots, V_n\}$  which determines the information gained and recalled by  $P_i$  at stage  $k$  of the game. Specification of  $r_k$  for all  $k \in K$  characterizes the information structure (pattern) of  $P_i$ , and the collection (over  $i \in N$ ) of these information structures is the information structure of the game.<sup>33</sup> Also known as an " $n$ -person deterministic multi-stage game".

220 T. BAŞAR AND G. J. OLSDER (ix) A set  $N_k$ , defined for each  $k \in K$  and  $i \in N$  as an appropriate subset of  $\{(Y^1 \times \dots \times Y^k) \times (U_1 \times \dots \times U_{k-1}) \times (U_1 \times \dots \times U_n) \times C^j\}$ , compatible with  $r_k$ .  $N_k$  is called the

information space of  $P_i$  at stage  $k$ , induced by his information  $r_{ik}$ . (x) A prespecified class  $T_k$  of mappings  $\gamma^k : N_k \rightarrow U_k$  which are the permissible strategies of  $P_i$  at stage  $k$ . The aggregate mapping  $\gamma^* = \{\gamma^1, \gamma^2, \dots, \gamma^K\}$  is a strategy for  $P_i$  in the game, and the class  $T$  of all such mappings  $\gamma^k$  so that  $\gamma^k \in T_k$ ,  $k \in \{1, 2, \dots, K\}$ , is the strategy set (space,) of  $P_i$ . (xi) A functional  $L_i : (X \times U^1 \times \dots \times U^N) \times \{X \times U^1 \times \dots \times U^N\} \times \dots \times (X \times U^1 \times \dots \times U^N) \rightarrow \mathbb{R}$  defined for each  $i \in N$ , and called the cost functional of  $P_i$  in the game of fixed duration. The preceding definition of a deterministic discrete-time infinite dynamic game is clearly not the most general one that could be given, first because the duration of the game need not be fixed, but be a variant of the players' strategies, and second because a "quantitative" measure might not exist to reflect the preferences of the players among different alternatives. In other words, it is possible to relax and/or modify the restrictions imposed by items (ii) and (xi) in Def. 5.1, and still retain the essential features of a dynamic game. A relaxation of the requirement of (ii) would involve introduction of a termination set  $A \subset X \times \{1, 2, \dots, K\}$ , in which case we say that the game terminates, for a given  $N$ -tuple of strategies, at stage  $k$ , if  $k$  is the smallest integer for which  $(x_j, k) \in A$ .<sup>34</sup> Such a more general formulation clearly also covers fixed duration game problems in which case  $A = X \times \{K\}$ , where  $K$  denotes the number of stages involved. A modification of (xi), on the other hand, might for instance involve a "qualitative" measure (instead of the "quantitative" measure induced by the cost functional), thus giving rise to the so-called qualitative games (as opposed to "quantitative games" covered by Def. 5.1). Any qualitative game (also called game of kind) can, however, be formulated as a quantitative game (also known as game of degree) by assigning a fixed cost of zero to paths and strategies leading to preferred states, and positive cost to the remaining paths and strategies. For example, in a two-player game, if  $P_1$  wishes to reach a certain subset  $A$  of the state set  $X$  after  $K$  stages and  $P_2$  wishes to avoid it, we can choose  $\gamma_i = f_0$  if  $x_{K+i} \in A$ ,  $L_2 = f_0$  if  $x_{K+i} \in A$ ,  $[1$  otherwise  $]$  —  $1$  otherwise and thus consider it as a zero-sum quantitative game. Now, returning to Def. 5.1, we note that it corresponds to an extensive form description of a dynamic game, since the evolution of the game, the information gains and exchanges of the players throughout the decision process, and the interactions of the players among themselves are explicitly displayed in such a <sup>34</sup>It is, of course, implicit here that  $x_k$  is determined by the given  $A$ -tuple of strategies, and the strategies are defined as in Def. 5.1, but by taking  $K$  sufficiently large.

GENERAL FORMULATION 221 formulation. It is, of course, also possible to give a normal form description of such a dynamic game, which in fact readily follows from Def. 5.1. More specifically, for each fixed initial state  $x^0$  and for each fixed  $N$ -tuple permissible strategies  $\{s_i^* \in F^i; i \in N\}$  the extensive form description leads to a unique set of vectors  $\{u^i = 1_k(rfk), x_k^+; i \in N, k \in K\}$  because of the causal nature of the information structure and because the state evolves according to a difference equation. Then, substitution of these quantities into  $U^i(i \in N)$  clearly leads to a unique  $N$ -tuple of numbers reflecting the corresponding costs to the players. This further implies existence of a composite mapping  $J = (J_1, \dots, J_N) : T_1 \times \dots \times T_N \rightarrow R$ , for each  $i \in N$ , which is also known as the cost functional of  $P_i$  ( $i \in N$ ). Hence, the permissible strategy spaces of the players (i.e.,  $T_1, \dots, T_N$ ) together with these cost functions ( $J_1, \dots, J_N$ ) constitute the normal form description of the dynamic game for each fixed initial state vector  $x^0$ . It should be noted that, under the normal form description, there is no essential difference between infinite discrete-time dynamic games and finite games (the complex structure of the former being disguised in the strategy spaces and the cost functionals), and this permits us to adopt all the noncooperative equilibrium solution concepts introduced in Chapters 2 and 3 directly in the present framework. In particular, the reader is referred to Defs 2.8, 3.12 and 3.27 which introduce the saddle-point, Nash and Stackelberg equilibrium solution concepts, respectively, which are equally valid for infinite dynamic games (in normal form). Furthermore the feedback Nash (cf. Def. 3.22) and feedback Stackelberg (cf. Def. 3.29) solution concepts are also applicable to discrete-time infinite dynamic games under the right type of interpretation, and these are discussed in Chapters 6 and 7, respectively. Before concluding our discussion on the ingredients of a discrete-time dynamic game as presented in Def. 5.1 we now finally classify possible information structures that will be encountered in the following chapters, and also introduce a specific class of cost functions—the so-called stage-additive cost functions. Definition 5.2 In an  $N$ -person discrete-time deterministic dynamic game of prespecified fixed duration, we say that  $P_i$ 's information structure is a(n) (i) open-loop (OL) pattern iff  $I_i = \{x^0\}, k \in K$ , (ii) closed-loop perfect state information (CLPS) pattern iff  $I_i = \{x^0, \dots, x_k\}, k \in K$ , (iii) closed-loop imperfect state information (CLIS) pattern iff  $I_i = \{y^0, \dots, y_k\}, k \in K$ , (iv) memoryless perfect state information (MPS) pattern iff  $I_i = \{x_i, x_k\}, k \in K$ , (v) feedback (perfect state) information (FB) pattern iff  $I_i = \{x_k\}, k \in K$ , (vi) feedback imperfect state information (FIS) pattern iff  $I_i = \{y_k\}, k \in K$ ,

222 T. BAŞAR AND G. J. OLSDER (vii) one-step delayed CLPS (1DCLPS pattern iff). =  $\{x^k, \dots, x^{k-1}\}$ , (viii) one-step delayed observation sharing (1DOS) pattern if  $f_k = \{y^k, \dots, y^{k-1}, y^k\}$ ,  $k \in K$ , where  $V_j = \{y^j, y^j, \dots, y^j\}$ .

**Definition 5.3** In an  $N$ -person discrete-time deterministic dynamic game of prespecified fixed duration (i.e.,  $K$  stages),  $P_i$ 's cost functional is said to be stage-additive if there exist  $g^k : X \times X \times U^k \times \dots \times U^1 \rightarrow \mathbb{R}$ , ( $k \in K$ ), so that  $K$   
 $L_i(u^1, \dots, u^N) = \sum_{k=0}^{K-1} g^k(x^{k+1}, u^k, \dots, u^0, x^k)$ , (5.10)  $f^k = i$  where  $v^k = (u^1, \dots, u^N)$ .  
 Furthermore, if  $L_i(v^k, \dots, u^N)$  depends only on  $x^{k+1}$  (the terminal state), then we call it a terminal cost functional. **Remark 5.1** It should be noted that every stage-additive cost functional can be converted into a terminal cost functional, by introducing an additional variable  $Z^k$  ( $k \in K$ ) through the recursive relation  $Z^{k+1} = Z^k + g^k\{f^k(x^k, u^k, \dots, u^1), u^k, \dots, u^1, x^k\}$ ,  $Z^0 = 0$ , and by adjoining  $Z^k$  to the state vector  $x^k$  as the last component. Denoting the new state vector as  $x^k = \{x^k, Z^k\}$ , the stage-additive cost functional (5.10) can then be written as  $L_i(u^1, \dots, u^N) = (Q, \dots, Q, 1) x^{K+1}$  which is a terminal cost functional. **D Games with chance moves:** Stochastic games Infinite discrete-time dynamic games which also incorporate chance moves (the so-called stochastic games) can be introduced by modifying Def. 5.1 so that we now have an additional player, called "nature", whose actions influence the evolution of the state of the game. These are, in fact, random actions which obey an a priori known probability law, and hence what replaces the state equation (5.9b) in stochastic games is a conditional probability distribution function of the state given the past actions of the players and the past values of the state. An equivalent way of saying this is that there exists a function  $F^k : X \times U^1 \times \dots \times U^N \times \Theta \rightarrow X$ , defined for each  $k \in K$ , so that  $Z^{k+1} = F^k\{x^k, u^k, \dots, u^N, \theta^k\}$ ,  $k \in K$ ,

**GENERAL FORMULATION 223** where  $\theta^k$  is the action variable of nature at stage  $k$ , taking values in  $\Theta$ ; the initial state  $x^0$  is also a random variable and the joint probability distribution function of  $\{x^0, \theta^0, \dots, \theta^k\}$  is known. A precise formulation of a stochastic dynamic game in discrete time would then be as follows. **Definition 5.4** An  $N$ -person discrete-time stochastic infinite dynamic game of prespecified fixed duration involves all but items (v) and (xi) of Def 5.1, and in addition (0) A finite or infinite set  $\Theta$ , with some topological structure, which denotes the action set of the auxiliary ( $N+1$ st) player, nature. Any permissible action  $\theta^k$  of nature at stage  $k$  is an element of  $\Theta$ . (v) A function  $F^k : X \times U^1 \times \dots \times U^N \times \Theta \rightarrow X$ , defined for each  $k \in K$ , so that  $x^{k+1} = F^k(x^k, u^1, \dots, u^N, \theta^k)$ ,  $k \in K$ , (5.11) where  $x^0$  is a random variable taking

values in  $X$ , and the joint probability distribution function of  $\{x^1, x^2, \dots, x^N\}$  is specified. (xi) A functional  $V : \{X \times U^1 \times \dots \times U^N \times G\} \times (X \times U^1 \times \dots \times U^N \times 0) \times \dots \times (X \times U^1 \times \dots \times U^N \times G) \rightarrow \mathbb{R}$  defined for each  $i \in N$ , and called the cost functional of  $P_i$  in the stochastic game of fixed duration. To introduce the noncooperative equilibrium solution concepts for stochastic games as formulated in Def. 5.4, we again have to transfer the original game in extensive form into equivalent normal form, by first computing  $L^*(-)$  ( $i \in N$ ) for each  $N$ -tuple of permissible strategies  $\{s^i \in T_i, i \in N\}$  and as a function of the random variables  $\{x^1, x^2, \dots, x^N\}$ , and then by taking expectation of the resulting  $N$ -tuple of cost functions over the statistics of these random variables. The resulting deterministic functions  $J^1, \dots, J^N$  ( $i \in N$ ) are known as the expected (average) cost functionals of the players, and they characterize the normal form of the original stochastic game together with the strategy spaces  $\{P_i; i \in N\}$ . We now note that, since such a description is free of any stochastic nature of the problem, all the solution concepts applicable to the deterministic dynamic game are also applicable to the stochastic dynamic game formulated in Def. 5.4, and hence the stochastic case need not be treated separately while introducing the equilibrium solution concepts. Remark 5.2 Definition 5.4 does not cover the most general class of stochastic dynamic games since (i) the state measurements of the players could be of stochastic nature by taking  $h_k$  to be a mapping:  $X \times 0 \rightarrow Y_k$ , (ii) the duration of the game might not be fixed a priori, but be a variant of the players' actions as well as the outcome(s) of the chance move(s), and (iii) the order in which the players act might not be fixed a priori, but again depend on the outcome of the chance move as well as on the actions of the players (Witsenhausen, 1971a). Such extensions will, however, not be considered in this book. Yet another

224 T. BAŞAR AND G. J. OLSDER possible extension which is also valid in deterministic dynamic games (and which is included in the formulation of finite games in extensive form in Chapters 2 and 3) is the case when the action sets  $X_k^i$  ( $i \in N, k \in K$ ) of the players are structurally dependent on the history of the evolution of the game. Such games will also not be treated in the following chapters. □

5.3 Continuous-Time Infinite Dynamic Games Continuous-time infinite dynamic games, also known as differential games in the literature, constitute a class of decision problems wherein the evolution of the state is described by a differential equation and the players act throughout a time interval. Hence, as a counterpart of Def. 5.1, we can formulate such games of

prespecified fixed duration as follows. Definition 5.5 A quantitative TV-person differential game of prespecified fixed duration involves the following: (i) An index set  $N = \{1, \dots, TV\}$  called the players' set. (ii) A time interval  $[0, T]$  which is specified a priori and which denotes the duration of the evolution of the game. (iii) An infinite set  $S_q$  with some topological structure, called the trajectory space of the game. Its elements are denoted as  $\{x(t), 0 < t < T\}$  and constitute the permissible state trajectories of the game. Furthermore, for each fixed  $t \in [0, T]$ ,  $x(t) \in S^\circ$ , where  $S^\circ$  is a subset of a finite dimensional vector space, say  $R^n$ . (iv) An infinite set  $U_i$  with some topological structure, defined for each  $i \in N$  and which is called the control (action) space of  $P_i$ , whose elements  $\{u_i(t), 0 < t < T\}$  are the control functions or simply the controls of  $P_i$ . Furthermore, there exists a set  $S_i \subset R^{m_i}$  ( $i \in N$ ) so that, for each fixed  $t \in [0, T]$ ,  $u_i(t) \in S_i$ . (v) A differential equation  $\dot{x} = f(t, x(t), u_1(t), \dots, u_N(t)), x(0) = x_0$  (5.12) whose solution describes the state trajectory of the game corresponding to the  $N$ -tuple of control functions  $\{u_i(t), 0 < t < T\}$  ( $i \in N$ ) and the given initial state  $x_0$ . (vi) A set-valued function  $r_i\{-\}$  defined for each  $i \in N$  as  $r_i\{t\} = \{x\{s\}, 0 < s < e\}$ ,  $0 < e < t$ , (5.13) where  $e$  is nondecreasing in  $t$ , and  $r_i\{t\}$  determines the state information gained and recalled by  $P_i$  at time  $t \in [0, T]$ . Specification of  $r_i\{-\}$  (in fact,

GENERAL FORMULATION 225  $t$  in this formulation) characterizes the information structure (pattern) of  $P_i$ , and the collection (over  $i \in N$ ) of these information structures is the information structure of the game. (vii) A sigma-field  $\mathcal{N}_i$ , in  $S^\circ$ , generated for each  $i \in N$  by the cylinder sets  $\{x \in S^\circ, x(s) \in B\}$  where  $B$  is a Borel set in  $S^\circ$  and  $0 < s < t$ . The sigma-field  $\mathcal{N}_i, t > t_0$ , is called the information field of  $P_i$ . (viii) A prespecified class  $P$  of mappings  $Y : [0, T] \times S^\circ \rightarrow S_i$ , with the property that  $u_i(t) = \gamma_i(t, x)$  is  $\mathcal{N}_i$ -measurable (i.e., it is adapted to the information field  $\mathcal{N}_i$ ).  $P$  is the strategy space of  $P_i$  and each of its elements  $\gamma_i^*$  is a permissible strategy for  $P_i$ . (ix) Two functionals  $q_i : S^\circ \rightarrow R$ ,  $g_i : [0, T] \times S^\circ \times S_1 \times \dots \times S_N \rightarrow R$  defined for each  $i \in N$ , so that the composite functional  $L\{u_1, \dots, u_N\} = \int_0^t g\{t, x\{t\}, u_1\{t\}, \dots, u_N\{t\}\} dt + \phi(x\{T\})$  (5.14) is well defined for every  $u^*(t) = Y\{t, x\}, Y \in P$  ( $j \in N$ ), and for each  $i \in N$ .  $L_i$  is the cost function of  $P_i$  in the differential game of fixed duration. A differential game, as formulated above, is yet not well defined unless we impose some additional restrictions on some of the terms introduced. In particular, we have to impose conditions on  $r_i$  and  $P$  ( $i \in N$ ), so that the differential equation (5.12) admits a unique solution for every  $N$ -tuple



$\{u_i(t) = \gamma^*(t, x, u_{-i}), i \in N\}$ , with  $\gamma^* \in P$ . A nonunique solution to (5.12) is clearly not allowed under the extensive form description of a dynamic game, since it corresponds to nonunique state trajectories (or game paths) and thereby to a possible nonuniqueness in the cost functions for a single  $i$ -tuple of strategies. We now provide below in Thm. 5.1 a set of conditions under which this uniqueness requirement is fulfilled. But first we list down some information structures within the context of deterministic differential games, as a counterpart of Def. 5.2.

**Definition 5.6** In an  $N$ -person continuous-time deterministic dynamic game (differential game) of prespecified fixed duration  $[0, T]$ , we say that  $P_i$ 's information structure is a(n) (i) open-loop (OL) pattern if  $\{x_0\}, t \in [0, T]$ , (ii) closed-loop perfect state (CLPS) pattern if  $\forall t \in [0, T], \{x(s), 0 < s < t\}$ , This term will be made precise in the sequel.

226 T. BAŞAR AND G. J. OLSDER (iii)  $\epsilon$ -delayed closed-loop perfect state (eDCLPS) pattern if  $\{x_0\}, 0 < t < T - \epsilon, \{x(s), 0 < s < t - \epsilon\}, \epsilon < T$  where  $\epsilon > 0$  is fixed, (iv) memoryless perfect state (MPS) pattern if  $\{x_0, x(t)\}, t \in [0, T]$ , (v) feedback (perfect state) (FB) pattern if  $\{x(t)\}, t \in [0, T]$ .

**Theorem 5.1** Within the framework of Def 5.5, let the information structure for each player be any one of the information patterns of Def 5.6. Furthermore, let  $S_0 \in \mathbb{R}^n$ . Then, if (i)  $f(t, x, u_1, \dots, u_N)$  is continuous in  $t \in [0, T]$  for each  $x \in S_0, i \in N$ , (ii)  $f(t, x, u_1, \dots, u_N)$  is uniformly Lipschitz in  $x, u_1, \dots, u_N$ ; i.e., for some  $k > 0, \forall t \in [0, T], \forall x, \tilde{x} \in S_0, \forall u_1, \dots, u_N, \tilde{u}_1, \dots, \tilde{u}_N \in W$  (i.e.,  $W$  is a compact subset of  $\mathbb{R}^m$ ),  $\|f(t, x, u) - f(t, \tilde{x}, u)\| \leq k \|x - \tilde{x}\|$ , (iii) for  $-f \in V$  ( $i \in N$ ),  $\forall \{t, x\}$  is continuous in  $t$  for each  $x(-) \in C^1[0, T]$  and uniformly Lipschitz in  $x(-) \in \mathbb{R}^n[0, T]$ , if the differential equation (5.12) admits a unique solution (i.e., a unique state trajectory) for every  $\gamma^* \in \Gamma$  ( $i \in N$ ), so  $\exists \gamma^*(t, x) = \gamma^*(t, x)$ , and furthermore this unique trajectory is continuous. **Proof.** It follows from a standard result on the existence of unique continuous solutions to differential equations. See for instance Coddington and Levinson (1955).

**Remark 5.3** Theorem 5.1 provides only one set of conditions which are sufficient to ensure existence of a unique state trajectory for every  $T$ -tuple of strategies  $\{\gamma^* \in \Gamma; i \in N\}$ , which necessarily implies a well-defined differential game problem within the framework of Def. 5.5. Since these conditions are all related to existence of a unique solution to (5.12) they can definitely be relaxed (but slightly) by making use of the available theory on functional differential equations (see, for example, Hale (1977)). But, Note that (iv) and (v) are not covered by (5.13) in Def. 5.5.  $\|\cdot\|$  denotes here the Euclidean norm for the

vector  $v$ .

GENERAL FORMULATION 227 inevitably, these conditions all involve some sort of Lipschitz-continuity on the permissible strategies of the players.

However, although such a restriction could be reasonable in the extreme case of one-person differential games (i.e., optimal control problems), it might be quite demanding in an  $N$ -player ( $N > 2$ ) differential game. To illustrate this point, consider, for instance, the one-player game described by the scalar differential equation  $\dot{z} = u$ ,  $z(0) = 0$ , and adopt the strategy  $u(t) = \text{sgn}\{x(t)\}$ . The solution to (i) with  $u(t) = \text{sgn}\{x(t)\}$  is clearly not unique; a multitude of solutions exists. If we adopt yet another strategy, viz.  $u(t) = -\text{sgn}\{x(t)\}$ , then the solution to (i) does not even exist in the classical sense. Hence, a relaxation of the Lipschitz-continuity condition on the permissible strategies could make an optimal control problem quite ill-defined. In such a problem, the single player may be satisfied with smooth (but sub-optimal) strategies. In differential games, however, it is unlikely that players are willing to restrict themselves to smooth strategies voluntarily. If one player would restrict his strategy to be Lipschitz, the other player(s) may be able to exploit this.

Specifically in pursuit evasion games (see Chapter 8) the players play "on the razor's edge", and such a restriction could result in a drastic change in the outcome of the game. Extensive discussions on these issues, and on various different definitions of "solution", can be found in the book by Krasovskii and Subbotin (1988). In conclusion, non-Lipschitz strategies cannot easily be put into a rigorous mathematical framework. On the other hand, in many games, we do not want the strategy spaces to comprise only smooth mappings. These difficulties may show up especially under the Nash equilibrium solution concept. In two-person Stackelberg games, however, non-Lipschitz strategies can more easily be handled in the general formulation, since one of the player's (follower's) choice of strategy is allowed to depend on the other player's (leader's) announced strategy.  $\square$  The saddle-point, Nash, Stackelberg and consistent conjectural variations equilibrium concepts introduced earlier for finite games are equally valid for (continuous-time) differential games if we bring them into equivalent normal form. To this end, we start with the extensive form description of a differential game, as provided in Def. 5.5 and under the hypotheses of Thm. 5.1, and for each fixed TV-tuple of strategies  $\{u^i \in U^i; i \in N\}$  we obtain the unique solution of the functional differential equation (5.15) and determine the corresponding action (control) vectors  $u^i(-) = u^i(-, x), i \in N$

$N, x \in S^\circ$ . Substitution of these into (5.14), together with the corresponding unique state trajectory, thus yields an  $N$ -tuple of numbers  $\{L_i; i \in N\}$ , for each choice of strategies by the players — assuming of course that functions  $g_i (i \in N)$  are integrable, so that (5.14) are well defined. Therefore, we have mappings  $J_i^\circ : T_1 \times \dots \times T_N \rightarrow R (i \in N)$  for each fixed initial state vector  $x_0$ , which we call the cost functional of  $P_i$  in a differential game in normal form. These cost functionals, together with the strategy spaces  $\{P_i; i \in N\}$  of the players, then constitute the equivalent

228 T. BAŞAR AND G. J. OLSDER normal form description of the differential game, which is the right framework to introduce noncooperative equilibrium solution concepts, as we have done earlier for other classes of dynamic games. Termination Definition 5.5 covers differential games of fixed prespecified duration; however, as discussed in Section 5.2 within the context of discrete-time games, it is possible to extend such a formulation so that the end point in both state and time is a variable. Let  $S^\circ$  again denote a subset of  $R^n$ , and  $R^+$  denote the half-open interval  $[0, \infty)$ . Let a closed subset  $A \subset S^\circ \times R^+$  be given, which we call a terminating (target) set, so that  $(:r_0, 0) \in A$ . Then, we say that the differential game terminates, for a given  $N$ -tuple of strategies, at time  $T \in R^+$  if  $T$  is the smallest element of  $R^+$  with the property  $(x(T), T) \in A$ , i.e.,  $T = \min\{t \in R^+ : (x(t), t) \in A\}$ . (5.16) This positive number  $T$  is called the terminal time of the differential game, corresponding to the given  $N$ -tuple of strategies. Terminal time is sometimes defined in a slightly different way, as the smallest time at which  $x(\cdot)$  penetrates,  $A$ , i.e.,  $T = \min\{t \in R^+ : (x(t), t) \in A\}$ , where  $A$  denotes the interior of  $A$ . The two definitions can only differ in situations where the trajectory  $x(\cdot)$  belongs to the boundary of  $A$  for a while or only touches  $A$  at one instant of time. Unless stated differently, we shall adopt (5.16) as the definition of the terminal time. The question of whether a given differential game necessarily terminates is one that requires some further discussion. If there is a finite time, say  $t_0$ , such that  $(x, t) \in A$  for all  $x \in S^\circ$ , then the game always terminates, in at most  $t_0$  units of time. Such a finite  $t_0$ , however, does not always exist, as elucidated by the following optimal control (one-player differential game) problem. Example 5.1 Consider the optimal control problem described by the two-dimensional state equation  $\dot{z}_i = u, z_i(0) = 1, \dot{z}_2 = -z_2 + u, z_2(0) = 1$ , where  $u(\cdot)$  is the control (action) variable satisfying the constraint  $0 < u(t) < 3$  for all  $t \in [0, \infty)$ . Let the terminal time  $T$  be defined as  $T = \min\{t \in [0, \infty) : z_2(t) = 2\}$ , and a permissible OL strategy 7

be defined as a mapping from  $[0, \infty)$  into  $[0, 3]$ . Now, if the cost function is  $L(u) = \int_0^T 1 dt = T$ , then the minimizing strategy is  $j^*(t) = 3, t > 0$ , with the corresponding terminal time (and thereby the minimum value of  $L$ ) being  $\ln 2$ .  
D

GENERAL FORMULATION 229 If the cost function to be minimized is  $L(u) = x(T)$ , however, the player may not have an incentive to terminate this "one-player game" since, as long as termination has not been achieved, his cost is not defined. So, on the one hand, he does not want to terminate the game; while on the other hand, if he has an incentive to terminate it, he should do it as soon as possible, which dictates him to employ the strategy  $j^*(t) = 3, t > 0$ . The player is faced with a dilemma here, since the latter cost function is not well defined for all strategies available to the player. This ambiguity can, however, be removed by either (i) restricting the class of permissible strategies to the class of so-called playable strategies which are those that terminate the game in finite time, or (ii) extending the definition of the cost functional so that  $L(x) = \int_0^T 1 dt$  if  $T$  is finite,  $\infty$  otherwise, which eliminates any possible incentive for the player not to terminate the game. In both cases, the optimal strategy will be  $j^*(t) = 3, t > 0$ . The difficulties encountered in the preceding optimal control example, as well as the proposed ways out of the dilemma, are also valid for differential games, and this motivates us to introduce the following concept of "playability".

Definition 5.7 For a given  $N$ -person differential game with a target set  $A$ , a strategy  $N$ -tuple is said to be playable (at  $(t_0, x_0)$ ) if it generates a trajectory  $x(t)$  such that  $(x(t), t) \in A$  for finite  $t$ . Such a trajectory  $x(t)$  is called terminating.

Differential games with chance moves<sup>38</sup> Chance moves can be introduced in the formulation of differential games by basically following the same lines as in the discrete-time case (cf. Def. 5.4), but this time one has to be mathematically more precise. In particular, if we assume the chance player (nature) to influence the state trajectory throughout a given time interval, then actions of this additional player will be realizations of a stochastic process  $\{w(t), t > 0\}$  whose statistics are known a priori. If we adjoin such a stochastic process to (5.12), then the resulting "differential equation"  $\dot{x} = F(t, x(t), u(t), w(t)), x(0) = x_0$ , might not be well defined, in the sense that even though its solution might be unique for each realization (sample path) of  $\{w(t), t > 0\}$ , it might not exist as a stochastic process. To obtain a meaningful formulation and tractable results, one has to impose some restrictions of  $F$  and  $\{w(t), t > 0\}$ . One particular such

38 Here it is assumed that the reader has some prior knowledge of stochastic processes. If this is not the case, either this part may be skipped (without much loss, since this formulation is used later only in Section 6.7) or the standard references may be consulted for background knowledge (Fleming and Rishel, 1975; Gikhman and Skorohod, 1972; Wong and Hajek, 1985).

230 T. BAŞAR AND G. J. OLSDER assumption is to consider equations of the form  $\dot{x} = F(s, x, v)(s), \dots, u_N(s) ds + a(s, x) dB_s$ , (5.17a) where  $F$  satisfies the conditions imposed on  $f$  in Def. 5.5, the function  $a$  satisfies similar conditions in its arguments  $s$  and  $x$  and  $\{B_t, t > 0\}$  is a special type of a stochastic process called the Wiener process. Equation (5.17a) can also be written symbolically as  $dx_t = F(*, :rt, u_1(\xi), \dots, u_N(\xi)) d* - f - cr(i, x_t) du_t$ ,  $x_t|_{t=0} = x_0$ , (5.17b) where we have used  $w_t$ , instead of  $B_t$ , to denote that it is the Wiener process. It should further be noted that in both (5.17a) and (5.17b), the function  $x(-)$  is written as  $\#(\cdot)$ , to indicate explicitly that it now stands for a stochastic process instead of a deterministic function. Equation (5.17b) is known as a stochastic differential equation, and existence and uniqueness properties of its solution, whenever  $u^i(-) = \gamma_i(-, x), \gamma_i \in P (i \in N)$ , are elucidated in the following theorem, whose proof can be found in Gikhman and Skorohod (1972). Theorem 5.2 Let  $T_1, \dots, T_N$  denote the strategy spaces of the players under any one of the information patterns of Def. 5.6. Furthermore let  $S_0 \in C^n[0, T]$ ,  $F$  satisfy the requirements imposed on  $f$  in Thm. 5.1, and  $\gamma^i \in P (i \in N)$  satisfy the restrictions imposed on  $\gamma^i$  in Thm. 5.1 (Hi). If, further,  $a$  is a nonsingular  $(n \times n)$  matrix, whose elements are continuous in  $t$  and uniformly Lipschitz in  $x$ , the stochastic differential equation (5.17b) with  $u^i(-) = \gamma^i(-, x), \gamma^i \in P (i \in N)$ , admits as its solution a unique stochastic process with continuous sample paths, for every such  $N$ -tuple of strategies. To complete the formulation of a differential game with chance moves (i.e., a stochastic differential game), it will now be sufficient to replace IS in Def. 5.5 (ix) with the expected value of the same expression, where the expectation operation is taken with respect to the statistics of the Wiener process  $\{w_t, t > 0\}$  and the initial state  $x_0$ . The game can then be converted into normal form by determining, in the usual sense, functions  $J^* : T_1 \times \dots \times T_N \rightarrow R (i \in N)$ , which is now the suitable form to introduce the solution concepts already discussed.

5.4 Mixed and Behavioral Strategies in Infinite Dynamic Games In Chapter 2, we have defined a mixed strategy for a player as a probability distribution on the space of his pure strategies, or equivalently, as a random

variable whose values are the player's pure strategies (cf. Def. 2.2), which was also adopted in Chapters 3 and 4, for finite games and static infinite games, respectively. Defined in this way, a mixed strategy is a mathematically well-established

GENERAL FORMULATION 231 object, mainly because the strategy spaces are finite for the former class of problems, and at most a continuum for the latter class of games, thereby allowing one to introduce (Borel-) measurable subsets of these strategy spaces, on which probability measures can be defined. An attempt to extend this directly to infinite dynamic games, however, is hampered by several measure-theoretic difficulties (Aumann, 1964). To illustrate the extent of these difficulties, let us consider a simple two-stage two-person dynamic game defined by the state equation  $x_2 = 1 + w_1$ ,  $x_3 = x_2 + u_2$ , where  $-1 < u_1 < 0$ ,  $0 < u_2 < 1$ , and  $P_2$  has access to closed-loop perfect state information (i.e., he knows the value of  $x_2$ )- A mixed strategy for  $P_1$  can easily be defined in this case, since his pure strategy space is  $[-1, 0]$  which is endowed with a measurable structure, viz. its Borel subsets. However, for  $P_2$ , the permissible (pure) strategies are measurable mappings from  $[0, 1]$  into  $[0, 1]$ , and thus the permissible strategy space of  $P_2$  is  $T_2 = I_1$ , where  $I$  denotes the unit interval. In order to define a probability distribution on  $I_1$ , we have to endow it with a measurable structure, but no such structure exists which is suitable for the problem under consideration (see Aumann, 1961). Intuitively, such a difficulty arises because the set of all probability distributions on  $I$  is already an extremely rich class, so that if one wants to define a similar object on a domain  $I_1$  whose cardinality is higher than that of  $I$ , such an increase in cardinality cannot reflect itself on the set of probability distributions. An alternative, then, is to adopt the "random variable" definition of a mixed strategy. This involves a sample space, say  $\Omega$ , and a class of measurable mappings from  $\Omega$ , into  $I_1$ , with each of these mappings being a candidate mixed strategy. The key issue here, now, is the choice of the sample space  $\Omega$ . In fact, since  $\Omega$  stands for a random device, the question can be rephrased as the choice of a random device whose outcomes are rich enough to be compatible with  $I_1$ . But the richest one is the continuous roulette wheel which corresponds to a sample space  $\Omega$  as a copy of the unit interval  $I$ .<sup>39</sup> Such a consideration then leads to the conclusion that the random variable (mixed strategy) should be a measurable mapping  $f$  from  $\Omega$ , into  $I_1$ . Unfortunately, this approach also leads to difficulties, since one still has to define a measurable

structure on the range space  $I_1$ . But now, if we note that, to every function  $f: Q \rightarrow I_1$ , there corresponds a function  $g: CI \times I \rightarrow I$  defined by  $g(u, x) = f(u)(x)$ , then the "measurability" difficulty is resolved since one can define measurable mappings from  $CI \times I$  into  $I$ . This conclusion of course remains valid if the action set  $I$  is replaced by any measurable space  $U$ , and the information space  $I$  is replaced by any measurable space  $N$ . Moreover, an extension to multi-stage infinite dynamic games is also possible as the following definition elucidates. <sup>39</sup>This follows from the intuitive idea of a random device (Aumann, 1964). It is, of course, possible to consider purely abstract random devices which are even richer.

232 T. BAŞAR AND G. J. OLSDER Definition 5.8 Given a  $K$ -stage discrete-time infinite dynamic game within the framework of Def 5.1, let  $U^k, N^k$  be measurable spaces for each  $i \in N, k \in K$ . Then, a mixed strategy  $\gamma_i$  for  $P_i$  is a sequence  $\gamma_i^* = (\gamma_i^1, \dots, \gamma_i^K)$  of measurable mappings  $\gamma_i^j : Q \times N^j \rightarrow f^j$  ( $j \in K$ ), so that  $\gamma_i^j(w, \cdot) \in P^j$  for every  $w \in \Omega$ , where  $\Omega$  is a fixed sample space (taken as a copy of the unit interval). A direct extension of this definition to continuous-time systems would be as follows. Definition 5.9 Given a differential game that fits the framework of Def. 5.5, let  $S_0$  and  $S^i$  ( $i \in N$ ) be measurable spaces. Then, a mixed strategy  $\gamma_i^*$  for  $P_i$ , in this game, is a measurable transformation  $\gamma_i^* : \Omega \times [0, T] \times S_0 \rightarrow S^i$ , so that  $\gamma_i^*(a, \cdot, \cdot) \in P^i$  for every  $w \in \Omega$ , where  $Q$  is a fixed sample space (taken as a copy of the unit interval). Equivalently,  $\gamma_i^*(\cdot, \cdot, x)$  is a stochastic process for each  $x \in S_0$ . A behavioral strategy, on the other hand, has been defined (cf. Def. 2.9) as a collection of probability distributions, one for each information set of the player. In other words, using the notation of Section 2.4, it is a mixed strategy  $\gamma_i(\cdot)$  with the property that  $\gamma_i^1$  and  $\gamma_i^2$  are independent random variables whenever  $I_1$  and  $I_2$  belong to different information sets. If we extend this notion directly to multi-stage infinite dynamic games (and within the framework of Def. 5.8), we have to define a behavioral strategy as a mixed strategy  $\gamma_i^*(\cdot, \cdot)$  with the property that the collection of random variables  $\{\gamma_i^j(\cdot, \cdot, \cdot), \gamma_i^j \in N^j; j \in K\}$  is mutually independent. But, since  $\Omega$  is in general a non-denumerable set, in infinite dynamic games this would imply that we have a non-denumerable number of mutually independent bounded random variables on the same sample space. This is clearly not possible, since any bounded random variable defined on our sample space  $\Omega$  should have a countable basis (Loeve, 1963). Then, a way out of this difficulty is to define a behavioral

strategy as a mixed strategy which is independent from stage to stage, but not necessarily stagewise. Aumann actually discusses that stagewise correlation is quite irrelevant (Aumann, 1964), and the expected cost function is invariant under such a correlation. Therefore, we have the following definition.

**Definition 5.10** Given a  $K$ -stage discrete-time infinite dynamic game within the framework of Def. 5.1, let  $U_i$ ,  $A_i$  be measurable spaces for each  $i \in N$ ,  $k \in K$ . Then, a behavioral strategy  $\sigma_i^*$  for  $P_i$  is a mixed strategy (cf. Def. 5.8) with the further property that the sequence of random variables  $\{t_j^i(-, \sigma_j^*), j \in K\}$  is mutually independent for every fixed  $r_i^i \in N_j$  ( $j \in K$ ). This definition of a behavioral strategy can easily be extended to multi-stage games with a (countably) infinite number of stages; however, an extension to continuous-time infinite dynamic games is not possible since the time set  $[0, T]$  is not denumerable.

## GENERAL FORMULATION 233 5.5 Tools for One-Person Optimization

Since optimal control problems constitute a special class of infinite dynamic games with one player and one criterion, the mathematical tools available for such problems may certainly be useful in dynamic game theory. This holds particularly true if the players adopt the noncooperative Nash equilibrium solution concept, in which case each player is faced with a single criterion optimization problem (i.e., optimal control problem) with the strategies of the remaining players taken to be fixed at their equilibrium values. Hence, in order to verify whether a given set of strategies is in Nash equilibrium, we inevitably have to utilize the tools of optimal control theory. We, therefore, present in this section some important results on dynamic one-person optimization problems so as to provide an introduction to the theory of subsequent chapters. The section comprises three subsections. The first two deal with the dynamic programming (DP) technique applied to discrete-time and continuous-time optimal control problems, and the third is devoted to the "minimum principle". For more details on, and a rigorous treatment of, the material presented in these subsections the reader is referred to Fleming and Rishel (1975), Fleming and Soner (1993) and Boltyanski (1978).

### 5.5.1 Dynamic programming for discrete-time systems

The method of dynamic programming is based on the principle of optimality which states that an optimal strategy has the property that, whatever the initial state and time are, all remaining decisions (from that particular initial state and particular initial time onwards) must also constitute an optimal strategy. To exploit this principle, we work backwards in time,



starting at all possible final states with the corresponding final times. Such a technique has already been used in this book within the context of finite dynamic (multi-act) games, specifically in Sections 2.5 and 3.5, in the derivation of noncooperative equilibria, though we did not refer explicitly to dynamic programming. We now discuss the principle of optimality within the context of discrete-time systems that fit the framework of Def. 5.1, but with only one player (i.e.,  $N = 1$ ). Toward that end, we consider equation (5.9b), assume feedback perfect state information and a stage-additive cost functional of the form (5.10), all for  $N \sim 1$ , i.e.,  $Z_{f_{k+1}} = f_k(x_k, u_k)$ ,  $u_k \in U_k$ ,  $k \in K$ , (5.18a)  $L\{u\} = \sum_{k=0}^{N-1} g_k(x_k, u_k)$ , (5.18b)  $f_0 = x_0$  where  $u = \{u_k, k \in K\}$ ,  $u_k = u|_k = f_k(x_k)$ ;  $f^*(-)$  denotes a permissible (control) strategy at stage  $k \in K$ , and  $K$  is a fixed positive integer. In order to determine the minimizing control strategy, we shall need the expression for the minimum (or minimal) cost from any starting point at any initial time. This is also called

234 T. BAŞAR AND G. J. OLSDER the value function, and is defined as  $K$   
 $V(k, x) = \min_{u \in U_k} \{g_k(x_k, u_k) + V(k+1, f_k(x_k, u_k))\}$  with  $U_i = \{u_i(x_i) \mid G U_i \text{ and } x_k = x\}$ . A direct application of the principle of optimality now readily leads to the recursive relation  $V(k, x) = \min_{u \in U_k} [g_k(f_k(x, u), u, x) + V(k+1, f_k(x, u))]$ . (5.19)  
 $u_k \in U_k$  If the optimal control problem admits a solution  $u^* = \{u_k, k \in K\}$ , then the solution  $V(1, X)$  of (5.19) should be equal to  $L(u^*)$ , and furthermore each  $u_k^*$  should be determined as an argument of the RHS of (5.19). Affine-quadratic problems As a specific application, let us consider the so-called affine-quadratic (or linear-quadratic) discrete-time optimal control problem, which is described by the state equation  $X_{k+1} = A_k x_k + B_k u_k + C_k f_k$ , (5.20a) and cost functional  $J(u) = \sum_{k=0}^{K-1} (x_k^T Q_k x_k + u_k^T R_k u_k)$ . (5.20b) Let us further assume that  $x_k \in \mathbb{R}^n$ ,  $u_k \in \mathbb{R}^m$ ,  $C_k \in \mathbb{R}^n$ ,  $R_k > 0$ ,  $Q_{k+1} > 0$  for all  $f_k \in K$ , and  $A_k, B_k$  are matrices of appropriate dimensions. Here, the corresponding expression for  $f_k$  is obvious, but the one for  $g_k$  is not uniquely defined; though, it is convenient to take it as  $g_k(v_k, X_k) = \frac{1}{2} v_k^T R_k v_k + \frac{1}{2} x_k^T Q_k x_k$ ,  $k=0, \dots, K-1$ ,  $\frac{1}{2} x_K^T Q_K x_K$ ,  $k=K$ . We now obtain, by inspection, that  $V(k, x)$  should be a general quadratic function of  $x$  for all  $f_k \in K$ , and that  $V\{K+1, x\} = \frac{1}{2} x^T Q_K x$ . This leads to the structural form  $V(k, x) = \frac{1}{2} x^T S_k x + x^T s_k + q_k$ . Substituting this in the recursive relation (5.19) we obtain the unique solution of the optimal control problem (5.20a)-(5.20b) as follows.  
**Proposition 5.1** The optimal control problem (5.20a)-(5.20b) admits the unique solution  $u_k^* = -R_k^{-1} (B_k^T S_{k+1} A_k x_k + S_{k+1} C_k)$ , (5.21a)

GENERAL FORMULATION 235 for all  $k \in \mathbb{K}$ , where  $P_k = [R_k + B'kS_{k+1}B_k] -$   
 $LB'k$ ,  $S_k = Q_k + A'kS_{k+1}[I - B_kP_kS_{k+1}]A_k$ ;  $S_{K+1} = Q_{K+1}$ , } (5.21b)  $s_{fc} =$   
 $A'k[I - B_kP_kS_{k+1}][s_{k+1} + S_{k+1}C_k]$ ;  $s_{K+1} = 0$ . Furthermore, the minimum value  
of (5.20b) is  $L\{u^*\} = -x_i S_m + z_{i+1} g_i$ , (5.22) where  $1 \leq i \leq J$  ( $c'k s_{k+1} + i c_k$   
 $- (s_{fc+i} + 5f_{c+i} c_{fc})' P_{fc} S_{fc} (s_{fc+i} + S_{k+1} c_k) + 2c'k s_{k+1}$ ). Remark 5.4 If the  
requirement  $Q_k > 0$  is not satisfied, a necessary and sufficient condition for  
(5.21a)-(5.21b) to provide a unique solution to the affine-quadratic discrete-  
time optimal control problem is  $R_k + B'kS_{k+1}B_k > 0$ ,  $\forall k \in \mathbb{K}$ . D Infinite horizon  
linear-quadratic problems To formulate a possibly meaningful linear-quadratic  
optimal control problem when  $K \rightarrow \infty$ , we take  $c_k \rightarrow 0$ , and the matrices  
 $A_k, B_k, Q_k$  and  $R_k$  to be independent of  $k$  (which will henceforth be written  
without the index  $k$ ). Thus the optimization problem is  $\min - Y^{\wedge}(x'k+i Q x_{k+i} +$   
 $u'k R u_k)$ ;  $Q > 0$ ,  $R > 0$   $f_{c=i}$  subject to  $x_{k+1} = A x_k + B u_k$ . It is natural to assume  
here that the problem is well defined, in the sense that there exists at least one  
control sequence that renders it a finite cost. Conditions which ensure this are  
stabilizability of the matrix pair  $(A, B)$ , and detectability of the matrix pair  $(A,$   
 $D)$ , where  $D$  is a matrix such that  $D'D = Q$ .<sup>40</sup> These notions of stabilizability  
and detectability belong to the realm of the theory of linear systems; see for  
instance Kailath (1980) or Anderson and Moore (1989). The pair  $(A, B)$  is  
stabilizable if an  $m \times n$  matrix  $F$  exists such that  $(A+BF)$  is a stable matrix, i.e.,  
all its eigenvalues lie strictly within the unit circle. The pair  $(A, D)$  is  
detectable if its "dual pair",  $(A', D')$  is stabilizable. The following result is now  
a standard one, which can be found in any textbook on linear control systems.  
<sup>40</sup>Using standard terminology, we will also refer to the latter condition as  
detectability of the pair  $(A, Q)$ .

236 T. BAŞAR AND G. J. OLSDER Proposition 5.2 Assume that the pair  $(A,$   
 $B)$  is stabilizable and the pair  $(A, Q)$  is detectable. Then, there exists an  $n \times n$   
nonnegative-definite matrix  $S$  such that (i) for fixed  $k$ ,  $S_j \rightarrow S$  as  $K \rightarrow \infty$ ,  
where for each  $K$ , and arbitrary  $Q_{k+i} > 0$ ,  $S_{fc}$  is recursively defined by  
(5.21a)-(5.21b); (ii)  $S$  is the unique solution of the algebraic Riccati equation  
(ARE)  $S = Q + A'S[I - B\{R + B'SB\} - LB'S]A$  within the class of nonnegative-  
definite matrices; (iii) the (closed-loop) matrix  $A - B(R + B'SB)^{-1}B'S$  is  
stable, i.e., all its eigenvalues lie strictly within the unit circle; (iv) the  
minimum value of the cost functional is  $\sum_{i=0}^{\infty} x' S x_i$ ; (v) the stationary optimal  
control law is  $u^*(z) = -[R + B'SB]^{-1}B'SAx$ . Under the stronger assumption that  
 $(A, D)$  is observable (a sufficient condition for which is  $Q > 0$ ), the solution of

the ARE is positive definite. 5.5.2 Dynamic programming for continuous-time systems The dynamic programming approach, when applied to optimal control problems defined in continuous time, leads to a partial differential equation (PDE) which is known as the Hamilton-Jacobi-Bellman (HJB) equation.

Toward this end we consider the optimal control problem defined by  $\dot{x}(t) = f(t, x(t), u(t))$ ,  $x(0) = x_0$ ,  $t \in [0, T]$ ,  $u(t) \in U$ ,  $L(u) = \int_0^T g(t, x(t), u(t)) dt + q(T, x(T))$ : (5.23)  $T = \{t \in [0, T] : I(t, x(t)) = 0\}$ , where  $I$  is a scalar function, defining an  $n$ -dimensional smooth manifold in the product space  $\mathbb{R}^n \times \mathbb{R}^+$ , and  $T$  is taken to be the class of all admissible feedback strategies. The minimum cost-to-go from any initial state  $(x)$  and any initial time  $(t)$  is described by the so-called value function which is defined by  $V(t, x) = \min_{u(s), t \leq s < T} \int_t^T g(s, x(s), u(s)) ds + q(T, x(T))$  (5.24a) satisfying the boundary condition  $V(T, x) = q(T, x)$  along  $I(T, x) = 0$ . (5.24b)

GENERAL FORMULATION 237 A direct application of the principle of optimality on (5.24a), under the assumption of continuous differentiability of  $V$ , leads to the HJB equation  $dV(t, x) = \min_u \{ -\dot{V}(t, x) = g(t, x, u) + \sum_{i=1}^n \lambda_i (f_i(t, x, u) - \dot{x}_i) \}$  (5.25) which takes (5.24b) as the boundary condition. In general, it is not easy to compute  $V(t, x)$ . Moreover, the continuous differentiability assumption imposed on  $V(t, x)$  is rather restrictive (see, for instance, Example 5.2 in subsection 5.5.3, for which it does not hold). Nevertheless, if such a function exists, then the HJB equation (5.25) provides a means of obtaining the optimal control strategy. This "sufficiency" result is now proven in the following theorem. Theorem 5.3 If a continuously differentiable function  $V(t, x)$  can be found that satisfies the HJB equation (5.25) subject to the boundary condition (5.24b), then it generates the optimal strategy through the static (pointwise) minimization problem defined by the RHS of (5.25). Proof. If we are given two strategies,  $u^*$  (the optimal one) and  $u$  (an arbitrary one), with the corresponding terminating trajectories  $x^*$  and  $x$ , and terminal times  $T^*$  and  $T$ , respectively, then (5.25) reads  $\dot{V}(t, x) = \min_u \{ g(t, x, u) + \sum_{i=1}^n \lambda_i (f_i(t, x, u) - \dot{x}_i) \}$ , (5.26a)  $\dot{V}(t, x^*) = g(t, x^*, u^*) + \sum_{i=1}^n \lambda_i (f_i(t, x^*, u^*) - \dot{x}_i^*)$ , (5.26b) where  $u^*$  and  $u$  have been replaced by the corresponding controls  $u^*$  and  $u$ , respectively. Integrating (5.26a) from 0 to  $T$  and (5.26b) from 0 to  $T^*$ , we obtain  $\int_0^T g(t, x, u) dt + V(T, x(T)) - V(0, x_0) > 0$ , (5.27)  $\int_0^{T^*} g(t, x^*, u^*) dt + V(T^*, x^*(T^*)) - V(0, x_0) = 0$  (5.28) Elimination of  $V(0, x_0)$  yields  $\int_0^T g(t, x, u) dt + q(T, x(T)) > \int_0^{T^*} g(t, x^*, u^*) dt + q(T^*, x^*(T^*))$ , (5.29) from which it readily follows that  $u^*$  is the optimal control, and therefore  $u^*$  is the optimal strategy.

## D

238 T. BAŞAR AND G. J. OLSDER Remark 5.5 If, in the problem statement (5.23), the time variable  $t$  does not appear explicitly, i.e., if  $f$ ,  $g$ ,  $q$  and  $r$  are time-invariant, the corresponding value function will also be time-invariant. This then implies that  $dV/dt = 0$ , and the resulting optimal strategy can be expressed as a function of only  $x(t)$ , i.e.,  $u^*(t) = \gamma^*(t, x(t))$ . In such cases, we will write  $V(x)$  for  $V(t, x)$ .  $\square$  Remark 5.6 There exists an alternative derivation of the HJB equation (5.25), which is of a geometrical nature. This will not be discussed here, since it is a special case of the more general derivation to be given in Section 2 of Chapter 8 for two-player zero-sum differential games.  $\square$  Affine-quadratic problems We now consider an important class of problems—the so-called affine-quadratic (or linear-quadratic) continuous-time optimal control problems—for which  $V(t, x)$  is continuously differentiable, so that Thm. 5.3 applies. Toward that end, let the system be described (as a continuous-time counterpart of (5.20a)-(5.20b)) by  $\dot{x}(t) = A(t)x(t) + B(t)u(t) + c(t)$ ;  $x(0) = x_0$ , (5.30a) and the cost functional to be minimized be given as  $J(u) = \int_0^T L(u) dt + x'(T)Qx(T) + x'(0)Q_0x(0) + \int_0^T (x'Qx + 2x'p + u'Ru) dt$ , (5.30b) where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$ ,  $0 < t < T$  and  $T$  is fixed.  $A(-)$ ,  $B(-)$ ,  $Q(-) > 0$ ,  $R(-) > 0$  are matrices of appropriate dimensions and with continuous entries on  $[0, T]$ . The matrix  $Q$  is nonnegative-definite, and  $c(-)$  and  $p(-)$  are continuous vector-valued functions, taking values in  $\mathbb{R}^n$ . Furthermore, we adopt the feedback information pattern and take a typical control strategy as a continuous mapping  $\gamma : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Denote the space of all such strategies by  $\mathcal{T}$ . Then the optimal control problem is to find a  $\gamma^* \in \mathcal{T}$  such that  $J(\gamma^*) \leq J(\gamma)$ ,  $\forall \gamma \in \mathcal{T}$ , (5.31a) where  $J(\gamma) = J(u)$ , with  $u(-) = \gamma(-, x)$ . (5.31b) Several methods exist to obtain the optimal strategy  $\gamma^*$  or the optimal control function  $u^*(-) = \gamma^*(-, x)$ . We shall derive the former by making use of Thm. 5.3. Simple arguments (see Anderson and Moore, 1989) lead to the conclusion that  $\min_{\gamma} J(\gamma)$  is quadratic in  $x_0$ . Moreover, it can be shown that, if the system is positioned at an arbitrary  $t \in [0, T]$  at an arbitrary point  $x \in \mathbb{R}^n$ , the minimum cost-to-go, starting from this position, is quadratic in  $x$ . Therefore, we may assume existence of a continuously differentiable value function of the form  $V(t, x) = -x'S(t)x + k'(t)x + m(t)$ , (5.32)

GENERAL FORMULATION 239 that satisfies (5.25). Here  $S(-)$  is a symmetric ( $n \times n$ ) matrix with continuously differentiable entries,  $f_c(-)$  is a

continuously differentiate  $n$ -vector and  $m(\cdot)$  is a continuously differentiable function. If we can determine such  $S(\cdot)$ ,  $f_c(\cdot)$  and  $m(\cdot)$  so that (5.32) satisfies the HJB equation (5.25), then Thm. 5.3 justifies the assumption of the existence of a value function quadratic in  $x$ . Substitution of (5.32) into (5.25) leads to —

$$x'Sx - x'k - m^2 = \min_{u} \{ (Sx + k)'(Ax + Bu + c) + -x'Qx + x'p + -u'Ru \} \quad (5.33)$$

(5.33)  $S(T) = Qf$ ,  $k(T) = 0$ ,  $m(T) = 0$ . (5.35a) (5.35b) (5.35c) Carrying out the minimization on the RHS yields  $u^*\{t\} = y^*(t, x(t)) = -R^{-1}B'SMx(t) + k\{t\}$ , (5.34) substitution of which into (5.33) leads to an identity relation which is readily satisfied if  $S + SA + A'S - SBR^{-1}B'S + Q = 0$ ,  $k + (A - BR^{-1}B'S)'k + Sc + p = 0$ ,  $m + k'c - k'BR^{-1}B'k = 0$ . Thus we arrive at the following proposition.

**Proposition 5.3** The linear-quadratic optimal control problem (5.30a)-(5.30b) admits a unique optimum feedback controller  $u^*$  which is given by (5.34), where  $S(\cdot)$ ,  $k(\cdot)$  and  $m(\cdot)$  uniquely satisfy (5.35a)-(5.35c). The minimum value of the cost functional is  $J\{1^*\} = S(0)x_0 + k'(0)x_0 + m(0)$ . Proof Except for existence of a unique  $S(\cdot) > 0$ ,  $f_c(\cdot)$  and  $m(\cdot)$  that satisfy (5.35a)-(5.35c), the proof has been given prior to the statement of the proposition. Furthermore, if there exists a unique  $S(\cdot)$  that satisfies (5.35a), which is known as the matrix Riccati equation, existence of unique solutions to the remaining two differential equations of (5.35a)-(5.35c) is assured since they are linear in  $k$  and  $m$ , respectively. Then, what remains to be proven is that a unique solution  $S(\cdot) > 0$  to (5.35a) exists on  $[0, T]$ . This can be verified by either making use of the theory of differential equations (cf. Reid, 1972) or utilizing the specific form of the optimal control problem with  $Qf$  and  $Q(\cdot)$  taken to be nonnegative-definite (Anderson and Moore, 1989).  $\square$

**Remark 5.7** The solution (5.34) can be obtained by other methods as well. Two of these are the minimum principle (to be discussed shortly; see also (Bryson and Ho, 1975)), and the "completion of squares" method (see Brock et al., 1970). Since the latter will be discussed in Chapter 6 in the context of affine-quadratic differential games, it will not be covered in this chapter.  $\square$

240 T. BAŞAR AND G. J. OLSDER **Remark 5.8** The nonnegative-definiteness requirements on  $Qf$  and  $Q(\cdot)$  may be relaxed, but then we have to assume from the outset the existence of a unique bounded solution to (5.35a) in order to ensure the existence of a unique minimizing control as given by (5.34).<sup>41</sup> Otherwise (5.35a) might not admit a solution; more precisely its solution may exhibit finite escape (depending on the length of the time interval), implying in turn that the "optimal" cost might tend to  $-\infty$ . To exemplify this situation

consider the scalar example:  $\dot{x} = x + u$ ,  $x(0) = x_0$ ,  $J = \int_0^T (x^2 + u^2) dt$ . (5.36)  
 The Riccati equation (5.35a), for this example, becomes  $\dot{S} + 2S - 1 = 0$ ,  
 $S(T) = 0$  (5.37) which admits the solution  $S = \frac{1}{2}(1 - e^{-2(T-t)})$  (5.38) on the interval  $(T - 1, T]$ . Hence, for  $T > 1$ , a continuously differentiable solution on the interval  $[0, T]$  to (5.37) does not exist. This non-existence of a solution to the matrix Riccati equation is directly related to the well-posedness of the optimal control problem (5.36), since it can readily be shown that, for  $T > 1$ ,  $J(u)$  can be made arbitrarily small (negative) by choosing a proper  $u(t)$ . In such a case we say that the Riccati equation has a conjugate point in  $[0, T]$  (Sagan (1969) and Brockett (1970)). This topic will be revisited in Chapter 6 in the context of affine-quadratic zero-sum differential games.  $\square$  Infinite horizon linear-quadratic problems Meaningful continuous-time linear-quadratic optimal control problems where  $T \rightarrow \infty$  can also be formulated. Toward that end we take  $c(t) = 0$ , and the matrices  $A(t), B(t), Q(t)$  and  $R(t)$  to be independent of  $t$ . Then the problem becomes  $J^* = \min_u \int_0^\infty (x'Qx + u'Ru) dt$ ;  $Q > 0, R > 0, u \in \mathbb{R}^m$  subject to  $\dot{x} = Ax + Bu$ . Conditions which ensure this problem to be well defined (with a finite minimum) are precisely those of the discrete-time analogue of this problem, i.e., the matrix pair  $(A, B)$  must be stabilizable and the matrix pair  $(A, D)$ , where  $D$  is a matrix such that  $D'D = Q$ , must be detectable; see for instance Kailath (1980) 41 See Chapter 6, and particularly Lemma 6.4 and Remark 6.15 for further elaboration of this point.

GENERAL FORMULATION 241 or Davis (1977) for discussions on these notions in continuous time. 42 The pair  $(A, B)$  is stabilizable if an  $m \times n$  matrix  $F$  exists such that  $(A + BF)$  is stable, i.e., all its eigenvalues lie in the open left half of the complex plane (note the difference of this definition of stability from the one for its discrete-time counterpart). Furthermore, the pair  $(A, D)$  is detectable if its "dual pair"  $(A', D')$  is stabilizable. The following proposition, which can be found in any textbook on linear control systems, summarizes the main result. Proposition 5.4 Assume that the pair  $(A, B)$  is stabilizable and the pair  $(A, Q)$  is detectable. Then, there exists an  $n \times n$  nonnegative definite matrix  $S$  such that (i) for fixed  $t$ , and for every  $Q_f > 0$ ,  $S_f(t) \rightarrow S$  (uniformly in  $t < T$ ) as  $T \rightarrow \infty$ , where  $S_f(t)$  is the unique nonnegative definite solution to (5.35a), parameterized in  $T$ ; (ii)  $S$  is the unique solution of the algebraic Riccati equation (ARE)  $SA + A'S - SBR^{-1}B'S + Q = 0$  in the class of nonnegative definite matrices; (iii) the closed-loop matrix  $A - BR^{-1}B'S$  is stable, i.e., all its eigenvalues lie in the open left half of the complex plane;

(iv) the minimum value of the cost functional (v) the stationary optimum control law is  $u^*(x) = -R^{-1}B'Sx$ . Under the stronger condition that  $(A, D)$  is observable (a sufficient condition for which is  $Q > 0$ ), the solution to the ARE is positive definite. 5.5.3 The minimum principle Continuous-time systems In this subsection our starting point will be the HJB equation (5.25) under the additional assumption that  $V$  is twice continuously differentiable, and we shall convert it into a series of pointwise optimization problems, indexed by the parameter  $t$ . This new form is in some cases more convenient to work with and it is closely related to the so-called minimum principle, as also discussed here. Toward this end we first introduce the function  $H(t, x, u) = \frac{1}{2}x^T Q x + f(t, x, u) + g(t, x, u)$ , (5.39) 42As in the discrete-time case, we will refer to the latter condition also as detectability of the pair  $(A, Q)$ .

242 T. BAŞAR AND G. J. OLSDER in terms of which equation (5.25) can be written as  $\frac{dV}{dt} = -\min_u H(t, x, u)$ . (5.40) at  $u$  Being consistent with our earlier convention, the minimizing  $u$  will be denoted by  $u^*$ . Then  $H(t, x, u^*) + \frac{dV}{dt}(t, x^*(t)) = 0$ . (5.41) Since this is an identity in  $x$ , its partial derivative with respect to  $x$  is also zero. This leads to, by also interchanging the orders of second partial derivatives (which is allowed if  $V$  is twice continuously differentiable):  $\frac{d}{dt} \left( \frac{dH}{du} \right) = \frac{d}{dx} \left( \frac{dH}{du} \right) \frac{dx}{dt}$  If there are no constraints on  $u$ , then  $\frac{dH}{du} = 0$  for  $u = u^*$  according to equation (5.40). If there are constraints on  $u$ , and  $u^*$  is not an interior point, then it can be shown that  $(\frac{dH}{du}) \cdot (\frac{du^*}{dx}) = 0$  (because of optimality,  $\frac{dH}{du}$  and  $\frac{du^*}{dx}$  are orthogonal; for specific problems we may have  $\frac{du^*}{dx} = 0$ ). In view of this, equation (5.42) becomes  $\frac{d}{dt} \left( \frac{dV}{dx} \right) = -\frac{d}{dx} H(t, x, u^*)$  By introducing the so-called costate vector,  $p'(t) = \frac{dV}{dx}(t, x^*(t))$ , where  $x^*$  denotes the state trajectory corresponding to  $u^*$ , (5.43) can be re-written as  $\frac{d}{dt} p' = -\frac{\partial}{\partial x} H(t, x^*, u^*) + p'(t) \frac{\partial}{\partial x} f(t, x^*, u^*) = -\frac{\partial}{\partial x} H(t, p, x^*, u^*)$ , (5.44) where  $H$  is defined by  $H(t, p, x, u) = g(t, x, u) + p' f(t, x, u)$ . (5.45) The final time  $T$  is determined by the scalar relation  $l(T, x) = 0$  and hence can be regarded as a function of the state:  $T(x)$ . The boundary condition for  $p(t)$  is determined from  $\frac{dx}{dt} = f(t, x, u)$  In conclusion, we have arrived at the following necessary condition for the optimal control  $u^*(t)$ : under the assumption that the value function  $V(t, x)$  is twice continuously differentiable, the optimal control  $u^*(t)$  and corresponding

GENERAL FORMULATION 243 (5.47) trajectory  $x^*(t)$  must satisfy the following so-called canonical equations  $\dot{x}^*(t) = f(t, x^*, u^*), x^*(t_0) = x_0; \dot{p}'(t) = -\frac{\partial}{\partial x} H(t, p, x^*, u^*), p'(t_0) = p_0$ .

$dH(t, p, x^*, u^*) = p'(T) = 9qilf$  along  $l(T, x) = 0$  -  $H(t, p, x, u) = g(t, x, u) + p'f(t, x, u)$ ,  $u^*(t) = \arg \min H(t, p, x^*, u)$ .  $u \in S$  In the derivation of (5.47), the controls have been assumed to be functions of time and state; i.e.,  $u(t) = 7(4, x(i))$ . If, instead, one starts with the class of control functions which depend only on time, i.e.,  $u(t) = 7(4)$ , the set of necessary conditions (5.47) can be derived in quite a different way (by using perturbation functions, which is a standard procedure in the classical calculus of variations), and under milder assumptions. Following such a derivation, one obtains the following result which can, for instance, be found in Pontryagin, et al. (1962) and which is referred to as the minimum principle.

**Theorem 5.4** Consider the optimal control problem defined by (5.23) and under the OL information structure. If the functions  $f$ ,  $g$ ,  $q$  and  $I$  are continuously differentiable in  $x$  and continuous in  $t$  and  $u$ , then relations (5.47) provide a set of necessary conditions for the optimal control and the corresponding optimal trajectory to satisfy.

**Remark 5.9** A particular class of optimal control problems which are not covered by Thm. 5.4 are those with a fixed terminal condition, i.e., with the terminal time  $T$  and the terminal state  $x(T)$  prescribed, since the function  $J$  is not differentiable in this case. However, the necessary conditions for the optimal control are still given by (5.47), with the exception that the condition on  $p(T)$  is absent and instead the terminal state constraint  $x(T) = X_f$  replaces it. We also note, in passing, that in the theory of zero-sum differential games, reasonable problems with state space dimension  $> 2$  will (almost) never be formulated with a fixed prescribed terminal state, since it is not possible to enforce this end-point constraint on both players if they have totally conflicting objectives.  $\square$

The following example now illustrates how Thm. 5.4 can be employed to obtain the solution of an optimal control problem for which the value function  $V$  is not continuously differentiable. It also exhibits some features which are frequently met in the theory of differential games, and introduces some terminology which will be useful later in Chapter 8.

244 T. BAŞAR AND G. J. OLSDER **Example 5.2** Consider an open-loop control problem with systems dynamics described by  $\dot{x} = -Gi + Gu$  where the scalar control satisfies the constraint  $-1 < u(t) < 1$  for all  $t > 0$ . The objective is to steer the system from an arbitrary but known initial point in the  $(x_1, x_2)$  plane to the line  $x_1 = x_2$  in minimum time. Hence, the cost function can be written as  $J(u) = \int_0^T 1 dt = T$ , where  $T$  is defined as the first instant for which the function  $l(t, x_1(t), x_2(t)) = x_1(t) - x_2(t)$  becomes zero. In the



derivation to follow, we shall restrict ourselves to initial points satisfying  $x_1 - x_2 > 0$ . Initial points with the property  $x_1 - x_2 < 0$  can be dealt with analogously. Now, application of relations (5.47) yields (with  $P = (P_1, P_2)'$ )  $H = 1 + p_1 x_1 + p_2 u$ ,  $P_1 = 0$ ,  $p_2 = -p_1$ , (5.49)  $u^*(t) = -\text{sgn}\{p_2(t)\}$ . Where the costate variable  $p_2(t)$  here is also known as the switching function, since it determines the sign of  $u^*(t)$ . For points on the line  $x_1 = x_2$ , which can be reached by optimal trajectories, obviously  $P_1 = 0$  and therefore the costate vector  $p(T) = (p_1(T), p_2(T))'$  will be orthogonal to the line  $x_1 = x_2$  and point in the direction "south-east". Since the magnitude of the costate vector is irrelevant in this problem, we take  $P_1(T) = 1$ ,  $p_2(T) = -1$ , (5.50) which leads to assuming, of course, that  $T > 1$ . Now, by integrating (5.48) backwards in time from an arbitrary terminal condition on the line  $x_1 = x_2$  (which we take as  $x_1(T) = x_2(T) = a$  where  $a$  is the parameter), and by replacing  $u(t)$  with  $u^*(t)$  as given by (5.51), we obtain  $x_1(t) = (T-t)^a + a$ ,  $x_2(t) = a - (T-t)^a$ , for  $T-1 < t < T$ . For  $t = T-1$  we get  $x_1(T-1) = x_2(T-1) = a-1$ . The line  $x_1 = x_2 = a-1$  is called, for varying  $a$ , the switching line.  $\square$

GENERAL FORMULATION 245 Since we only consider the "south-east" region of the state space, the trajectory  $(x_1(t), x_2(t))$  should move into this region from  $x_1(T) = x_2(T) = a$  if  $t$  goes backwards in time, which is only true for  $a < 1$ . For  $a > 1$ , the trajectories first move into the region  $x_2 > x_1$ , in retrograde time. Those trajectories cannot be optimal, since we only consider the region  $x_2 < x_1$ , and, moreover, boundary conditions (5.50) are not valid for the region  $x_2 > x_1$ . Hence only points  $x_1(t) = x_2(T) = a$  with  $a < 1$  can be terminal points. The half line  $x_1 = x_2 = a$ ,  $a < 1$ , is called the usable part (UP) of the terminal manifold or target set. Integrating the state equations backwards in time from the UP, and substituting  $t = T - 1$ , we observe that only the line segment  $x_1 = x_2 = a - 1$ ,  $a < 1$ , is a switching line. Using the points on this switching line as new terminal state conditions, we can integrate the state equations further backwards in time with  $u^*(t) = -1$ . The complete picture with the optimal trajectories is shown in Fig. 5.1. These solutions are the only ones which satisfy the necessary conditions of Thm. 5.4 and therefore they must be optimal, provided that an optimal solution exists. Figure 5.1: The solution of Example 5.2. The dotted lines denote lines of "equal time-to-go". Once the optimal trajectories are known, it is not difficult to construct  $V(x)$ . Curves of "equal time-to-go", also called isochrones, have been indicated in Fig. 5.1. Such a curve is determined from the equation  $V(x) = \text{constant}$ . It is

seen that along the curve AB the value function  $V(x)$  is discontinuous. Curves along which  $V(x)$  is discontinuous are called barriers. The reason behind the usage of such a terminology is that no optimal trajectory can ever cross a curve along which  $V(x)$  is discontinuous in the direction from large  $V$  to small  $V$ . Note that, because of the discontinuity in  $V(x)$ , Thm. 5.3 cannot be used in

246 T. BAŞAR AND G. J. OLSDER this case. The corresponding feedback strategy  $j^*(t,x)$ , which we write as  $j^*(x)$  since the system is autonomous (i.e., time-invariant), can readily be obtained from Fig. 5.1: to the right of the switching line and barrier, we have  $j^*(x) = -1$ , and to the left of the switching line and barrier, and also on the barrier, we have  $j^*(x) = +1$ . Discrete-time systems We now conclude this subsection by stating a counterpart of Thm. 5.4 for discrete-time optimal control problems. Such problems have earlier been formulated by (5.18a) and (5.18b), but now we also assume that  $f_k$  is continuously differentiable in  $x_k$ , and  $g_k$  is continuously differentiable in  $x_k$  and  $x_{j+i}$  ( $f_k \in K$ ). Then, the following theorem, whose derivation can be found in either Canon et al. (1970) or Boltyanski (1978), provides a minimum principle for such systems. Theorem 5.5 For the discrete-time optimal control problem described by (5.18a) and (5.18b), let (i)  $f_k(\cdot; u_k)$  be continuously differentiable on  $R^n$ , ( $k \in K$ ), ( $H$ )  $H_k(p_{k+1}, u_k, x_k)$  be continuously differentiable on  $R^n \times R^n$ , ( $k \in K$ ), ( $H_i$ )  $f_k(\cdot, 0)$  be convex on  $R^n \times R^m$ , ( $k \in K$ ). Then, if  $\{u_k, k \in K\}$  denotes an optimal control sequence, and  $\{x_{j+1}, k \in K\}$  denotes the corresponding state trajectory, there exists a finite sequence  $\{p_2, \dots, p_{k+i}\}$  of  $n$ -dimensional costate vectors so that the following relations are satisfied:  $x_{k+1} \sim f_k(x_k, u_k)$ ,  $u_k^* = \arg \min_{u_k \in U_k} H_k(p_{k+1}, u_k, x_k)$ ,  $p_{k+1} = -\lambda_{k+1} f_k(x_k, u_k) + \lambda_k g_k(x_{k+1}, u_k, x_k)$  (5.52)  $p_{k+i} = 0$ ,  $H_k(p_{k+i}, x_k, u_k) = g_k\{f_k(x_k, u_k), u_k, x_k\} + p_{k+1} f_k(x_k, u_k)$ ,  $k \in K$ . The reader should now note that Prop. 5.1 also follows as a special case of this theorem.

GENERAL FORMULATION 247 5.6 Representations of Strategies Along Trajectories, and Time Consistency of Optimal Policies In our discussion of some important results of optimal control theory in the previous section, we have confined our treatment to two specific information patterns and thereby to two types of strategies; namely, (i) the open-loop information pattern dictates strategies that depend only on the initial state and time, (ii) the feedback information pattern forces the permissible strategies to depend only on the

current value of the state and time. The former class of strategies is known as open-loop strategies, while the latter class is referred to as feedback strategies. Referring back to the discrete-time affine-quadratic optimal control problem described by (5.20a)-(5.20b) (and assuming that  $c_k = 0$  for all  $k \in K$ , for the sake of brevity in the discussion to follow), the solution presented in Prop. 5.1, i.e.,  $u^k = -P_k S_{k+1} A_k x^k$ ,  $k \in K$ , (5.53) is a feedback strategy, and in fact it constitutes the unique solution of that optimal control problem within the class of feedback strategies. If the optimal control problem had been formulated under the open-loop information pattern, then (5.53) would clearly not be a solution candidate. However, the unique optimal open-loop strategy, for this problem, can readily be obtained from (5.53) as  $u^k = -P_k S_{k+1} A_k x^k$ ,  $k \in K$ , (5.54a) where  $x^k$  is the value of the state at stage  $k$ , obtained by substitution of (5.53) in the state equation; in other words,  $x^k$  can be solved recursively from the difference equation  $x^{k+i} = (A_k - B_k P_k S_{k+i} A_k) x^k$ ,  $x^0 = x$ . (5.54b) It should be noted that (5.54a) is an open-loop strategy since it only depends on the discrete-time parameter  $k$  and the initial state  $x_0$  (through (5.54b)). We can also interpret (5.54a) as the realization of the feedback strategy (5.53) on the control set  $U_k$ , that is to say, (5.54a) is the action dictated by (5.53). For this reason, we also sometimes refer to (5.54a) as the open-loop realization or open-loop value of the feedback strategy (5.53). The question now is whether we can obtain other solutions (optimal strategies) for the optimal control problem under consideration by extending the strategy space to conform with some other information structure, say the closed-loop perfect state pattern. If we denote by  $T_k$  the class of all permissible closed-loop strategies (under the CLPS pattern) at stage  $k$  (i.e., strategies of the form  $u^k = f^k(x^k, x^k_{-1}, \dots, x^k_0)$ ), then (5.53) is clearly an element of that space, and it

248 T. BAŞAR AND G. J. OLSDER constitutes an optimal solution to the optimal control problem (5.20a)-(5.20b) also on that enlarged strategy space. As a matter of fact, any element  $y^k$  of  $T_k$  ( $k \in K$ ), which satisfies the boundary condition  $l^k(x^k, x^k_{-1}, \dots, x^k_0) = -P_k S_{k+1} A_k x^k$ ,  $f^k \in K$ , (5.55) also constitutes an optimal solution. One such optimal policy is the one-step memory strategy  $l^k\{x^k, z^k_{-i}\} = -P_k S_{k+i} A_k x^k + C_k [x^k - F_{k-i} X_{k-i}]$ , where  $F_k = A_k - B_k P_k S_{k+1} A_k$ , and  $C_k$  is a matrix of appropriate dimensions. It should be noted that all such strategies have the same open-loop value (which is the RHS of (5.55)) and generate the same state trajectory (5.54b), both of which follow from (5.55). Because of these properties, we call them different

representations of a single strategy, a precise definition of which is given below. Since the foregoing analysis also applies to continuous-time systems, this definition covers both discrete-time and continuous-time systems.

**Definition 5.11** For a control problem with a strategy space  $T$ , a strategy  $\gamma \in T$  is said to be a representation of another strategy  $\gamma' \in T$  if (i) they both generate the same unique state trajectory, and (ii) they both have the same open-loop value on this trajectory. **Remark 5.10** The notion of representation of a strategy, as introduced above, involves no optimality. As a matter of fact, it is a property of the strategy space itself, together with the control system for which it is defined, and no cost functional has to be introduced.  $\square$  A significance of the notion of "representations" in control problems is that it enables one to construct equivalence classes (of equal open-loop value control laws) in the general class of closed-loop strategies, so that there is essentially no difference between the elements of each equivalence class. In fact, for an optimal control problem, we can only talk about a unique optimal "equivalence class", instead of a unique optimal "strategy", since as we have discussed earlier within the context of the optimal control problem (5.20a)-(5.20b), every representation of the feedback strategy (5.53) constitutes an optimal solution to the optimal control problem, and there are uncountably many such (optimal) strategies. As a converse to the italicized statement, we can say that every solution of the optimal control problem (5.20a)-(5.20b) is a representation of the feedback strategy (5.53). Both of these statements readily follow from Def. 5.1, and they are valid not only for the specific optimal control problem (5.20a)-(5.20b) but for the general class of control problems defined in either discrete or continuous time.

**GENERAL FORMULATION 249** We now address the problem of constructing equivalence classes of representations of strategies for a control problem with a given space ( $T$ ) of closed-loop strategies, and first for discrete-time systems. The procedure, which follows directly from Def. 5.11, is as follows. For such a system, first determine the set ( $T_0$ ) of all elements of  $T$ , which are strategies that depend only on the initial state  $x_0$ , and the discrete time parameter  $k$ . The set  $T_0$ , thus constructed, is the class of all permissible open-loop controls in  $T$ . Now, let  $\gamma = \{u(k; x_0); k \in K\}$  be a chosen element of  $T_0$ , which generates (by substitution into (5.18a)) a (necessarily) unique trajectory  $\{x^*(k); k \in K\}$ . Then, consider all elements  $\gamma' = \{u'(k; x_0); k \in K\}$  of  $T$  with the properties (i)  $\gamma'$  generates the same state trajectory as  $\gamma$ , and (ii)

$\gamma(x_0, u_1, \dots, u_n) = \gamma(x_0, k) \in K$ . The subset of  $V$  thus constructed constitutes an equivalence class of representations which, in this case, have the open-loop value  $\gamma$ . If this procedure is executed for every element of  $\mathcal{L}_1$  then the construction of all equivalence classes of  $T$  becomes complete. For continuous-time systems, essentially the same procedure can be followed to construct equivalence classes of representations of strategies. In both discrete-time and continuous-time infinite dynamic systems such a construction leads, in general, to an uncountable number of elements in each equivalence class which necessarily contains one, and only one, open-loop strategy. The main reason for the "non-denumerable" property of equivalence classes is that in deterministic problems the CLPS information pattern (which involves memory) exhibits redundancy in information—thus giving rise to existence of a plethora of different representations of the same open-loop policy. This "informational nonuniqueness" property of the CLPS pattern is of no real concern to us in optimal control problems, since every representation of an optimal control strategy also constitutes an optimal solution. However, in dynamic games, the concept of "representations" plays a crucial role in the characterization of noncooperative equilibria, and the preceding italicized statements are no longer valid in a game situation. We defer details of an investigation in that direction to Chapters 6 and 7, and only provide here an extension of Def. 5.11 to dynamic games.

**Definition 5.12** For an  $N$ -person dynamic game with strategy spaces  $\{P_i; i \in N\}$  let the strategies of all players, except  $P_i$ , be fixed at  $\gamma_j \in P_j, j \in N, j \neq i$ . Then, a strategy  $\gamma_i \in P_i$  for  $P_i$  is a representation of another strategy  $\gamma_i^* \in P_i$ , with  $\gamma_j \in P_j (j \in N, j \neq i)$  fixed, if (i) the  $N$ -tuples  $\{\gamma_i, \gamma_j \in P_j, j \neq i\}$  and  $\{\gamma_i^*, \gamma_j \in P_j, j \neq i\}$  generate the same unique state trajectory, and (ii)  $\gamma_i$  and  $\gamma_i^*$  have the same open-loop value on this trajectory.

Time consistency In (one-player) optimal control problems, even though all representations of an optimal control strategy constitute optimal solutions, as discussed above,

250 T. BAŞAR AND G. J. OLSDER some of these representations may be preferred over others due to additional appealing properties they carry, such as robustness to modeling and decision errors. One such refinement scheme that brings in a selection among different representations of a given optimal strategy is provided by the notion of time consistency, which is valid not only for one-player dynamic decision problems (such as optimal control), but also for dynamic games in both discrete and continuous time. This notion manifests

itself in two forms—as weak and strong time consistency, which we introduce below in the general framework of dynamic games. Toward that end, let us first introduce the notation  $\Gamma(r; [0, T]; \text{sol})$  (5.56) to denote an  $i$ -person dynamic game, where  $T$  is the product strategy space,  $[0, T]$  is the decision interval,<sup>43</sup> and  $\text{sol}$  stands for any particular solution concept according to which (optimal or equilibrium) policies are determined.<sup>44</sup> Furthermore, let  $\Gamma_M(r; [s, t])$  (5.57) denote, respectively, the truncations of  $\Gamma$  to the time interval  $[s, t] \subset [0, T]$ , and let  $D[s, t] = D(\Gamma; [s, t]; \text{sol})$  (5.58) denote a version of  $D(\Gamma; [0, T]; \text{sol})$ , where the policies of all players in the intervals  $[0, s)$  and  $(t, T]$  are fixed as  $\{f_i^*, f_i^T, i \in N\}$ . An underlying assumption here is that the cost functions of the players are of additive type, such as (5.10) or (5.14). Then, we have the following two definitions on the "time consistency" of a solution obtained for the original game according to the concept "sol".

**Definition 5.13** An  $N$ -tuple of policies  $\{f_i^*, i \in N\}$  solving the dynamic game  $D(\Gamma; [0, T]; \text{sol})$  is weakly time consistent (WTC) if its truncation to the interval  $[s, T]$ ,  $\{f_i^*, i \in N\}$ , solves the truncated game  $D[s, T]$ , this being so for all  $s \in (0, T)$ . If a solution  $\{f_i^*, i \in N\}$  is not WTC, then it is time inconsistent.

**Definition 5.14** An  $N$ -tuple of policies  $\{f_i^*, i \in N\}$  solving the dynamic game  $D(\Gamma; [0, T]; \text{sol})$  is strongly time consistent (STC) if its truncation to the interval  $[s, t]$ ,  $\{f_i^*, i \in N\}$ , solves the truncated game  $D[s, t]$ , for every  $\{f_i^*, i \in N\} \in F[s, t]$ , this being so for every  $s \in (0, T)$ . Both these refinements essentially say that for a solution to be time consistent (in the weak or strong sense), the players should have no rational reason, <sup>43</sup>By an abuse of notation, we will let this also denote the discrete interval for the discrete-time problem, in which case  $[0, T] = \{1, \dots, K\}$ . <sup>44</sup>In an optimal control problem, for example,  $\text{sol}$  could stand for minimization or maximization, or reachability to a target set; in a dynamic game it could stand for saddle-point, Nash or Stackelberg equilibrium.

**GENERAL FORMULATION** 251 at any future stage of the game, to deviate from the adopted policies. The difference between weak and strong time consistency lies in the consistency (or inconsistency) of the past actions with the adopted policies. In the former (that is, weak time consistency) there is no ground for renegeing at future stages only if the past actions are consistent with the original solution,  $\{f_i^*, i \in N\}$ ; whereas in the latter this is true even if there have been deviations in the past from the actions which are dictated by the original solution. Solutions of optimal control problems of the types discussed in

Section 5.5 (and of course also their representations) are weakly time consistent—a fact that easily follows from Def. 5.13. Not every solution is strongly time consistent, however; a case in point being the optimal open-loop control  $U^* = \{u^*(t), t \in [0, T]\}$  obtained using the minimum principle (cf. subsection 5.5.3). If a non-optimal open-loop policy  $U$  is applied to the system in the time interval  $[0, s)$ , which brings the system to a state, say  $x(s)$ , that is not on the optimal trajectory, then the control  $u^*$  will not necessarily be optimal for the new truncated control problem  $D^s$ . The reason for this is that  $u^*$  will no longer satisfy the necessary conditions given in Thm. 5.4 or Thm. 5.5 on the interval  $[s, T]$ , as the initial state at  $t = s$  is not on the optimal trajectory. Hence, optimal open-loop control is only WTC. The optimal feedback control, obtained from the dynamic programming equation (5.19) or (5.25), however, is STC—a fact that follows from Def. 5.14. For the discrete-time linear-quadratic problem (with  $C_{fe} = 0$ ), for example, the policy (5.53), which is obtained from a dynamic programming equation, is STC, whereas its open-loop version (5.54a) is only WTC. For the more general affine-quadratic problem described by (5.20a), the optimal feedback policy (5.21a) is STC. In a way, an STC policy is one that is not only independent of the initial state  $x_0$  but also independent of other past values of the state. Not every optimal policy that depends only on the current value of the state is necessarily STC, however, especially if the initial state is a fixed known quantity. To make the point here, consider the optimal feedback policy  $u^* = \{u^k, k \in K\}$  obtained from (5.19), for a discrete-time system whose initial state  $x_0$  is fixed and given, under which the optimal trajectory,  $\{x^k, k \in K\}$ , is just a sequence of known vectors. Now, if  $T$  is a policy space compatible with the feedback (FB) information pattern, and we seek an optimal solution in the class  $Y$ , then not only  $u^*$ , but every  $u \in T$  given by  $u^k(x^k) = u^*(x^k, x_0)$ ,  $k \in K$ , will be an optimal policy, where  $V^k$  is any function satisfying the boundary condition  $V^k(x^k, x_0) = V^k(x^k)$ ,  $k \in K$ . Different choices of  $V^k$  lead to different representations of the same (optimal) policy on the optimal trajectory, and they all (in this case) depend only on the current value of the state. Because of this feature, one may be tempted to call all these different representations "optimal feedback policies", but this would not be correct because only one of them, namely  $u^*$ , is STC. The others cannot be obtained directly from (5.19) and do not meet the "permanent optimality" property associated with an STC optimal solution. Here, the feedback control is obtained by minimization of the right-hand side of (5.19) or (5.25), depending on whether we have a discrete-time

or a continuous-time problem.

252 T. BAŞAR AND G. J. OLSDER There are two points worth mentioning here in connection with the notion of an STC optimal policy. The first one is that it is quite possible for a constant policy to be STC, which would arise if the optimization problem in (5.19) (or (5.25)) admits a solution which is independent of  $x$  for every  $t$  (or  $k$ ). In such a case, the optimal feedback and the optimal open-loop solutions would coincide, and hence the open-loop solution would also be STC. The second point is that a given dynamic optimization problem may not admit an STC optimal solution, even under the CL information pattern. This would arise if the policy space has additional structural restrictions imposed on the permissible policies. For example, one may require the policies to depend only linearly on the current value of the state, with no such restriction imposed on their dependence on past history. If the unique optimal FB solution obtained from (5.19) (or (5.25)), i.e.,  $u^*$ , is not linear in its argument, then the problem would not admit an STC solution. However, the problem would still admit a WTC optimal solution, such as (in the discrete-time case) any policy of the form  $u_k = M_k(x_k, x^k) - M_k x_k + \tilde{f}_k(x^k) \sim M_k x_k$ ,  $u_k \in U$ , where  $\{M_k\}$  is an arbitrary matrix sequence (of compatible dimensions) and  $\{x^k\}$  is the trajectory generated by  $u^*$ , in general as a function of  $x_0$ . The notions of weak and strong time consistency are of course also valid in systems which incorporate lag variables, either in the state or in the control. Consider, for example, the system where the evolution is governed by (as a counterpart of (5.18a)):  $x_{k+1} = f_k(x_k, x_{k-1}, u_k)$ ,  $x_0$  and  $x_{-1}$  given. This could be converted into the form (5.18a) by introducing a new state vector  $\tilde{x}_k$  which corresponds to the pair  $(x_k, x_{k-1})$ , thus making our earlier discussion apply here, with  $x_k$  replaced by  $\tilde{x}_k$ . Note that, in view of this, an STC solution will exist for the discrete-time optimal control problem (under the cost function (5.18a)) only if one allows the permissible policies at time  $t$  to depend not only on  $x_t$  but also on  $x_{t-1}$ . This can best be seen if one writes down the counterpart of the dynamic programming (DP) equation (5.19) in terms of the original variables:  $V(k, x_k, x_{k-1}) = \min_{u \in U_k} \{V(k+1, f_k(x_k, x_{k-1}, u), x_k) + g_k(x_k, x_{k-1}, u)\}$  with the boundary condition being  $V(K+1, x_{K+1}, x_K) = 0$ . Clearly, an STC solution will have to be of the form  $u_k = \tilde{f}_k(x_k, x_{k-1})$ . If the information pattern is FB, then the problem may not admit a solution at all. This concludes our discussion on WTC and STC optimal solutions in one-person dynamic decision problems. The topic will be revisited later in



Chapters 6 and 7, in the context of saddle-point, Nash and Stackelberg equilibria in dynamic games. Stochastic systems We devote the remaining portion of this section to a brief discussion on the representation of closed-loop strategies in stochastic control problems. In par-

GENERAL FORMULATION 253 ticular, we consider, as a special case of Def. 5.4, the class of control problems described by the state equation  $X_{k+1} = f_k(x_k, u_k) + e_k$  (5.59) where  $f_k$  ( $k \in K$ ) is a random vector taking values in  $R^n$ , and the joint probability distribution function of  $\{x_0, \theta_1, \dots, \theta_k\}$  is known. Then, given a feedback strategy  $\{j_k(\cdot), k \in K\}$ , it is in general not possible to write down a representation of this strategy that also involves past values of the state, since now there is no redundancy in the CLPS information, due to existence of the disturbance vector  $\theta_k$ . In other words, while in deterministic systems  $X_{k+i}$  can easily be recovered from  $X_k$  if one knows the strategy employed, this is no longer possible in stochastic systems because  $x_{k+i}$  also contains some information concerning  $\theta_k$ . This is particularly true if the random vectors  $x_0, \theta_1, \dots, \theta_k$  are Gaussian, mutually independent and  $\text{cov}(\theta_k) > 0$  for all  $k \in K$ . Hence, we have the following proposition. Proposition 5.5 //, for the stochastic system (5.59), all random vectors are Gaussian, mutually independent and the covariance of  $\theta_k$  is of full rank for all  $k \in K$ , then every equivalence class of representations is a singleton. Remark 5.11 The reason why we do not allow some random vectors to be statistically dependent and/or  $\text{cov}(\theta_k)$  to have lower rank is because then some components of  $X_{k+i}$  can be expressed in terms of some components of  $X_k, x_{k-i}$ , etc., independent of  $\theta_k, \theta_{k-i}, \dots$ , which clearly gives rise to nonunique representations of a strategy. Furthermore, we should note that the requirement of having a Gaussian probability distribution with positive definite covariance can be replaced by having any probability distribution that assigns positive probability mass to every nonempty open subset of  $R^n$ . The latter requirement is important because, if any open subset of  $R^n$  receives zero probability, then nonunique representations are possible on that particular open set.  $\square$  Optimal control of stochastic systems described by (5.59) can be performed by following the basic steps of dynamic programming applied to the deterministic system (5.18a)-(5.18b). Toward this end, we first replace the cost functional (5.18b) by  $J(l) = E \sum_{k=0}^{K-1} g_k(x_{k+1}, u_k, X_k) - \gamma_k(x_i, l < k), k \in K$ , (5.60) where  $\{f_k, j_k \in K\}$  is any permissible closed-loop strategy, and  $E[\cdot]$  denotes the expectation operation defined with respect to the underlying statistics of  $X_0, \theta_1, \dots, \theta_k$  which

we assume to be mutually independent. Then, the value function  $V(k, x)$  is defined by  $V(k, x) = \min_{u_k \in U_k} E \{ \sum_{i=0}^{K-k-1} g(x_{k+i}, u_{k+i}) + V(k+i, x_{k+i}) \mid x_k = x \}$  (5.61a) where  $g(x_k, u_k)$  stands for the cost function at time  $k$ , i.e.,  $E[(x_k - E[x_k])^2]$ .

254 T. BAŞAR AND G. J. OLSDER Following the steps of dynamic programming, and by utilizing the assumption of independence of the random vectors involved, it can readily be shown that the value function satisfies the recursive relation  $V(k, x) = \min_{u_k \in U_k} \{ E \{ g(x_k, u_k) + V(k+1, x_{k+1}) \mid x_k = x, u_k = u \} \}$  (5.61b) which also leads to the conclusion that the optimal strategy  $u^*(k \in K)$  depends only on the current value of the state, i.e., is a feedback strategy. This is also a strongly time consistent solution under the CLPS information pattern, where the notion (introduced by Def. 5.14) applies equally to stochastic dynamic decision problems, since its definition used strategies, and not actions. We now apply this result to the so-called linear-quadratic discrete-time stochastic optimal control problem described by the linear state equation  $x_{k+1} = A_k x_k + B_k u_k + \delta_k$  (5.62a) and the quadratic cost functional  $J(u) = E \{ \sum_{k=0}^{K-1} x_k^T Q_k x_k + u_k^T R_k u_k \mid x_0 = x_0 \}$  (5.62b) where  $R_k > 0$ ,  $Q_k \geq 0$ ;  $(k \in K)$ ;  $\delta_k$  ( $k \in K$ ) are  $n$ -dimensional mutually independent Gaussian vectors with mean value  $c_k$ , and covariance  $A_k > 0$ , and  $u^*(k \in K)$  are permissible closed-loop strategies. For this special case,  $V(k, x)$  again turns out to be a general quadratic function in  $x$ , and this leads to the following counterpart of Prop. 5.1. Proposition 5.6 For the linear-quadratic stochastic control problem described by (5.62a)-(5.62b), and under the closed-loop perfect state information pattern, there exists a unique minimizing solution (which is a feedback strategy) given by  $u_k = -R_k^{-1} B_k^T (P_k x_k + s_k)$  (5.63) where all terms are defined in Prop. 5.1. Remark 5.12 The solution of the stochastic linear-quadratic optimal control problem depends only on the mean value ( $c_k, k \in K$ ) of the additive disturbance term but not on its covariance ( $A_k, k \in K$ ). Moreover, the solution coincides with the optimal feedback solution of the deterministic version of the problem (as presented in Prop. 5.1), which takes the mean value of the Gaussian disturbance term as a constant known input. Finally, in view of Remark 5.11, uniqueness of the solution is valid under the assumption that the covariance matrix  $A_k$  is positive definite for all  $k \in K$ .  $E_k$  denotes the expectation operation with respect to statistics of  $\delta_k$ . The minimum expected value of the cost function does, however, depend on

$\{A^*, f_c \in K\}$ .

GENERAL FORMULATION 255 Remark 5.13 Optimal solutions to stochastic control problems can be obtained under other deterministic information patterns also, but they are not representations of each other under the full-rank condition of the independent additive noise. Hence, each different information pattern requires a different derivation for the optimal solution. These solutions are not all STC, but they are necessarily WTC. See Problem 5 in Section 5.8 for an illustrative example.

5.7 Viscosity Solutions In this section another solution approach to the HJB equation (5.25) will be given. This approach has its origin in the theory of stochastic optimal control theory (in continuous time) as will be explained soon. If in (5.25), or in its time-invariant analogue  $f(x,u) + g(x,u) = 0$ , (5.64) the minimizing  $u$ , as a function of  $dV/dx$ ,  $x$  (and  $t$  for (5.25)), is substituted into the same expressions, the resulting equations will be written formally as  $dV - dV - H(x, V, t) = 0$ , (5.65)  $dV - H(x, V) = 0$ , (5.66) respectively. For both equations we assume appropriate boundary conditions. Sometimes finding a solution of (5.65) is referred to as a Cauchy problem and of (5.66) as a Dirichlet problem. Note that, strictly speaking, the function  $H$  in (5.65) and (5.66) does not directly depend on the argument  $V$  in our case and that the format of (5.65) and (5.66) is more general than strictly necessary. Rather than studying (5.65) it pays sometimes to study  $v - \epsilon \Delta^2 V - dV - \epsilon \Delta V = 0$  (5.67) where  $\epsilon$  is a (small) positive parameter. The first term of this expression is also written compactly as  $\epsilon \Delta V$ . In the same way one could add the same term with the second derivatives to (5.66):  $v - \epsilon \Delta^2 V - dV$  The theory of partial differential equations tells us that existence and unicity results are easier to obtain for (5.67) than for (5.65). One may hope to get results

256 T. BAŞAR AND G. J. OLSDER for (5.67) and then study their limits as  $\epsilon$  tends to zero. Hopefully the solution of (5.67) (respectively, (5.68)) will resemble a solution of (5.65) (respectively, (5.66)) in some sense if  $\epsilon \rightarrow 0$ . Quite often solutions to (5.65), (5.66), viewed as abstract mathematical equations, are not unique, among others depending on the definition of solution that one adopts. With the addition of a proper understanding of the underlying optimal control problem, the solution becomes usually unique. Before continuing along these lines, we will first indicate that (5.67) is not only an auxiliary equation, but that it represents a condition for the value function of a

related stochastic optimal control problem. Consider the one-person version of the AT-person stochastic differential equation as described by (5.17b):  $dx_t = F(t, x_t, u(t))dt + \Sigma(t, x_t)dW_t$ ,  $x_t|_{t=0} = x_0$ , (5.69) and the cost functional  $J(u) = E \int_0^T g(t, x_t, u(t))dt + q(x_T)$  (5.70) where  $E$  stands for the expectation operator (see Appendix B) and where  $u\{t\} = u(t, x_t) \in S$ , i.e., only feedback strategies are allowed. The following result from Fleming (1969) (see also Fleming and Rishel (1975)) provides a set of sufficient conditions and a characterization for the optimal solution of this stochastic control problem. Proposition 5.7 Let there exist a suitably smooth solution  $W(t, x)$  of the semi-linear parabolic partial differential equation  $dW/dt + \sum_{i,j} a_{ij} \frac{\partial^2 W}{\partial x_i \partial x_j} + m[VxW - F(t, x, u) + g(t, x, u)] = 0$ ;  $V^* ; W(T, x) = q(T, x)$ , where  $a_{ij}$  denotes the  $ij$ th element of the matrix  $\Sigma\Sigma^*$ , and  $x^i$  denotes the  $i$ th component of the vector  $x$ . If there is a function  $u^o(t, x, p)$  that minimizes  $\{p^*F(t, x, u) + g(t, x, u)\}$  on  $S$ , then  $u^*(t, x) = u^o(t, x, VxW)$  is an optimal control, with the corresponding expected cost given by  $J^* = W(0, x_0)$ . This proposition yields a partial differential equation which the value function of the stochastic optimal control problem should satisfy. The resemblance between (5.67) and (5.71) will be obvious. Thus (5.67), for the time-invariant case, can also be considered as a condition which the value function of an appropriately defined stochastic optimal control problem should satisfy. The formulation of the stochastic optimal control problem and the statement of the proposition have been rather informal. The rigorous underlying mathematics is beyond the scope of this book and can for instance be found in the references already mentioned and in Lions (1982) and in Crandall and Lions (1983).

GENERAL FORMULATION 257 We now return to (5.67) and will study its solution when  $\epsilon \downarrow 0$ . To express the fact that the solution of (5.67) depends on  $\epsilon$ , it will be indicated by  $V^\epsilon$ . If  $\lim_{\epsilon \downarrow 0} V^\epsilon(x)$  exists, this limit might be what is called a viscosity solution of (5.66). The name "viscosity solution" comes from a celebrated method in fluid dynamics where the term  $\epsilon \Delta u + g(x, u)$  represents physically a viscosity. In optimal control theory, however, the term "viscosity solution" is reserved for a particular solution of (5.66), which may equal the limit solution (with  $\epsilon \downarrow 0$ ) of (5.68) if it exists. More properly, it then could have been called the vanishing viscosity solution. Definition 5.15 Suppose that  $H$  is continuous in its arguments. A continuous function  $V(x)$  is called a viscosity subsolution of (5.66) provided that for all continuously differentiable  $\phi(x)$ , which map  $R^n$  or an appropriate subset of  $R^n$  into  $R$  (as

does  $V(x)$ ), the following holds: if  $V \in C^2$  attains a local maximum at  $x_0$ , then  $H(x_0, V(x_0), \Delta) \geq 0$ . (5.72) A continuous function  $V(x)$  is called a viscosity supersolution of (5.66) provided that for all continuously differentiable  $\phi(x)$ , which map  $\mathbb{R}^n$  or an appropriate subset of  $\mathbb{R}^n$  into  $\mathbb{R}$  (as does  $V(x)$ ), the following holds: if  $V - \phi$  attains a local minimum at  $x_0$ , then  $H(x_0, V(x_0), \Delta) \leq 0$ . (5.73) A continuous function  $V(x)$  is called a viscosity solution of (5.66) if it is both a viscosity subsolution and a viscosity supersolution. An equivalent, somewhat more intuitive, definition of a viscosity subsolution is as follows. A continuous function  $V(x)$  is a viscosity subsolution of (5.66) if  $H(x, V(x), p) \geq 0$  for all  $p \in \mathbb{R}^n$  such that  $\liminf_{x \rightarrow x_0} \frac{V(x) - V(x_0) - p \cdot (x - x_0)}{|x - x_0|} > 0$ . Those vectors  $p$  which satisfy this latter equation are sometimes referred to as subdifferentials. A similar and somewhat more intuitive definition of a viscosity supersolution can also be given. The equivalences of these definitions have been shown in Crandall, Evans and Lions (1984). Remark 5.14 It should be noted that, in the definitions of viscosity sub- and supersolutions, in the references cited the inequalities (5.72) and (5.73) are both reversed in sign. The definitions given here, however, are consistent with (5.68), in which it is assumed that  $\lambda$  is positive. In the references cited, the second-order term in (5.68) is preceded by a minus sign. We have not adopted the addition of this minus sign because of the fact that the second-order term is directly

258 T. BAŞAR AND G. J. OLSDER related to the covariance of the noise in the system equations (and hence has to be nonnegative). Another way to make the definitions given here and in the references cited fully equivalent is to change  $H$  into  $-H$ .  $\square$  The following simple result establishes the consistency of the notions of viscosity and "classical solutions". Theorem 5.6 If the continuously differentiable function  $V$  is a classical solution of (5.66), then it is a viscosity solution. The following theorem relates the solutions of (5.66) and (5.67). Theorem 5.7 Suppose that  $V_\epsilon$  is a continuously differentiable solution of (5.67) and that it converges uniformly, as  $\epsilon \rightarrow 0$ , to some continuous function  $V$ . Then  $V$  is a viscosity solution of (5.66). Proof. Let us check (5.72) first for twice continuously differentiable  $\phi$ . Assuming that  $V \in C^2$  has a local maximum at  $x_0$ , a function  $\xi$  is chosen, differentiable a sufficient number of times and with  $0 < \xi < 1$  for  $x \in \mathbb{R}^n$  and  $\xi(x_0) = 1$ , such that  $V - (\phi - \xi Q)$  has a strict local maximum at  $x_0$ . Thus for  $\xi$  sufficiently small,  $V_\epsilon - (\phi - \xi Q)$  has a local maximum at some  $x_\epsilon$  and  $x_\epsilon \rightarrow x_0$  as  $\epsilon \rightarrow 0$ . Then,  $V_\epsilon(x_\epsilon) = V(\xi(x_\epsilon) -$

$C(x, e), A(x, e) < A(x, 0) - C(x, e)$ , and therefore  $H(x, e, V(x, e), V(x, 0)) = -eA(x, e) > -eA(x, 0) - Q(x, e)$ . Now, if we take the limit  $e \rightarrow 0$ , this expression becomes (5.72). Suppose now that  $V$  is only once continuously differentiable. The first part of the proof, as just given, is now repeated for  $V \in C^1$ , where  $V$  is twice continuously differentiable and  $V \in C^1 \rightarrow C^2$  as  $\epsilon \rightarrow 0$ . This concludes the part of the proof that  $V$  is a viscosity subsolution. The fact that  $V$  is also a viscosity supersolution can analogously be shown, which then, together with  $V$  being a viscosity subsolution, concludes the proof. A weaker version of Thm. 5.7 exists in the sense that the function  $H$  may also depend on  $\epsilon$ , to be indicated by  $H_\epsilon$ . In the formulation of the extended version of Thm. 5.7 it is then necessary that  $H_\epsilon$  converges to  $H$  uniformly on compact subsets of  $\mathbb{R}^n \times \mathbb{R}^n$ , as  $\epsilon \rightarrow 0$ . This weaker version is useful, for instance, when the continuous-time formulation of the problem is approximated by a discretized one (for numerical purposes, for instance). Such a "discretized version" of  $H$  could then coincide with  $H_\epsilon$  (see Lions (1982), Bardi and Sartori (1992) and Pourtallier and Tolwinski (1992)). An issue not dealt with heretofore is that of uniqueness of viscosity solutions. Consider (5.67) again. For special forms of the function  $H$  the uniqueness has been shown (see Lions (1982) or Crandall, Evans and Lions (1984)). In those

GENERAL FORMULATION 259 cases it is essential that  $H$  depends explicitly on  $V$ , which is for instance the case if  $H$  has the following form:  $dV/dx H(x, V, -) = V + a(x) + \theta(x)$  for some scalar functions  $a$  and  $\theta$ . However, as remarked in the beginning of this section,  $H$  as introduced does not depend on  $V$ . It is, however, possible to create such a dependence by means of the so-called Kruzkov-transform of the value function  $V$ . This transform of the value function is defined as  $v(x) = \int_0^\infty t(V(x)) dt$  if  $V(x) < \infty$ ,  $v(x) = 1$  elsewhere, where  $\int_0^\infty t(V(x)) dt = 1 - \exp(-V)$ . It is important to realize that  $v$  is bounded and is itself the value function of an optimal control problem, viz.,  $v(x) = \inf_u \int_0^\infty (L(u)) dt$ ,  $u$  with  $L$  defined in (5.23). The transformed cost function  $\int_0^\infty (L(u)) dt$  can be rewritten in the "standard form" of integral and terminal part as  $\int_0^T (L(u)) dt + \exp(-q(T, x(T)))$ . As a side remark, this latter cost function resembles cost functions with discount factors, which have the form  $\int_0^T g(t, x, u) \exp\{-Xt\} dt + q\{T, x(T)\} \exp(-XT)$ , where the nonnegative constant  $A$  is the so-called discount factor. Such cost functions can be treated in the standard way but offer some advantages if  $T = \infty$  and the terminal term is not present. Under some mild conditions, see Bardi and

Soravia (1991), the following proposition can be shown to hold. Proposition 5.8 If  $V$  is a viscosity solution of (5.66), where  $H$  does not depend on the second argument  $V$ , and if there exists a constant  $g_0$  such that  $g(x, u) > g_0 > 0$ , then  $v$  is a viscosity solution of  $dv + H(x, v, \cdot) = 0$ , (5.74) where  $H(x, v, \cdot) = \min \{ g(x, u) - g_0, g(x, u) - (g(x, u) - g_0)v \}$ .

260 T. BAŞAR AND G. J. OLSDER This proposition shows how the original equation  $H(x, V, \cdot) = 0$ , where  $H$  does not depend on the value function  $V$  directly, can be transformed into (5.74), which has a linear term in the transformed value function  $v$ . For the latter kind of partial differential equations, it is easier to prove uniqueness of solution. It also offers a starting point for numerical procedures to solve for  $v$  (and hence for  $V$ ). Hitherto in this section, we have taken  $V$  to be continuous. For the case in which it is discontinuous the definition of viscosity solution can be modified as follows. Definition 5.16 Suppose that  $H$  is continuous in its arguments. An upper semicontinuous function  $V$  is called a viscosity subsolution of (5.66) provided that for all continuously differentiable  $\phi$ , which map  $\mathbb{R}^n$  or an appropriate subset of  $\mathbb{R}^n$  into  $\mathbb{R}$  (as does  $V$ ), the following holds: if  $V - \phi$  attains a local maximum at  $x_0$ , then  $H(x_0, V(x_0), \phi'(x_0)) \geq 0$ . (5.75) A lower semicontinuous function  $V$  is called a viscosity supersolution of (5.66) provided that for all continuously differentiable  $\phi$ , which map  $\mathbb{R}^n$  or an appropriate subset of  $\mathbb{R}^n$  into  $\mathbb{R}$  (as does  $V$ ), the following holds: if  $V - \phi$  attains a local minimum at  $x_0$ , then  $H(x_0, V(x_0), \phi'(x_0)) \leq 0$ . (5.76) A function  $V$  is called a viscosity solution of (5.66) if the upper and lower semicontinuous envelopes of  $V$ , defined, respectively, by  $V^*(x) = \limsup V(y)$ ,  $V_*(x) = \liminf V(y)$ , are respectively a viscosity subsolution and a viscosity supersolution. For value functions with discontinuities relatively few results are available as yet; see the notes section 5.9. 5.8 Problems 1. A company has  $x(t)$  customers at time  $t$  and makes a total profit of  $\int_0^t (cx(s) - u(s)) ds$ , up to time  $t$ , where  $c$  is a positive constant. The function  $u(t)$ , restricted by  $u(t) \geq 0$ ,  $f(t) \geq 0$ , represents the money put into advertising. The restriction  $f(t) \geq 0$  indicates that the company is not allowed to borrow

GENERAL FORMULATION 261 money. Advertisement helps to increase the number of customers according to  $\dot{x} = u$ ,  $x(0) = x_0 > 0$ . The company wants to maximize the total profit during a given time period  $[0, T]$ , where  $T$  is fixed. Obtain both the open-loop and feedback optimal solutions for this problem. 2.

Consider the two-dimensional system  $\dot{X} = -X^2 + u$ ,  $\dot{u} = -1 + Z^i u$ , where the control variable  $u$  is scalar and satisfies  $|u(t)| < 1$ . In polar coordinates, the equations of motion become  $\dot{r} = -\cos a$ ,  $\dot{a} = -ru + \sin a$ , where  $a$  is the angle measured clockwise from the positive  $X^2$ -axis. The target set  $A$  is  $|a| > a$ , where  $0 < a < \pi/2$ . (For an interpretation of this system see Example 8.3, later in Chapter 8, with  $V^2 = 0$ , or Lewin and Olsder (1979).) The decision maker wants the state  $(r, a)$  to stay away from  $A$  for as long as possible. Determine the time-optimal-control. Show that, for  $a > \pi/2$ , there is a barrier. 3. Consider the following discrete-time, discrete state space optimal control problem. The state space consists of the integers  $\{0, 1, 2, \dots\}$ . At time  $t_k$  the decision maker steers the system positioned at integer  $i > 1$ , either to  $\{0\}$  or to the integer  $\{i + 1\}$ , where it arrives at time  $t_{k+1}$ . The control problem terminates when the system is steered to the origin  $\{0\}$ . The pay-off is completely determined by the last move. If the system moves from  $\{i_k\}$  to  $\{0\}$ , then the pay-off is  $2i_k$ , to be maximized by the decision maker. Show that each playable pure strategy leads to a finite value, and that the behavioral strategy which dictates both alternatives with equal probability ( $1/2$ ) at each stage leads to an infinite expected pay-off (assuming that  $\gamma > 1$ ). 4. Consider the discrete-time system whose state evolves according to the dynamics  $X_{k+1} = f_k(x_k, U_k, U_{k-1})$ ,  $u_0 = 0$ ,  $f_k \in K$ . (i) Show that every STC optimal policy for this system, under the cost function (5.18b) and the information pattern that allows the control at stage  $k$  to depend on the current and past values of the state and past values of control, has to be in the form  $U_k = \gamma f_k(\cdot; x_k, u_{k-1})$ .

262 T. BAŞAR AND G. J. OLSDER (ii) Obtain the STC solution(s) explicitly for the linear-quadratic (LQ) problem where the cost is given by (5.20b), and the system dynamics (exhibiting lagged dependence on the control) by  $x_{k+1} = A_k x_k + B_k u_k + C_k u_{k-1}$ ,  $u_0 = 0$ ,  $k \in K$ . (iii) Does the LQ problem above admit (a) an STC optimal policy under the CLPS information pattern; (b) a WTC optimal policy under the CLPS information pattern? In the latter case obtain one such policy, if there exists one. 5. In Problem 4 above, replace the deterministic linear dynamics by the stochastic difference equation  $x_{k+1} = A_k x_k + B_k u_k + C_k u_{k-1} + 0_k$ ,  $u_0 = 0$ ,  $k \in K$ , where  $x \setminus 0, \dots, 0_k$  are Gaussian, mutually independent, and have zero mean and positive definite covariances. Further let the cost function be given by (5.62b). (i) Does the problem admit an STC optimal control under the CLPS information pattern? (ii) Obtain a WTC optimal policy for this stochastic control problem under the CLPS information



pattern. (iii) Can you generate an optimal OL control policy from the solution obtained above? Would it be WTC under the OL information pattern? 5.9 Notes Section 5.2. As already noted in Remark 5.2, the state-model description of a discrete-time (deterministic or stochastic) infinite dynamic game as presented here is clearly not the most general extension of Kuhn's finite game model (cf. Chapters 2 and 3) to the infinite case. One such extension which has not been covered in Remark 5.2 is the infinite-horizon problem wherein the number of stages  $K$  tends to infinity. In such games, the so-called stationary strategies (i.e., strategies which do not depend on the discrete-time parameter  $k$ ) become of real interest, and they have hitherto attracted considerable attention in the literature (see Shapley, 1953; Sobel, 1971; Maitra and Parthasarathy, 1970), but mostly for finite or at most denumerable state and strategy spaces. The general class of discrete-time stochastic dynamic games under the feedback information structure are also referred to as Markov games in the literature. For two survey articles on this topic, see Parthasarathy and Stern (1977) and Raghavan and Filar (1991). Section 5.3. Differential games were first introduced by Isaacs (1954-1956), within the framework of two-person zero-sum games, whose work culminated in the publication of his book about ten years later (see Isaacs, 1975). Nonzero-sum differential games were later introduced in the works of Starr and Ho (1969a,b), Case (1967) and

GENERAL FORMULATION 263 Friedman (1971), but under specific information structures—namely, the open-loop and feedback perfect state information patterns. Some representative references on stochastic differential games, on the other hand, are Friedman (1972), Elliott (1976) and Başar (1977b). The book by Krasovskii and Subbotin (1988) has an extensive discussions on non-Lipschitzian differential equations. A class of differential games not covered by Def. 5.5 are those whose state dynamics are governed by partial differential equations. These will not be treated in this book; for some examples of practical games which fit into this framework, the reader is referred to Roxin (1977). Yet another class of differential games not to be covered in subsequent chapters are those on which the players are allowed to employ impulsive controls; see Blaquiere (1977) for some results on this topic. Section 5.4. Mixed and behavioral strategies in infinite dynamic games were first introduced by Aumann (1964) who also provided a rigorous extension of Kuhn's extensive form description of finite games to infinite games, and proved in this context that if an  $JV$ -person nonzero-sum infinite

dynamic game is of "perfect recall" for one player, then that player can restrict himself to behavioral strategies. For a further discussion of mixed strategies in differential games and for an elaboration on the relation of this concept with that of "relaxed controls" used in control theory (cf. Warga, 1972), the reader is referred to Wilson (1977), who also proves the existence of mixed-strategy saddle points in suitably structured zero-sum differential games of fixed duration and with open-loop information structure. Some other references which deal with existence of mixed-strategy equilibrium solutions in differential games are Elliott (1976), Pachter and Yavin (1979), Levine (1979) and Kumar and Shiau (1981). Section 5.5. Some classic texts on the theory of optimal control where the reader can refer to for a more extensive coverage and references are Bryson and Ho (1975), Hermes and LaSalle (1969) and Fleming and Rishel (1975). Section 5.6. The concept of "representations of a strategy" was introduced by Başar in a series of papers, within the context of infinite dynamic games (see, e.g., Başar, 1974, 1976a, 1977b), and it was shown to be closely related to existence of nonunique Nash equilibria under the CLPS information pattern, as it will be elucidated in Chapter 6. The subsection on time consistency is based on the material in Başar (1989b), which the reader should refer to for extensive discussion on the motivation behind these notions, and some specific results. Also, Chapters 6 and 7 contain further discussion of time consistency in the context of Nash and Stackelberg equilibria. The issue of time consistency (or inconsistency) has pervaded the economics literature during the past two decades, following the stimulating paper by Kydland and Prescott (1977). Section 5.7. Viscosity solutions in the theory of optimal control (and also in the theory of zero-sum differential games) became well known with the publication of the book by Lions (1982). The theory presented in this section is based mainly on Crandall, Evans and Lions (1984). Another reference in the same vein is Crandall and Lions (1983). Viscosity solutions for minimum-time problems have been reported in Bardi (1989) and Staicu (1989). In all these references it was assumed that the value function  $V$  is continuous in its arguments, a property not often satisfied (see

264 T. BAŞAR AND G. J. OLSDER Example 5.2 and also Chapter 8). Recently progress has been made on problems with discontinuous value functions, notably for time-optimal control problems; see Ishii (1989), Bardi and Soravia (1991) and Bardi and Staicu (1991). The notion of discontinuous

viscosity solutions was introduced by Ishii (1987). For a more recent rigorous treatment of the topic of viscosity solutions in the context of optimal control, see Fleming and Soner (1993).

Chapter 6 Nash and Saddle-Point Equilibria of Infinite Dynamic Games 6.1  
Introduction This chapter discusses properties and derivation of Nash and saddle-point equilibria in infinite dynamic games of prescribed fixed duration. The analysis is first confined to dynamic games defined in discrete time, and with a finite number of stages, and then extended to differential games. Some results for infinite-horizon formulations are also presented, primarily for affine-quadratic structures. Utilization of the two standard techniques of optimal control theory, viz. the minimum principle and dynamic programming, leads to the so-called open-loop and feedback Nash equilibrium solutions, respectively. These two different Nash equilibria and their derivation and properties (such as existence and uniqueness) are discussed in Section 6.2, and the results are also specialized to affine-quadratic games, as well as to two-person zero-sum games. When the underlying information structure for at least one player is dynamic and involves memory, a plethora of (so-called informationally nonunique) Nash equilibria with different sets of cost values exists, whose derivation entails some rather intricate analysis—not totally based on standard techniques of optimal control theory. This derivation, as well as several important features of Nash equilibria under closed-loop perfect state information pattern, are discussed in Section 6.3, first within the context of a scalar three-person two-stage game (cf. subsection 6.3.1) and then for general dynamic games in discrete time (cf. subsection 6.3.2). Section 6.4 is devoted to derivation of necessary and sufficient conditions for Nash equilibria in stochastic nonzero-sum dynamic games with deterministic information patterns. Such a stochastic formulation eliminates informational

265

266 T. BAŞAR AND G. J. OLSDER nonuniqueness, thereby making the question of existence of unique Nash equilibrium under closed-loop perfect state information pattern a meaningful one. Section 6.5 presents the counterparts of the results of Section 6.2 in the continuous time, that is, for differential games with fixed duration, and the next one (Section 6.6) discusses some important applications of this theory to worst-case controller design (so-called  $JET^\circ$ -optimal control), first in continuous and then in discrete time.

Finally, Section 6.7 presents the counterparts of the results of Section 6.4 in the continuous time. 6.2 Open-Loop and Feedback Nash and Saddle-Point Equilibria for Dynamic Games in Discrete Time Within the context of Def. 5.1, consider the class of AT-person discrete-time deterministic infinite dynamic games of prescribed fixed duration ( $K$  stages) which are described by the state equation  $X_{k+1} = f(x_k, u_k, v_k)$ ,  $k \in K$ , (6.1) where  $n \in I = \mathbb{R}^n$  and  $X$  is specified a priori. Furthermore, the control sets  $U_i$  are taken as measurable subsets of  $\mathbb{R}^{m_i}$  ( $t \in C \mathbb{R}^{m_i}$ ;  $k \in K$ ,  $i \in N$ ), and a stage-additive cost functional  $K$   $V\{u^1, \dots, u^N\} = \sum_{k=1}^K J(x_k, u^1, \dots, u^N, v^1, \dots, v^N)$  (6.2)  $k=i$  is given for each player  $i \in N$ . In this section, we discuss derivation of Nash equilibria for this class of nonzero-sum games when the information structure of the game is either (i) open-loop or (ii) memoryless perfect state for all the players. In the former case, any permissible strategy for  $P_i$  at stage  $k \in K$  is a constant function and therefore can be considered to be an element of  $U_i$ , i.e.,  $U_i^k = U_i$ ,  $k \in K$ ,  $i \in N$ . Under the latter information structure, however, any permissible strategy for  $P_i$  at stage  $k \in K$  is a measurable mapping  $\gamma_i^k : X \rightarrow U_i$ ,  $i \in N$ ,  $k \in K$ . With  $\Gamma_i^k$  ( $k \in K$ ,  $i \in N$ ) taken as the appropriate strategy space in each case, we recall (from Def. 3.12) that an AT-tuple of strategies  $\{\gamma_i^k; i \in N, k \in K\}$  constitutes a Nash equilibrium solution if, and only if, the following inequalities are satisfied for all  $\{\gamma_i^k \in \Gamma_i^k; i \in N, k \in K\}$ :  $J_1^* = J_1(\gamma_1^1, \dots, \gamma_1^K; \gamma_2^1, \dots, \gamma_2^K; \dots; \gamma_n^1, \dots, \gamma_n^K; \gamma_{n+1}^1, \dots, \gamma_{n+1}^K; \dots; \gamma_N^1, \dots, \gamma_N^K) < J_1(\gamma_1^1, \dots, \gamma_1^K; \gamma_2^1, \dots, \gamma_2^K; \dots; \gamma_n^1, \dots, \gamma_n^K; \gamma_{n+1}^1, \dots, \gamma_{n+1}^K; \dots; \gamma_N^1, \dots, \gamma_N^K; \gamma_{n+1}^1, \dots, \gamma_{n+1}^K; \dots; \gamma_N^1, \dots, \gamma_N^K)$  (6.3)

NASH AND SADDLE-POINT EQUILIBRIA 267 Here,  $\gamma_i^k = \{\gamma_i^k, \dots, \gamma_i^k\}$ —the aggregate strategy of  $P_i$ —and the notation  $\gamma_i^k \in \Gamma_i^k$  stands for  $\gamma_i^k \in T^k \times U_i$ ;  $\Gamma_i^k \in K$ . Moreover,  $J_1^1, \dots, J_1^K$  is equivalent to  $Z/(u^1, \dots, u^N)$  with  $J$  replaced by  $J|_{(\cdot)}$  ( $i \in N, k \in K$ ). Under the open-loop information structure, we refer to the Nash solution as "open-loop Nash equilibrium solution", which we discuss in the first subsection to follow. For the memoryless perfect state information, the Nash solution will be referred to as "closed-loop no-memory Nash equilibrium solution", and under a further restriction, which is the counterpart of Def. 3.22 in the present framework. It will be called "feedback Nash equilibrium solution". Both types of equilibria will be discussed in subsection 6.2.2. 6.2.1 Open-loop Nash equilibria One method of obtaining the open-loop Nash equilibrium solution(s) of the class of discrete-time games formulated above is to view them as static infinite games and directly apply the analysis of Chapter 4. Toward this end we first note that, by backward recursive

substitution of (6.1) into (6.2), it is possible to express  $U$  solely as functions of  $\{u^j \mid j \in N\}$  and the initial state  $x$  whose value is known a priori, where  $u$  is defined as the aggregate control vector  $(u^1, \dots, u^N)'$ . This then implies that, to every given set of functions  $\{f_k \mid k \in K\}$ , there corresponds a unique function  $L_i : X \times U_1 \times \dots \times U_N \rightarrow R$ , which is the cost functional of  $P_i$ ,  $i \in N$ . Here,  $C_i$  denotes the aggregate control set of  $P_j$ , compatible with the requirement that if  $u \in C_i$  then  $u^k \in U^k$ ,  $\forall k \in K$ , and the foregoing construction leads to a normal form description of the original game, which is no different from the class of infinite games treated in Chapter 4. Therefore, to obtain the open-loop Nash equilibria, we simply have to minimize  $L_i(x, u_1, \dots, u_N)$  over  $U_i$ , for each  $i \in N$ , and then determine the intersection point(s) of the resulting reaction curves. In particular, if  $L_i(x, u_1, \dots, u_N)$  is continuous on  $U_1 \times \dots \times U_N$ , strictly convex in  $u^i$ , and further if  $U^*$  are closed, bounded and convex, an open-loop Nash equilibrium (in pure strategies) exists (cf. Thm. 4.3). Such an approach can sometimes lead to quite unwieldy expressions, especially if the number of stages in the game is large. An alternative derivation which partly removes this difficulty is the one that utilizes techniques of optimal control theory, by making explicit use of the stage-additive nature of the cost functionals (6.2) and the specific structure of the extensive form description of the game, as provided by the state equation (6.1). There is in fact a close relationship between derivation of open-loop Nash equilibria and the problem of solving (jointly)  $N$  optimal control problems, which can readily be observed from inequalities (6.3) since each one of them (together with (6.1) and (6.2)) describes an optimal control problem whose structure is not affected by the remaining players' control vectors. Exploring this relationship a little further, we arrive at the following result. Theorem 6.1 For an  $N$ -person discrete-time infinite dynamic game, let

268 T. BAŞAR AND G. J. OLSDER (i)  $f_k(-, u^1, \dots, u^N)$  be continuously differentiable on  $R^n$ , ( $A; \in K$ ), (ii)  $g_k(-, u^1, \dots, u^N, -)$  be continuously differentiable on  $R^n \times R^m$ , ( $k \in K, i \in N$ ), (Hi)  $f_k(\cdot, \dots, \cdot)$  be convex on  $R^n \times R^m \times \dots \times R^m$ , ( $k \in K$ ). Then, if  $\{Y^*(x) = u^i; i \in N\}$  provides an open-loop Nash equilibrium solution and  $\{x^{i+1}, k \in K\}$  is the corresponding state trajectory, there exists a finite sequence of  $n$ -dimensional (costate) vectors  $\{p^i \in R^n \mid p^i \in R^n\}$  for each  $i \in N$  such that the following relations are satisfied:  

$$\dot{p}^i = -\partial f_k / \partial x^i - \sum_{j \in N} p^j \partial g_j / \partial x^i, \quad p^i = 0, \quad (6.4a)$$

$$u^i = \arg \min_{u^i \in U^i} f_j(p^i, x^i, u^i, \dots, u^N), \quad (6.4b)$$
where  $u^i \in U^i$ ,  $N^i = \{j \in N \mid j \neq i\}$ .

$(6.4c) \quad P_i^{T+1} = \dots$

Every such Nash equilibrium solution is weakly time consistent. Proof. Consider the  $j$ 'th inequality of (6.3), which says that  $Y^*(x_i) = \dots$  minimizes  $L^*(u^1, \dots, u^N)$  over  $U^*$  subject to the state equation  $x_{k+1} = f_k(x_k, u_k^1, \dots, u_k^N)$ ,  $k \in K$ . This is a standard optimal control problem for  $P_i$  since  $u^j$  ( $j \in K, j \neq i$ ) are open-loop controls and hence do not depend on  $u^i$ . The result, then, follows directly from the minimum principle for discrete-time control systems (cf. Thm. 5.5). The "weak time consistency" of the solution is a direct consequence of Def. 5.13, where the "sol" operation is the Nash solution. D Theorem 6.1 thus provides a set of necessary conditions (solvability of a set of coupled two-point boundary value problems) for the open-loop Nash solution to satisfy; in other words, it produces candidate equilibrium solutions. In principle, one has to determine all solutions of this set of equations and further investigate which of these candidate solutions satisfy the original set of inequalities (6.3). If some further restrictions are imposed on  $f_k$  and  $g_k$  ( $i \in N, k \in K$ ) so that

NASH AND SADDLE-POINT EQUILIBRIA 269 the resulting cost functional  $U$  (defined earlier, in this subsection) is convex in  $u^i$  for all  $v, i \in U_j, j \neq i, j \in K$ , the latter phase of verification can clearly be eliminated, since then every solution set of (6.4a)-(6.4c) constitutes an open-loop Nash equilibrium solution. A specific class of problems for which this can be done, and the conditions involved expressed explicitly in terms of the parameters of the game, is the class of so-called "affine-quadratic" games which we first formally introduce below.

**Definition 6.1** An  $N$ -person discrete-time infinite dynamic game is of the affine-quadratic type if  $U_k = R_k^1$  ( $i \in N, k \in \sim K$ ), and

$$f_k(x_k, u_k^1, \dots, u_k^N) = A_k x_k + \sum_{i=1}^N B_{ki} u_i^k + c_k \quad (6.5a)$$

$$g_k(x_{k+1}, u_k^1, \dots, u_k^N, x_k) = -\frac{1}{2} x_{k+1}^T Q_{k+1} x_{k+1} + \sum_{i=1}^N u_i^k R_{ki} \quad (6.5b)$$

where  $A_k, B_k, Q_{k+1}, R_{ki}$  are matrices of appropriate dimensions,  $Q_{k+1}$  is symmetric,  $R_{ki} > 0$ ,  $c_k \in R^n$  is a fixed vector sequence, and  $k \in K, i \in N$ . An affine-quadratic game is of the linear-quadratic type if  $c_k = 0$ .

**Theorem 6.2** For an  $N$ -person affine-quadratic dynamic game with  $Q_{k+1} > 0$  ( $i \in N, k \in K$ ), let  $A_k, M_{k+1}$  ( $k \in K, j \in N$ ) be matrices of appropriate dimensions, defined by  $A_k = \dots$

$$M_i = Q_i + \sum_{t=i}^T A_{i+1}^T M_{t+1} A_{i+1} \quad (6.6a)$$

$$M^{T+1} = Q_{T+1} \quad (6.6b)$$

If the matrices  $M_k$  ( $k \in K$ ), thus recursively defined, are invertible, the game admits a unique open-loop Nash equilibrium solution given by  $u^i(x_1) = u^i = -M_i^{-1} B_i^T M_i^{-1} A_i x_1$ , ( $k \in$

$K, i \in \mathbb{N}$ ), (6.7a) where  $\{x^{k+1}; k \in K\}$  is the associated state trajectory determined from  $x^{k+1} = A^k x^k + B^k u^k$ ;  $x^0 = x^0$ , (6.7b) and  $B^k, T^k$  are defined by  $B^k = A^k B^0$ ,  $T^k = A^k T^0$ , (6.7c) The stagewise cost functions  $g^k$  can also be taken to depend on  $x^k$  instead of  $x^{k+1}$ , as in the proof of Prop. 5.1; but we prefer here the present structure (without any loss of generality) for convenience in the analysis to follow.

270 T. BAŞAR AND G. J. OLSDER  $g^k = c^k - B^k [I - A^k]^{-1} B^k m^{k+1}$ , (6.7d) with  $m^k$  recursively generated by  $m^k = A^k m^{k+1} + B^k u^k$ ,  $m^N = 0$ ,  $i \in \mathbb{N}$ ,  $c^k \in K$ . (6.7e) Proof. Since  $Q^k > 0$ , and  $B^k u^k > 0$ ,  $L^k(x^k, u^k, \dots, u^N)$  is a strictly convex function of  $u^k$  for all  $u^j \in R^{m^j}$  ( $j = i, i+1, \dots, N$ ) and for all  $x^i \in R^n$ .

Therefore, every solution set of (6.4a)-(6.4c) provides an open-loop Nash solution. Hence, the proof will be completed if we can show that (6.7a) is the only candidate solution. First note that  $H^k = [A^k x^k + B^k u^k + B^k m^{k+1} + c^k - B^k [I - A^k]^{-1} B^k m^{k+1}]^T$  and since  $Q^k > 0$ ,  $R^k > 0$ , minimization of this "Hamiltonian" over  $u^k \in R^{m^k}$  yields the unique relation  $\frac{\partial H^k}{\partial u^k} = -[B^k (I - A^k)^{-1} B^k m^{k+1} + c^k - B^k [I - A^k]^{-1} B^k m^{k+1}]^T$  (\*) where  $\frac{\partial H^k}{\partial x^k} = A^k x^k + B^k u^k + c^k - B^k [I - A^k]^{-1} B^k m^{k+1}$ ;  $x^k = x^k(u)$  Furthermore, the costate (difference) equation in (6.4c) reads  $P^k = A^k P^{k+1} + Q^k x^k + B^k u^k - P^{k+1} = 0$ . (ii) Let us start with  $k = K$  in which case (i) becomes  $\frac{\partial H^K}{\partial u^K} = -M^K x^{K+1}$  (ii) and if both sides are first premultiplied by  $B^K$  and then summed over  $i \in \mathbb{N}$  we obtain, by also making use of (ii) and (6.6a),  $x^{k+1} \sim A^k x^k = (I - A^k) x^{k+1} + c^k$  which further yields the unique relation  $x^{k+1} = [I - A^k]^{-1} c^k$  which is precisely (6.7b) for  $k = K$ . Substitution of this relation into (iv) then leads to (6.7a) for  $k = K$ .

NASH AND SADDLE-POINT EQUILIBRIA 271 We now prove by induction that the unique solution set of (i)-(iii) is given by (6.7a)-(6.7b) and  $p^k = A^k m^{k+1} + c^k - B^k [I - A^k]^{-1} B^k m^{k+1}$  ( $k \in \mathbb{N}$ ,  $A^k \in K$ ). Let us assume that this is true for  $k = l + 1$  (already verified for  $l = K - 1$ ) and prove its validity for  $k = l$ . First, using the solution  $p^{l+1} = A^{l+1} [M^{l+1} x^{l+1} + c^{l+1}]$  (i) with  $k = l$ , we obtain, after several algebraic manipulations,  $u^l = -[B^l (I - A^l)^{-1} B^l m^{l+1}]$ . (ii) Again, premultiplying this expression by  $B^l$ , and summing it over  $i \in \mathbb{N}$  leads to, also in view of (ii), (6.6a) and (6.7d),  $m^{l+1} = A^l m^{l+1} + B^l u^l$  which is (6.7b). If this relation is used in (v), we obtain the unique control vectors (6.7a), for  $k = l$ , and if it is further used in (iii) we obtain, in view of (6.7e), This then closes the induction argument, and thereby completes the proof of the theorem. D

Remark 6.1 An alternative derivation for the open-loop Nash equilibrium

solution of the affine-quadratic game is (as discussed earlier in this subsection in a general context) to convert it into a standard static quadratic game and then to make use of the available results on such games (cf. Prop. 4.6). By backward recursive substitution of the state vector from the state equation into the quadratic cost functionals, it is possible to bring the cost functional of  $P_i$  into the structural form as given by (4.23), which further is strictly convex in  $u_i = (u_i^1, \dots, u_i^K)$  because of assumptions  $Q_i > 0$ ,  $R_i > 0$ ,  $(A_i; \in K)$ .<sup>50</sup> Consequently, each player has a unique reaction curve, and the condition for existence of a unique Nash solution becomes equivalent to the condition for unique intersection of these reaction curves (cf. Prop. 4.6).<sup>51</sup> The existence condition of Thm. 6.2, i.e., nonsingularity of  $A^*$ ,  $(A; \in K)$ , is precisely that condition, but expressed in a different (more convenient, recursive) form. As in Prop. 4.6, it is of course possible for an affine-quadratic dynamic game to have multiple Nash equilibria, which would happen if  $A_{fc}$  is singular for some  $A_i$ , which corresponds to  $R$  being singular in (4.26a), and  $r$  being in its range space. For yet another derivation of the result of Thm. 6.2, by making use of Prop. 5.1, the reader is referred to Problem 1, Section 6.8.  $\square$  <sup>50</sup>The condition  $Q_{ik+1} > 0$  is clearly sufficient (along with  $R_i > 0$ ) to make  $U_i$  strictly convex in  $u_i$ , but is by no means necessary. It can be replaced by weaker conditions (which ensure convexity) under which the statements of Thm. 6.2 and this remark are still valid. One such condition, that is in fact tight, is given in Lemma 6.1 later, which should be interpreted for the present context. <sup>51</sup> For a derivation of open-loop Nash solution in two-person linear-quadratic games, and along these lines, the reader is referred to Başar (1976a) and Olsder (1977a).

272 T. BAŞAR AND G. J. OLSDER Zero-sum dynamic games We now turn our attention to another special class of nonzero-sum dynamic games—the two-person zero-sum games—in which case (assuming, in accordance with our earlier convention, that  $P_1$  is the minimizer and  $P_2$  is the maximizer),  $1 \leq 2$   $A_{ik} = -A_{ik} = A_{ik}$  (6.8a) and the "Nash" inequalities (6.3) reduce to the saddle-point inequality  $J(u_i, u_j) < J(u_i, u_j^*) < J(u_i^*, u_j)$ , (6.8b) where  $J^1 - J^2$ . Directly applying Thm. 6.1 in the present context, we first have from (6.4a)-(6.4c) (in view of (6.8a)) the relation  $p_k = -p_k$ , ( $k \in K$ ), and therefore from (6.4d),  $H_k = -H_k = H_k$ , ( $k \in K$ ). Hence, we have arrived at the following conclusion. Theorem 6.3 For a two-person zero-sum discrete-time infinite dynamic game, let (i)  $f_k(-, u_k^t)$  be continuously differentiable on



$R^n$ , ( $k \in K$ ),  $(H) \varphi_k(-, \cdot)$  be continuously differentiable on  $R^n \times R^n$ , ( $k \in K$ ),  $(H_i) f_k(-, \cdot)$  be convex on  $R^n \times U_1 \times U_2$ , ( $k \in K$ ). Then,  $\{u^i; i = 1, 2\}$  provides an open-loop saddle-point solution, and  $\{x_{k+1}; k \in K\}$  is the corresponding state trajectory, there exists a finite sequence of  $n$ -dimensional (costate) vectors  $\{p_2, \dots, p_{k+1}\}$  such that the following relations are satisfied:  $x_{k+1} = f_k(x^k, u^k, p^k)$ ,  $f_k(p^{k+1}, u^k, x^k) = H_k(p^{k+1}, u^k, x^k)$ ,  $\forall u^i \in U_i, i = 1, 2$   $\int_{t_0}^{t_1} p^k \cdot \dot{x}^k = g^k(x^k, u^k, p^k) - p^k \cdot (A^k x^k + B^k u^k + c^k)$  (6.9a)  $I, (6.9b) (6.9c) p_{K+1} = 0$ , where  $p^k = p^k(x^k, u^1, u^2), k \in K$ . (6.9d)

**NASH AND SADDLE-POINT EQUILIBRIA 273** Proof. The result follows from Thm. 6.1, as discussed earlier. D As a specific application of this theorem, we now consider the class of affine-quadratic two-person zero-sum dynamic games (cf. Def. 6.1) described by the state equation  $x_{k+1} = A^k x_k + B^k u^k + c^k, k \in K$ , (6.10a) and the objective functional  $J(x^0, u^1, u^2) = -\frac{1}{2}(x^k)^T Q^k x^k + \sum_{k=0}^{K-1} (f^k(u^1) - u^2 k)$  (6.10b) which  $P_1$  wishes to minimize and  $P_2$  attempts to maximize. It should be noted that we have taken the weighting matrices for the controls in (6.10b) as unity, without any loss of generality, since otherwise they can be absorbed in  $B^k$  and  $B^2$  provided, of course, that  $R^k \} > 0$  and  $R^2 > 0$  which was our a priori assumption in Def. 6.1. Let us also assume that  $Q^k > 0 (A; k \in K)$  which essentially makes  $L(u^1, u^2)$  strictly convex in  $u^1$ . In order to formulate a meaningful problem, we also have to require  $L(u^1, u^2)$  to be (strictly) concave in  $u^2$ , since otherwise  $P_2$  can make the value of  $L$  unboundedly large. The following lemma now provides a necessary and sufficient condition for (6.10b) to be strictly concave in  $u^2$ .

**Lemma 6.1** For the affine-quadratic two-person zero-sum dynamic game introduced above, the objective functional  $L(u^1, u^2)$  is strictly concave in  $u^2$  (for all  $u^1 \in R^m$ ) if and only if  $I - B^2 S^k B^2 > 0 (k \in K)$ , (6.11a) where  $S^k$  is given by  $S^k = Q^k + A^k S^{k+1} A^k + A^k S^{k+1} B^2 [I - B^2 S^{k+1} B^2]^{-1} B^2 S^{k+1} A^k$ ;  $S^{K+1} = Q^{K+1}$  (6.11b) Proof Since  $Z^k(u^1, u^2)$  is a quadratic function of  $u^2$ , the requirement of strict concavity is equivalent to existence of a unique solution to the optimal control problem  $\min_{u^2} [-L(u^1, u^2)]$  subject to (6.10a) and for each  $u^1 \in R^m$ . Furthermore, since the Hessian matrix (matrix of second partials) of  $L$  with respect of  $u^2$  is independent of  $u^1$  as well as of  $c$ , we can instead consider the optimal control problem  $\min_{u^2} J^k(x^k, u^2) = -\frac{1}{2} x^k T Q^k x^k + \sum_{k=0}^{K-1} (f^k(u^1) - u^2 k)$

274 T. BAŞAR AND G. J. OLSDER subject to  $x_{k+1} = A_k x_k + B_k u_k$ ; that is, we can take  $u_k = 0, c_k = 0$  without any loss of generality. Then, the result follows from the dynamic programming technique outlined in subsection 5.5.1, and in particular from (5.19), which admits the unique solution  $V(k, x) = -x' S_k x + x' Q_k x$ , if, and only if, (6.11a) holds. We are now in a position to present the open-loop saddle-point solution of the affine-quadratic two-person zero-sum dynamic game. Theorem 6.4 For the two-person affine-quadratic zero-sum dynamic game described by (6.10a)-(6.10b) and with  $Q_{k+i} > 0$  ( $k \in K$ ), let condition (6.11a) be satisfied and  $A_k, M_k$  ( $k \in K$ ) be matrices of appropriate dimensions, defined through  $A_{k+1} = [A_k + (B_k B_k' - B_{k+1} B_{k+1}') M_{k+1}]^{-1} M_{k+1} Q_{k+1} + A_k' M_{k+1} A_k$ ;  $M_{K+1} = Q_{K+1}$ . (6.12a) (6.12b) Then, (i) the matrices  $A_k$  ( $k \in K$ ), thus recursively defined, are invertible, (ii) the game admits a unique open-loop saddle-point solution given by  $u_k^* = -B_k' M_{k+1} A_k^{-1} (x_k - x_{k+1}) + c_k$ , (6.13a)  $x_{k+1} = A_k^{-1} (x_k - B_k u_k^* + c_k)$ ,  $k \in K$ , (6.13b) where  $\{x_{k+1}; k \in K\}$  is the corresponding state trajectory determined from  $x_{k+1} = A_k^{-1} [A_k x_k + B_k u_k^* + c_k]$ ,  $x_1 = x$  (6.13c) and  $\{r_k, r_k\}$  are given by  $r_k = M_{k+1} A_k^{-1} r_{k+1} + m_{k+1}$ , (6.13d)  $r_k = c_k - (B_k B_k' - B_{k+1} B_{k+1}') m_{k+1}$ , (6.13e) with  $m_k$  generated by the difference equation  $m_k = A_k^{-1} [A_k^{-1} m_{k+1} + c_k + x_k A_k^{-1} c_k]$ ,  $m_{K+1} = 0$ . (6.13f)

NASH AND SADDLE-POINT EQUILIBRIA 275 Proof. Let us first note that, with the open-loop information structure and under condition (6.11a) (cf. Lemma 6.1), the game is a static strictly convex-concave quadratic zero-sum game which admits a unique saddle-point solution by Corollary 4.5. Second, it follows from Thm. 6.3 that this unique saddle-point solution should satisfy relations (6.9a)-(6.9d), which can be rewritten, for the affine-quadratic game, as follows:  $x_{k+1}^* = A_k x_k^* + B_k u_k^* + c_k$ ,  $x_1 = x$ , (i)  $H_k(p_{k+1}, u_k^*, u_k^*, x_k^*) < H_k(p_{k+1}, u_k^*, u_k^*, x_k^*) < H_k(p_{k+1}, u_k^*, u_k^*, x_k^*)$  ( $\ll$ )  $P_k = A_k' p_{k+1} + Q_{k+1} x_k^*$ ,  $P_{K+1} = 0$ , (iii)  $H_k(p_{k+1}, u_k^*, u_k^*, x_k^*) = -[A_k x_k^* + B_k u_k^* + c_k]' p_{k+1} + B_k' p_{k+1} (A_k x_k^* + B_k u_k^* + c_k)$ ,  $k \in K$ . An inductive argument, as in the proof of Thm. 6.2, shows that this set of equations admits the solution:  $u_k^* = -(B_k' p_{k+1} + c_k) (A_k^{-1} B_k B_k' + Q_{k+1})^{-1} (A_k x_k^* + B_k u_k^* + c_k)$ ,  $k \in K, i = 1, 2$ , (iv)  $p_k = A_k' [M_{k+1} x_{k+1}^* + m_{k+1}]$ ,  $k \in K$ , where the corresponding value of the state vector,  $x_k$ , satisfies  $A_k x_{k+1}^* = A_k x_k^* + r_k$ . Since there exists a unique saddle-point solution, there necessarily exists a unique relationship between  $x_k$  and  $x_{k+1}$ , implying that the matrix  $A_k$  should be invertible for each  $k \in K$ . Hence,  $x_{k+1} = A_k^{-1} [A_k x_k^* + r_k]$ ,  $A_k \in K$ , which verifies (6.13c). When

this is used in (iv), it leads to (6.13a)-(6.13b), also in view of (6.13d)-(6.13f), thus completing the proof of the theorem. D Remark 6.2 The statement of Thm. 6.4 is valid even if the "nonnegative definiteness" condition on  $Q_{k+i}$  does not hold, provided that some other appropriate conditions are imposed on the game parameters to ensure that  $L$  defined by (6.10b) is strictly convex in  $u_1$ . One such set of conditions, which are in fact tight, can be obtained directly from Lemma 6.1, by simply replacing  $Q_k$  by  $-Q_k$ , and  $B_2$  by  $B_1$ , that is,  $I + B^{-1}S_{k+1}B > 0$  ( $f \in K$ ), (6.14a) where  $S_k$  is generated by  $S_i = Q_k + A^k S_{i+1} A^k + B^{-1} S_{i+1} B^{-1}$  ( $i \in N$ ),  $S_{N+1} = 0$ . (6.14b)

276 T. BAŞAR AND G. J. OLSDER 6.2.2 Closed-loop no-memory and feedback Nash equilibria We now turn our attention to a discussion and derivation of Nash equilibria under the memoryless perfect state information pattern which provides the players with only the current value of the state at every stage of the game (and of course also the initial value of the state, which is known a priori). For this class of dynamic games we first obtain the following counterpart of Thm. 6.1, which provides a set of necessary conditions for any closed-loop no-memory Nash equilibrium solution to satisfy: Theorem 6.5 For an  $N$ -person discrete-time infinite dynamic game, let (i)  $f_k(\cdot)$  be continuously differentiable on  $R^n \times U_k \times \dots \times U_j$ , ( $k \in K$ ), (ii)  $g_k(\cdot)$  be continuously differentiable on  $R^n \times U_k \times \dots \times U_j \times R^n$ , ( $k \in K, i \in N$ ). Then, if  $\{x^k(i) = u^k; k \in K, i \in N\}$  provides a closed-loop no-memory Nash equilibrium solution such that  $Y^k(\cdot)$  is continuously differentiable on  $R^n$  for all  $k \in K, i \in N$ ,  $f_k(x^k, u^1, \dots, u^N)$  is convex in  $(x^k, u^1, \dots, u^N)$  for every  $i \in N, k \in K$ , and if  $\{x^k(i); k \in K\}$  is the corresponding state trajectory, there exists a finite sequence of  $n$ -dimensional (costate) vectors  $\{p^k(i)\}$ , for each  $i \in N$  such that the following relations are satisfied:  $x^k(i+1) = f_k(x^k(i), u^k(i), \dots, u^j(i))$ ,  $x^k(0) = x^0$ , (6.15a)  $u^k(i) = \arg \min_{u^k} H^k(p^k(i+1), u^k(i), \dots, u^j(i))$ , (6.15b)  $p^k(i) = -\partial H^k(p^k(i+1), u^k(i), \dots, u^j(i)) / \partial x^k(i)$  where  $p^k(i+1) + dx^k(i) = g^k(p^k(i+1), x^k(i), u^k(i), \dots, u^j(i))$ ,  $p^k(N) = 0$ ,  $i \in N, k \in K$ , (6.15c)  $H^k(p^k(i+1), u^k(i), \dots, u^j(i)) = \lambda^k(i) + p^k(i+1) \cdot (f_k(x^k(i), u^k(i), \dots, u^j(i)) - x^k(i+1)) + p^k(i+1) \cdot (g^k(p^k(i+1), x^k(i), u^k(i), \dots, u^j(i)) - p^k(i+1))$ ,  $k \in K, i \in N$ , (6.15d)

NASH AND SADDLE-POINT EQUILIBRIA 277 and a similar convention applies to  $A^k(\cdot)$  and  $\lambda^k(\cdot)$ . Proof. The proof is similar to that of Thm. 6.1, but here we also have to take into account the possible dependence of  $u^k(i)$  on  $X^k$  ( $k \in K, i \in N$ ). Accordingly, the  $i$ th inequality of (6.3) now says that  $\lambda^k(i)$

$(x_k, x)$ ;  $k \in K$  } minimizes the function  $J_i(\gamma_1, \dots, \gamma_{i-1}, \gamma_i, \gamma_{i+1}, \dots, \gamma_A) = E \sum_{k=1}^K (x_k + W^*(x_k, i)) / c_k = E \sum_{k=1}^K (x_k + 1) / c_k$  over  $P$  subject to the state equation constraint  $Z_{k+1} = f_k(x_k, j_1^*(x_k, X_1), \dots, j_{i-1}^*(x_k, X_1), j_k(x_k, X_1), j_{k+1}^*(x_k, X_1), \dots, j_{k+1}^*(x_k, x_1)) = f_k(x_k, j_j(x_k, x_1))$ ,  $A \in K$ . Then, the result follows directly from Thm. 5.5 by taking  $f_k$  as  $f_k$  and  $g_k$  as  $1/c_k$  which are continuously differentiable in their relevant arguments because of the hypothesis that the equilibrium strategies are continuously differentiable in their arguments. If the set of relations (6.15a)-(6.15d) is compared with (6.4a)-(6.4d), it will be seen that they are identical, except for the costate equations—the latter equation (6.15c) having two additional terms which are due to the dynamic nature of the information structure, allowing the players to utilize current values of the state. Furthermore, it is important to notice that every solution of the set (6.4a)-(6.4c) also satisfies relations (6.15a)-(6.15c), since every such solution is associated with static information, thereby causing the last two terms in (6.15c) to drop out. But, the "open-loop" solution is not the only one that the set (6.15a)-(6.15c) admits; surely it could have other solutions which explicitly depend on the current value of the state, thus leading to nonunique Nash equilibria, all of which are weakly time consistent (a la Def. 5.13). This phenomenon of multiplicity of Nash equilibria is closely related to the "informational nonuniqueness" feature of Nash equilibria under dynamic information, as introduced in Section 3.5 (see, in particular, Prop. 3.10 and Remark 3.15), whose counterpart in infinite games will be thoroughly discussed in the next section. Here, we deal with a more restrictive class of Nash equilibrium solutions under memoryless perfect state information pattern—the so-called feedback Nash equilibrium—which is devoid of any "informational nonuniqueness" (see also Def. 3.22) and is also strongly time consistent (cf. Def. 5.14).

278 T. BAŞAR AND G. J. OLSDER Definition 6.2 For an  $N$ -person  $K$ -stage discrete-time infinite dynamic game with memoryless perfect state information pattern,<sup>52</sup> let  $J_i$  denote the cost functional of  $P_i$  ( $i \in N$ ) defined on  $F_1 \times \dots \times F_N$ . An  $N$ -tuple of strategies  $\{j_i^* \in T_x; i \in N\}$  constitutes a feedback Nash equilibrium solution if it satisfies the set of  $K$   $N$ -tuple inequalities (3.28) for all  $x_k \in X_k$ ,  $i \in N$ ,  $k \in K$ . Proposition 6.1 Under the memoryless perfect state information pattern, every feedback Nash equilibrium solution of an  $N$ -person discrete-time infinite dynamic game is a closed-loop no-memory Nash equilibrium solution (but not vice versa). Proof This result is the counterpart of

Prop. 3.9 in the present framework and therefore its proof parallels that of Prop. 3.9. The result basically follows by showing that, for each  $i \in N$ , the collection of all the  $i$ th inequalities of the  $K$   $i$ -tuples imply the  $i$ th inequality of (6.3). Definition of the feedback Nash equilibrium solution directly leads to a recursive derivation which involves solutions of static  $N$ -person nonzero-sum games at every stage of the dynamic game. Again, a direct consequence of Def. 6.2 is that the feedback equilibrium solution depends only on  $x_k$  at stage  $k$ , and dependence on  $X$  is only at stage  $k = 1$ .<sup>53</sup> By utilizing these properties, we readily arrive at the following theorem. Theorem 6.6 For an  $N$ -person discrete-time infinite dynamic game, the set of strategies  $\{f_i^k(x_k); k \in K, i \in N\}$  provides a feedback Nash equilibrium solution if, and only if, there exist functions  $V_i(k, \cdot): R^n \rightarrow R, k \in K, i \in N$ , such that the following recursive relations are satisfied: 
$$V_i(k, x) = \min_{u_i} [g_i(x, u_i) + V_i(k+1, r_i^k(x, u_i))] \quad (6.16)$$
 
$$V_i(K+1, x) = 0, i \in N, \text{ where } r_i^k(x, u_i) = (f_i^k(x, u_i), \dots, f_{i-1}^k(x, u_{i-1}), u_i, f_{i+1}^k(x, u_{i+1}), \dots, f_N^k(x, u_N)).$$
 Every such equilibrium solution is strongly time consistent, and the corresponding Nash equilibrium cost for  $P_i$  is  $V_i(1, X_i)$ . Proof. Let us start with the first set of  $N$  inequalities of (3.28). Since they have to hold true for all  $x^k \in T_k, i \in N, k < K - 1$ , this necessarily implies <sup>52</sup>The statement of this definition remains valid if the information pattern is instead "closed-loop perfect state". <sup>53</sup>This statement is valid also under the "closed-loop perfect state" information pattern. Note that the feedback equilibrium solution retains its equilibrium property also under the feedback information pattern.

NASH AND SADDLE-POINT EQUILIBRIA 279 that they have to hold true for all values of state  $x_k$  which are reachable by utilization of some combination of these strategies. Let us denote that subset of  $R^n$  by  $X_k$ . Then, the first set of inequalities of (3.28) becomes equivalent to the problem of seeking Nash equilibria of an  $N$ -person static game with cost functionals  $J_i^k(x_k, u_k, \dots, u_k), i \in N, (i)$  which should be valid for all  $x_k \in X_k$ . This is precisely what (6.16) says for  $k = K$ , with a set of associated Nash equilibrium controls denoted by  $\{u_i^k \in N\}$  since they depend explicitly on  $x_k \in X_k$ , but not on the past values of the state (including the initial state  $X$ ). Now, with these strategies substituted into (i), a similar argument (as above) leads to the conclusion that the second set of inequalities of (3.28) defines a static Nash game with cost functionals  $V_i^*(K, x_K) + g_{k-i}(x_K, u_{k-i}, \dots, u_{k-i}, X_k$

$i), i \in N$ , where  $X^k = f^{k-1}(x^{k-1}, u^{k-1}, \dots, u^{k-1})$ , and the Nash solution has to be valid for all  $x^{k-1} \in X^{k-1}$  (where  $X^{k-1}$  is the counterpart of  $X^k$  at stage  $k = K - 1$ ). Here again, we observe that the Nash equilibrium controls can only be functions of  $x^{k-1}$ , and (6.16) with  $k = K - 1$  provides a set of necessary and sufficient conditions for  $\{y^{k-1}(x^{k-1}) \in N\}$  to solve this static Nash game. The theorem then follows from a standard induction argument. Note that the "strong time consistency" property of the feedback Nash equilibrium, and the expression for the corresponding cost for each player, are direct consequences of the recursive nature of the construction of the solution. D The following corollary, which is the counterpart of Thm. 6.2 in the case of feedback Nash equilibrium, now follows as a special case of Thm. 6.6.

**Preliminary Notation for Corollary 6.1.** Let  $P^k, (i \in N, k \in K)$  be matrices of appropriate dimensions, satisfying the set of linear matrix equations  $W + B^k z^{k+1} + B^k z^{k+1} Y, B^k H = B^k Z^k + A^k, i \in N, (6.17a)$  where  $Z^k, (i \in N)$  are obtained recursively from  $Z^1 = F^k + Y, P^k + Q^k, Z^{k+1} = Q^{k+1} + P^{k+1}, i \in N, (6.17b)$  and  $F^k + A^k - Y, B^k P^k, k \in K, (6.17c)$

280 T. BAŞAR AND G. J. OLSDER Furthermore, let  $a^k \in R^m, (i \in N, k \in K)$  be vectors satisfying the set of linear equations:  $W + B^k z^{k+1} + B^k z^{k+1} Y, B^k H = B^k Z^k + A^k, i \in N, (6.17d)$  where  $Z^k, (i \in N)$  are obtained recursively from  $Z^1 = F^k + Y, P^k + Q^k, Z^{k+1} = Q^{k+1} + P^{k+1}, i \in N, (6.17e)$  and  $F^k + A^k - Y, B^k P^k, k \in K, (6.17f)$  Finally, let  $n^k, \in R^m, (i \in N, k \in K)$  be generated by  $4 = "i + ^Hj + 1 + Cfc + iA + ^£ l «*1 ««» nK + i = 0. (6.17g) J"6N$

**Corollary 6.1**  $n$ -person affine-quadratic dynamic game (cf. Def. 6.1) with  $Q^{k+1} > 0, (i \in N, k \in K)$  and  $J^k > 0, (i, j \in N, j \neq i, k \in K)$  admits a unique feedback Nash equilibrium solution  $i/j$  and only if, (6.17a) and (6.17d) admit unique solution sets  $\{P^{k*}; i \in N, k \in K\}$  and  $\{z^{k*}; i \in N, k \in K\}$ , respectively, in which case the equilibrium strategies are given by  $j^i(x^k) = -P^k z^{k*} - a^k, (k \in K, t \in N), (6.18a)$  and the corresponding feedback Nash equilibrium cost for each player is  $J^i(1^*, \dots, 1^N) = V^i(1, x^1) = \frac{1}{2} x^1 \sum_{z^1}^t f^k x^1 + n^i, (i \in N). (6.18b)$

**Proof.** Starting with  $k = K$  in the recursive equation (6.16), we first note that the functional to be minimized (for each  $i \in N$ ) is strictly convex, since  $R^k + B^k Q^{k+1} + B^k > 0$ . Then, the first-order necessary conditions for minimization are also sufficient and therefore we have (by differentiation) the unique set of equations  $= B^k Q^{k+1} [A^k x^k + c^k]; i \in N$ , which readily leads to the conclusion that any set of Nash equilibrium strategies at stage  $k = K$  has to be affine in  $x^k$ . Therefore, by substituting  $7^k = -P^k x^k - a^k, (i \in N)$  into the foregoing equation, and

by requiring it to be

NASH AND SADDLE-POINT EQUILIBRIA 281 satisfied for all possible  $x_k$ , we arrive at (6.17a) and (6.17d) for  $k = K$ . Further substitution of this solution into (6.16) for  $k = K$  leads to  $V^1(K, x) = \min_x \{Z_K x - Q_K x + Q_K x + n_K\}$  that is,  $V^1(K, \bullet)$  has a quadratic structure at stage  $k = K$ . Now, if this expression is substituted into (6.16) with  $k = K - 1$ , and the outlined procedure is carried out for  $k = K - 1$ , and this so (recursively) for all  $k < K - 1$ , one arrives at the conclusions that (i)  $V^1(k, x) = \min_x \{Z_k x - Q_k x + C_k x + n_k\}$  is the unique solution of the recursive equation (6.16) under the hypothesis of the corollary and by noting that  $Z_k > 0$  ( $i \in N, k \in K$ ), and (ii) the minimization operation in (6.16) leads to the unique solution (6.18a) under the condition of unique solvability of (6.17a) and (6.17d). The expression for the cost, (6.18b), follows directly from the expression derived for the "cost-to-go"  $V^1(k, x)$ . This, then, completes verification of Corollary 6.1.  $\square$  Remark 6.3 The result of Corollary 6.1 as well as the verification given above extends readily to more general affine-quadratic dynamic games where the cost functions of the players contain additional terms that are linear in  $X_k$ , that is, with  $g_i$  in (6.5b) replaced by  $g_i(x^{k+i}, u^1, \dots, u^m, x_k) = \alpha^i (x^{k+1} [Q_{k+1} x^{k+1} + 2l_{k+1}] + J_{2k} R_{k+1} u^k)$  where  $l_{k+1}$  ( $k \in K$ ) is a known sequence of  $n$ -dimensional vectors for each  $i \in N$ . Then, the statement of Corollary 6.1 remains intact, with only the equation (6.17e) that generates  $Q$  now reading  $a = \sum_{i=1}^m H_{ii} l_i + \sum_{i=1}^m z_{k+i} P_k + Y_{lp}(R_{lk} + h; C_{K+i} = J_{*+i}, t \in N$ , and the cost-to-go functions admitting the compatibly modified form  $V^1(k, x) = \min_x \{Z_k x - Q_k x + (a - l_k)' x + n_k\}$ ,  $i \in N$ .  $\square$  Remark 6.4 The "nonnegative-definiteness" requirements imposed on  $Q^{k+i}$  and  $R_k$ : ( $i, j \in N, j \wedge i; i \in K$ ) are sufficient for the functional to be minimized in (6.16) to be strictly convex, but they are by no means necessary. A set of less stringent (but more indirect) conditions would be  $R_t + B_{ik} Z_{l+1} B_{ik} > 0$  ( $i \in N, k \in K$ ), under which the statement of Corollary 6.1 still remains valid. Furthermore, it follows from the proof of Corollary 6.1 that, if (6.17a) admits more than one set of solutions, every such set constitutes a feedback Nash equilibrium solution, which is also strongly time consistent.  $\square$

282 T. BAŞAR AND G. J. OLSDER Remark 6.5 It is possible to give a precise condition for the unique solvability of the sets of equations (6.17a) and (6.17d) for  $P_k$  and  $a^k$  ( $i \in N, k \in K$ ), respectively. The said condition (which is the same for both) is the invertibility of matrices  $S_k^i, k \in K$ , which are

composed of block matrices, with the  $i$ th block given as  $R_{ii} + B_k Z_{k+1} B_k$  and the  $ij$ th block as  $B_j' Z_{k+1} B_j$ , where  $i, j \in N$ ,  $j \neq i$ .  $\square$  **Zero-sum dynamic games** We now consider the general class of two-person discrete-time zero-sum dynamic games and determine the "cost-to-go" equation associated with the feedback saddle-point solution, as a special case of Thm. 6.6. **Corollary 6.2** For a two-person discrete-time zero-sum dynamic game, the set of strategies  $\{Y_k^*(x_k) \in A; \epsilon K, i = 1, 2\}$  provides a feedback saddle-point solution if, and only if, there exist functions  $V\{k, x\} : R^n \rightarrow R, k \in K$ , such that the following recursive relation is satisfied:  $V(k, x) = \min \max [g_k(f_k(x, u_1, u_2), u_1, u_2, x) + V(k+1, f_k(x, u_1, u_2))] \max_j \min_i [g_j^*(f_j(x, u_1, u_2), u_1, u_2, x) + V(k+1, f_j(x, u_1, u_2))]$  (6.19)  $= g_k(f_k(x, u_1^*(x), u_2^*(x)), u_1^*(x), u_2^*(x)) + V(k+1, f_k(x, u_1^*(x), u_2^*(x)))$ ;  $V(K+1, x) = 0$ . Every such saddle-point solution is strongly time consistent, and the unique saddle-point value of the game is  $V\{1, X\}$ . **Proof** The recursive equation (6.19) follows from (6.16) by taking  $N = 2$ ,  $g^i = -g^j = g^k$  ( $k \in K$ ), since then  $V_1 = -V_2 = V$  and existence of a saddle point is equivalent to interchangeability of the min-max operations. **D** For the further special case of an affine-quadratic zero-sum game, the solution of (6.19) as well as the feedback saddle-point solution can be explicitly determined in a simple structural form, but after a rather lengthy derivation. We accomplish this in two steps: first we obtain directly the special case of Corollary 6.1 for the affine-quadratic zero-sum game (see Corollary 6.3 below), and second we simplify these expressions further so as to bring them into a form compatible with the results of Thm. 6.4 (see Thm. 6.7 in the sequel). **Corollary 6.3** For the two-person affine-quadratic zero-sum dynamic game described by (6.10a)-(6.10b), the unique solutions of (6.17a) and (6.17d) are given by  $P_i = [I + K_1 Z_k^{\wedge} B_1 Y' K_1 Z_k^{\wedge} A_k, (6.20a)$

**NASH AND SADDLE-POINT EQUILIBRIA 283** where  $P_i = -[I - K_2 Z_k^{\wedge} B_1]^{-1} K_2 Z_k^{\wedge} B_1 A_k, (6.20b)$   $a^i = [I + K_1 Z_k^{\wedge} B_1]^{-1} K_1 A_{i+1} + Z_{k+1} c_k, (6.20c)$   $a^i = -[I - K_1 Z_k^{\wedge} B_2 Z_k^{\wedge} B_1^{-1} K_1^{\wedge} k+1 + Z_{k+1} c_k], (6.20d)$   $K^i = B_1' [I + Z_{k+1} B_2 Z_k^{\wedge} (I - B_1' Z_{k+1} B_2 Z_k^{\wedge} B_1^{-1})^{-1}], (6.21a)$   $K_2 \pm B_2' [I - Z_{k+1} B_1 (I + B_1' Z_k^{\wedge} B_1^{-1} B_1')^{-1}], (6.21b)$  and  $Z_k = Z_{k+1}$  and  $\xi_k = \xi_{k+1}$  ( $\xi = \xi_k$  satisfy the recursive equations  $Z_k = F' Z_{k+1} F + P_1' P_1 - P_2' P_2 + Q_k; Z_{K+1} = Q_{K+1}, (6.22a)$  and  $a^i = \xi(a_{i+1} + Z_{k+1} C_k y, a_{i+1} = 0 (6.22b)$   $F_k^{\wedge} A_k - B_1 P_1 - B_2 K P_2, P_k^{\wedge} c_k - B_1 a^i - B_2 a^2$ . Furthermore, the set of conditions of Remark 6.4 is equivalent to  $I + B^{\wedge} Z_{k+i} B_1 > 0$  ( $A; \epsilon K$ ), (6.23a)  $/ - B_2 Z_{k+1} B_1 > 0$  ( $k \in K$ ). (6.23b) **Proof** By letting  $N = 2, Q_{j+1} = -Q_j, j \in N, Q_{N+1} = 0, R_{22} = -R_{k2} = I$  in



Corollary 6.1, we first observe that  $Z^1 = -Z^k$  and the equations for  $Z^k$  and  $f^k$  are the same as (6.22a) and (6.22b),<sup>54</sup> respectively, assuming of course that  $P^k$ 's and  $a^k$ 's in these equations are correspondingly the same—a property which we now verify. Toward this end, start with (6.17a) when  $i = 2$ , solve for  $P^k$  terms of  $P^k$  (this solution is unique under (6.23b)), substitute this into (6.17a) when  $i = 1$  and solve for  $P^k$  from this linear matrix equation. The result is (6.20a), assuming this time that the matrix  $[I + K^k Z^k + B^k x^k]$  is invertible. Furthermore, by following the same procedure with indices 1 and 2 interchanged, we arrive at (6.20b), on account of the invertibility of  $[I - K^k Z^k + B^k]$ . Repeating the same steps for  $a^1$  and  $a^2$ , this time by working with (6.17d), we arrive at (6.20c)-(6.20d), again under the same matrix invertibility.  
<sup>54</sup>Verification of (6.22b) requires some algebraic manipulations.

284 T. BAŞAR AND G. J. OLSDER conditions. Therefore, for this special case,  $P^k$  and  $P^k$  given by (6.20a)-(6.20b), and  $a^1$  and  $a^2$  given by (6.20c)-(6.20d), are indeed the unique solutions of (6.17a) and (6.17d), respectively, provided that the two matrices in question are invertible. A direct manipulation on these matrices actually establishes nonsingularity under conditions (6.23a)-(6.23b); however, we choose here to accomplish this by employing an indirect method which is more illuminating. First note that (6.23a) and (6.23b) correspond to the existence conditions given in Remark 6.4, and therefore they make the functional to be minimized in (6.16) strictly convex in the relevant control variables. But, for the zero-sum game, (6.16) is equivalent to (6.19), and consequently (6.23a) and (6.23b) make the kernel in (6.19) strictly convex in  $u^1$  and strictly concave in  $u^2$  ( $k \in K$ ). Since every strictly convex-concave quadratic static game admits a (unique) saddle point (cf. Corollary 4.5), the sets (6.17a) and (6.17d) have to admit unique solutions for the specific problem (zero-sum game) under consideration. Hence, the required inverses in (6.21a) and (6.21b) should exist. This, then, completes the proof of Corollary 6.3. D To further simplify (6.20a)-(6.22b), we now make use of the following matrix identity. Lemma 6.2 Let  $Z = Z'$  and  $B$  be matrices of dimensions  $(n \times n)$  and  $(n \times m)$ , respectively, and with the further property that  $B'ZB$  does not have any unity eigenvalues. Then,  $[I_n + ZB\{I_m - B'ZB\}^{-1}XB']^{-1} = \{I_n - ZBB'\}^{-1}$ . (6-24) Proof. First note that the matrix inverse on the RHS of this identity exists, under the hypothesis of the lemma, since nonzero eigenvalues of  $B'ZB$  and  $ZBB'$  are the same (see Marcus and Mine, 1964, p. 24). Then, the result follows by multiplying both sides of the identity by  $I_n - ZBB'$  from the right

and by straightforward manipulations. D Application of this identity to (6.21a) and (6.21b) readily yields (by identifying  $B \setminus$  with  $B$  and  $Z_{k+i}$  with  $Z$  in the former, and  $B \setminus$  with  $B$  and  $-Z_{k+i}$  with  $Z$  in the latter)  $K_k - B_k [I - Z_{k+i} B_k B_k]^{-1}$ ,  $K_k = B_k [I + Z_{k+1} B' B_k]^{-1}$ . Furthermore, if these are substituted into (6.20a)-(6.20d), some extensive, but straightforward, matrix manipulations which involve repeated application of (6.24) lead to the expressions  $P_k = B_k [Z_{k+1} [I + (B_1 B_1' - B_2 B_2) Z_{k+1}]^{-1} A_k]$ , (6.25a)  $P_k = B_k [Z_{k+1} [I + (B_1 B_1' - B_2 B_2) Z_{k+1}]^{-1} A_k]$ , (6.25b) Here we suppress the dimensions of the identity matrices since they are clear from the context.

NASH AND SADDLE-POINT EQUILIBRIA 285  $a_k = B_k [I + Z_{k+1} \{B_1 B_1' - B_2 B_2\}^{-1} Q_{k+1} + Z_{k+1} c_k]$ , (6.25c)  $a_k = -B_k [I + Z_{k+1} (B_1 B_1' - B_2 B_2)^{-1} (a_{k+1} + Z_{k+1} c_k)]$ , (6.25d) and  $F_k$ , given in Corollary 6.3, is then expressed as  $F_k = [I - (B_1 B_1' - B_2 B_2) Z_{k+1}]^{-1} A_k$ . (6.25e) Now, finally, if (6.25a)-(6.25b) are substituted into (6.22a), we arrive at the conclusion that  $Z_k$  satisfies the same equation as  $M_k$ , given by (6.12b), and hence  $F_k = A^{-1} A_k$ , and furthermore  $f_k$  generated by (6.22b) is identical with  $m_k$  generated by (6.13f). Therefore, we have the following. Theorem 6.7 The two-person affine-quadratic zero-sum dynamic game described by (6.10a)-(6.10b) admits a unique feedback saddle-point solution if, and only if,  $I + B_1 M_{k+1} B_1 > 0$  ( $k \in K$ ), (6.26a)  $B_1 M_{k+1} B_1 > 0$  ( $k \in K$ ), (6.26b) in which case the unique equilibrium strategies, which are also strongly time consistent, are given by  $7^*(x^k) = -B_1' M_{k+1} A^{-1} [A_k x^k + c_k] - B_2' (A' f_c)_{k+1} C_{k+1}^{-1} f_c$ , (6.27a)  $7^f(x^k) = B_2 M_{k+1} k^{-1} [A_k x^k + c_k + B_1' C_{k+1}^{-1} f_c]$ ,  $k \in K$ , (6.27b) where  $O_k = (A^{-1})' (h_k + M_{k+1} c_k)$ ,  $C_{k+1} = 0$ , and  $M_{k+1}$ ,  $A^{-1} (A; k \in K)$  are as defined in Thm. 6.4- The corresponding unique state trajectory  $\{x^{k+1}; k \in K\}$  satisfies the difference equation  $x_{t+1} = K X [A_k x_t + V_k]$ ,  $x_t = x_u$  (6.27c) where  $r_k$  ( $k \in K$ ) is as defined in Thm. 6.4, with  $m_{k+1}$  replaced by the equivalent vector  $O_{t+1}$ . The corresponding saddle-point value is  $J^* = -f(7_1^*, 7_2^*) = -x_i' M_i X_i + C_1 X_1 + n_x$ . Furthermore, if  $Q_{k+i} > 0 \forall i; i \in K$ , then  $M_{k+i} > 0 \forall i; i \in K$ , and hence condition (6.26a) becomes superfluous.

286 T. BAŞAR AND G. J. OLSDER Proof. This result, with the exception of the last statement, follows from Corollary 6.1, Corollary 6.3 and the discussion given prior to the statement of the theorem. To prove the last statement, take  $c_k = 0$ , without any loss of generality, and note that with  $Q^{k+i} > 0$  the lower value of any stage-truncated (from below) version of the game is

nonnegative (choose, e.g.,  $u = 0$ ), and hence the feedback saddle-point value of each such truncated game, which is  $\|x_k\| M_k X_k$ , is nonnegative for any arbitrary  $x_k \in \mathbb{R}^n$ . Hence,  $M_k > 0$  for all  $k \in K$ .<sup>56</sup> D Remark 6.6 As in the case of Nash equilibria (cf. Remark 6.3), the result above extends naturally to the more general affine-quadratic zero-sum dynamic games which have in the cost function an additional linear term in  $i$ , to be denoted  $x_{k+1} l_{k+1}^i$ . Then, the only modification will have to be made in the equation for  $\lambda$ , which will now read  $O_t = (A^{-1} A_k)'(C_{k+1}^i + M_{k+1} c_k) + l_k$ ,  $C_{k+1}^i = I_{k+1}^i$ . Otherwise, the statement of Thm. 6.7 remains intact.  $\square$  Remark 6.7 A comparison of Thms. 6.4 and 6.7 now readily reveals an important property of the saddle-point solution in such games; namely, whenever they both exist, the unique open-loop saddle-point solution and the unique feedback saddle-point solution generate the same state trajectory in affine-quadratic zero-sum games. Furthermore, the open-loop values (cf. Section 5.6) of the feedback saddle-point strategies (6.27a)-(6.27b) are correspondingly the same as the open-loop saddle-point strategies (6.13a)-(6.13b). These two features of the saddle-point solution under different information structures are in fact characteristic of not only affine-quadratic games, but of the most general zero-sum dynamic games treated in this section, as it will be verified in the next section (see, in particular, Thm. 6.9). In nonzero-sum dynamic games, however, the Nash equilibrium solutions obtained under different information structures do not exhibit such a feature, as it will also be clear from the discussion of the next section (see, in particular, subsection 6.3.1).  $\square$  Even though the two saddle-point solutions of Thms. 6.4 and 6.7 generate the same state trajectory, the existence conditions involved are not equivalent—a result that follows from a comparison of (6.11a) and (6.14b) with (6.26b) and (6.26a). This point is further elaborated on in the following proposition for the case  $Q_{k+1} > 0$ , under which (6.14b) and (6.26a) are automatically satisfied, which therefore leaves only (6.11a) and (6.26b) for comparison. Proposition 6.2 For the affine-quadratic two-person zero-sum dynamic game described by (6.10a)-(6.10b) and with  $Q_{k+1} > 0 \forall A; \forall K$ , condition (6.11a) implies (6.26b), but not vice versa. In other words, every affine-quadratic two-person zero-sum dynamic game (with nonnegative cost on the state) that admits <sup>56</sup>It is possible to verify this property of  $M_k$  also by direct matrix manipulations; see Prop. 6.2 later for a hint in this direction.

NASH AND SADDLE-POINT EQUILIBRIA 287 a unique open-loop saddle point also admits a unique feedback saddle point, but existence of the latter

does not necessarily imply existence of a saddle point in the open-loop strategies. Proof. The proof of the implication from (6.11a) to (6.26b) is by induction (on  $k$ ), where we have to show that for arbitrary  $k \in K$ ,  $S_{k+1} > M_{k+1} \Rightarrow S_k > M_k$ . To save from indices, let  $S_{k+1} = S$ ,  $M_{k+1} = M$ , and likewise for  $B_1$  and  $B_2$ , and start with the relation  $S > M > 0$ , where the nonnegative definiteness of  $M$  is a property that was already proven in Thm. 6.7. Consider the perturbed matrices  $S_\epsilon = S + \epsilon I$ ,  $M_\epsilon = M + \epsilon I$ ,  $\epsilon > 0$ , both of which are positive definite for each  $\epsilon > 0$ . The following sequence of implications ( $\Rightarrow$ ) and equivalences ( $\Leftrightarrow$ ) now follows from standard properties of matrices:  $S_\epsilon > M_\epsilon \Rightarrow S_\epsilon - B_2 B_2' > M_\epsilon - B_2 B_2' + B_1 B_1' \Leftrightarrow S_\epsilon (I - B_2 B_2' S_\epsilon^{-1}) > M_\epsilon \{I + (B_1 B_1' - B_2 B_2') M_\epsilon^{-1}\}$ . Since the last inequality holds  $\forall \epsilon > 0$ , and both sides are well defined as  $\epsilon \rightarrow 0$ , we finally have (using the continuous dependence of eigenvalues of the matrices above on the parameter  $\epsilon$ )  $(S - B_2 B_2')^{-1} > M (I + (B_1 B_1' - B_2 B_2') M)^{-1}$  and pre- and post-multiplying both sides by  $A'$  and  $A$  and then adding  $Q$  yields the desired result  $S_k > M_k$ . This completes the proof by induction, since for  $k = K + 1$  both matrices equal  $Q_{k+1}$ . To verify that (6.26b) does not necessarily imply (6.11a), we simply produce a counter example. Consider the scalar two-stage game (i.e.,  $K = 2$ ) with  $Q_{k+1} = B_k = A_k = 1$ ,  $B_k = 1/\sqrt{2}$ , ( $k = 1, 2$ ). Both (6.11a) and (6.26b) are satisfied for  $k = 2$ . For  $k = 1$ , however,  $S_2 = 3$ ,  $M_1 = 5/3$ , and while (6.26b) is satisfied, the condition (6.11a) fails to hold true. **D Remark 6.8** The condition (6.26b) of Thm. 6.7 is quite tight for affine-quadratic dynamic games (with nonnegative cost on the state) under the CLPS information structure, in the sense that if the matrix in (6.26b) has any negative eigenvalues, then the upper value of the game becomes unbounded. This is <sup>57</sup>For two positive definite matrices  $V$  and  $W$ , the inequality  $V > W$  means that the matrix difference  $V - W$  is nonnegative definite, which also implies that the difference of their inverses,  $V^{-1} - W^{-1}$ , is nonpositive definite. For a proof of this last inequality, see, for example, Bellman (1970).

288 T. BAŞAR AND G. J. OLSDER easy to see when PI uses only the current value of the state, because then existence of a negative eigenvalue for the matrix in (6.26b) at some stage  $k$  would imply that the recursive derivation of the feedback saddle-point equilibrium solution encounters at stage  $k$  a static game that is not concave in the maximizing variable—making the upper value of that game unbounded (as in Remark 4.7). This argument can be extended also to the general CLPS information structure, implying that even if the

minimizer is allowed to use also the past values of the state the upper value would still be unbounded when the matrix in (6.26b) has at least one negative eigenvalue.<sup>58</sup> The lower value of the game, however, could still be bounded, which means that if the minimizer is also allowed to have access to the maximizer's actions (controls), then condition (6.26b) can be further relaxed; this relaxed version of (6.26b) can in fact be obtained quite readily by allowing (in the recursive derivation of the feedback saddle-point equilibrium) the minimizer (PI) to choose  $u_j$ , dependent not only on  $X^f$  but also on  $u^{\setminus}$ , and then requiring the resulting minimum value of the stagewise cost (in the truncated game) to be strictly concave in  $u^{\setminus}$ . If the information structure is CLPS, and the matrix in (6.26b) is nonnegative (but not positive) definite, then whether the upper value is bounded or not depends on the specific structures of the cost and system matrices, which we do not further discuss here.  $\square$

**6.2.3 Linear-quadratic games with an infinite number of stages** We now consider stationary dynamic games with an infinite number of stages, and restrict attention to linear-quadratic structures, which means that in the formulation of Def. 6.1, all matrices are time invariant,  $c = 0$ , and  $K \rightarrow \infty$  or  $K \rightarrow \infty$ . Feedback Nash equilibria of such games can be obtained in two different ways: as the limit of the feedback Nash solution of any time-truncated version (with, say,  $K$  stages) as the number of stages ( $K$ ) goes to infinity, or from the outset as the Nash equilibrium of an infinite-horizon game under CLNM information pattern. Of course, a natural procedure to follow would be a combination of these two methods: find the limit of the finite-horizon solution as  $K \rightarrow \infty$ , and then verify that this limiting solution (if it exists) provides a Nash equilibrium solution for the infinite-horizon game. Toward this end, first note that a candidate solution can easily be obtained from Corollary 6.1, by simply taking  $c = 0$  and dropping the time indices from the various matrices. Let  $z^k(i \in N)$  denote the solution of (6.17b) when there are  $K$  stages (where we show explicit dependence on  $k$ , as in the one-player case discussed in subsection 5.5.1, Prop. 5.2), and let  $Z^0(i \in N)$  be its limit as  $K \rightarrow \infty$  for fixed  $k$  (assuming that this limit exists, and is independent of  $A$ ). Likewise, let  $P_i(i \in N)$  denote the limit of  $P^k(i \in N)$  as  $K \rightarrow \infty$ . Then, these limiting See Basfar (1991b) or Başar and Bernhard (1995) for details of this argument.

**NASH AND SADDLE-POINT EQUILIBRIA 289** matrices necessarily satisfy the following two algebraic matrix equations  $[RH + B_i Z_i B_i' P_i + B_i' Z_i J_2 B_i P_i J = S^* Z M, i \in N, (6.28a) p = F' P F + J_2 P_i' R_{ii} P_i + Q_i; i \in N, (6.28b)$  where  $F = A$

-  $J_2 B_i^{-1}$  (6.28c) The corresponding limiting feedback strategy for  $P_i$  is then (from (6.18a))  $\gamma_i^*(x^k) = -P_i x^k$  ( $i \in N, A_i = 1, 2, \dots$ ). (6.29) Conditions that will guarantee that such an  $N$ -tuple of stationary policies are in Nash equilibrium can be obtained readily from Prop. 5.2, by simply holding all but one (say,  $i$ th) players' strategies fixed at (6.29) and requiring that the resulting infinite-horizon optimal control problem be well defined, as explained in subsection 5.5.1. This result is given below in Prop. 6.3 after introducing some notation. Let  $F_i$  and  $Q_i$  ( $i \in N$ ) be defined by  $F_t = A - J_2 B_i P_i \backslash Q_t * Q \{ + \wedge^2 P_j' \&gt; P_j$  ( $i \in N$ ).  $J_i^* J_i'$  Proposition 6.3 Let there exist two  $N$ -tuples of matrices  $\{Z_i, P_i \mid i \in N\}$  satisfying (6.28a) and (6.28b), and further satisfying the conditions that for each  $i \in N$  the pair  $(F_i, B_i)$  is stabilizable and the pair  $(F_i, Q_i)$  is detectable. Then, (i) the  $N$ -tuple of stationary feedback policies (6.29) provides a Nash equilibrium solution for the linear-quadratic nonzero-sum dynamic game of this subsection, leading to the finite infinite-horizon Nash equilibrium cost  $\pm x^1 Z_i x^1$  for  $P_i$ , (ii) the resulting system dynamics, described by  $x^{k+1} = F x^k$ ,  $A_i = 1, 2, \dots$ , are stable. A few remarks regarding the solution presented above are now in order. First, we have not given any conditions on the parameters of the game that will guarantee the existence of a solution set to (6.28a)-(6.28b); obtaining such

290 T. BAŞAR AND G. J. OLSDER conditions seems to be quite a challenging task. Second, even though the set of equations (6.28a)-(6.28b) was obtained by taking a limit on the solution for the time-truncated version of the game, it is quite possible that these equations will admit (other) solutions that are not necessarily related to the solution(s) of the finite-horizon game. But, these would also provide Nash equilibria, as long as the conditions on stabilizability and detectability are satisfied. For the special class of linear-quadratic zero-sum games it is possible to obtain stronger results. First, let us rewrite the "stationary" counterparts of the feedback saddle-point policies (6.27a)-(6.27b), along with the "stationary" counterparts of the relevant matrices ( $A$  and  $M$ ) which were introduced in Thm. 6.4, where we again use an "overbar" to denote the limiting values (as  $K \rightarrow \infty$ ):  $\bar{K}^1(x^k) = -B^{-1} M A^{-1} A x^k$ , (6.30a)  $\bar{K}^2(x^k) = B^{-2} M A^{-1} A x^k$ ,  $k = 1, 2, \dots$ ; (6.30b)  $\bar{A} = [I + \{B_1 B_1' - B_2 B_2'\} M]$ , (6.30c)  $\bar{M} = Q + A' M A$ . (6.30d) Furthermore, the counterpart of the concavity condition (6.26b) is  $I - B_2' M B_2 > 0$ , (6.30e) and for convenience we also introduce the following relaxed version of (6.30e):  $I - B_2' M B_2 > 0$ . (6.30f) Now, the following lemma, whose proof can be found in (Başar and Bern- hard,

1995), provides precise relationships between the limit of the sequence generated by (6.12b) and solutions of (6.30d) as well as the value of the underlying infinite-horizon game, when  $Q > 0$ . Before stating the lemma, let us introduce the notation  $M_k$  to denote the unique nonnegative-definite matrix sequence generated by (6.12b) for a if-stage time-invariant linear-quadratic game, with the terminal condition (at  $k = K + 1$ ) taken as  $Q$ , and with condition (6.30f) satisfied. Lemma 6.3 For the linear-quadratic infinite-horizon zero-sum dynamic game with  $Q > 0$ , let the pair  $(A, Q)$  be observable (respectively, detectable). Then, (i) if there exists a nonnegative-definite solution to the generalized algebraic Riccati equation (GARE) (6.30d),  $M$ , which also satisfies (6.30e), then  $M$  is necessarily positive definite, (ii) if the GARE (6.30d) does not admit a positive (respectively, nonnegative) definite solution satisfying (6.30f), then the upper value of the game is unbounded,

NASH AND SADDLE-POINT EQUILIBRIA 291 (Hi) if  $M$  denotes a positive (respectively, nonnegative) definite solution of (6.30d) satisfying (6.30e), then for all  $K > 1$ ,  $M > M(kK) > M^+$ , for all  $k < K$ . An important consequence of this lemma is that if the GARE (6.30d) admits multiple positive (respectively, nonnegative) definite solutions satisfying (6.30e), there is a minimal such solution (minimal in the sense of matrix partial ordering), denoted  $M^+$ , to which the sequence  $\{M_k\}$  converges as  $K \rightarrow \infty$ . In view of this observation, and the result of Lemma 6.3, the following theorem can be established. Its proof can be found in Başar (1991b) and Başar and Bernhard (1995). Theorem 6.8 Consider the infinite-horizon discrete-time linear-quadratic zero-sum dynamic game, with  $(A, Q)$  constituting an observable (respectively, detectable) pair. Then, we have the following. (i) The game has equal upper and lower values if, and only if, the GARE (6.30d) admits a positive (respectively, nonnegative) definite solution satisfying (6.30e), and only if it admits a positive (respectively, nonnegative) definite solution satisfying (6.30f). (ii) If the GARE admits a positive (respectively, nonnegative) definite solution, satisfying (6.30e), then it admits a minimal such solution, to be denoted  $M^+$ . Then, the finite value of the game is  $x^T M^+ x$ . (Hi) The upper (minimax) value of the game is finite if, and only if, the upper and lower values are equal. (iv) If  $M^+ > 0$  (respectively,  $> 0$ ) exists, as given above, the controller  $u^*$  given by (6.30a), with  $M$  replaced by  $M^+$ , attains the finite upper value, in the sense that  $\sup_{u^*} (J^*(u^*, v^*)) = x^T M^+ x$ , and a maximizing solution above is the stationary feedback policy given by (6.30b), again with  $M$

replaced by  $M^+$ . (v) If  $M^+ > 0$  (respectively,  $> 0$ ), exists, the closed-loop system under the policies (6.30a)-(6.30b), with  $M$  replaced by  $M^+$ , is asymptotically stable, that is, the matrix  $F$  defined below is Hurwitz:  $F = (I - (BB' - DD^M + iA + y^A$ . Note that the theorem above does not say that the feedback policies (6.30a)- (6.30b), with  $M$  replaced by  $M^+$ , are in saddle-point equilibrium. It only says that  $P_i$ 's feedback policy as given assures a finite upper value, which is also

292 T. BAŞAR AND G. J. OLSDER equal to the lower value. This finite value of the game may not, however, be assured by the corresponding feedback policy of  $P_2$ ; more will be said on this point in the context of continuous-time (differential) games, to be discussed in subsection 6.5.3.

### 6.3 Informational Properties of Nash Equilibria in Discrete-Time Dynamic Games

This section is devoted to an elaboration on the occurrence of "informationally nonunique" Nash equilibria in discrete-time dynamic games, and to a general discussion on the interplay between information patterns and existence-uniqueness properties of noncooperative equilibria in such games. First, we consider, in some detail, a scalar three-person dynamic game which admits uncountably many Nash equilibria, and which features several important properties of infinite dynamic games. Then, we discuss these properties in a general context.

#### 6.3.1 A three-person dynamic game illustrating informational nonuniqueness

Consider a scalar three-person two-stage linear-quadratic dynamic game in which each player acts only once. The state equation is given by  $x_3 = x_2 + u_1 + u_2$ ;  $x_2 = x_1 + u_3$ , (6.31) and the cost functionals are defined as  $L_1 = (x_3)^2 + (u_1)^2$ ;  $L_2 = -L_3 = -(x_3)^2 + 2(u_2)^2 - (u_3)^2$ . (6.32) In this formulation,  $u_i$  is the scalar unconstrained control variable of  $P_i$  ( $i = 1, 2, 3$ ), and  $x_1$  is the initial state whose value is known to all players.  $P_1$  and  $P_2$ , who act at stage 2, have also access to the value of  $x_2$  (i.e., the underlying information pattern is CLPS (or, equivalently, MPS) for both  $P_1$  and  $P_2$ ), and their permissible strategies are taken as twice continuously differentiable mappings from  $\mathbb{R} \times \mathbb{R}$  into  $\mathbb{R}$ . A permissible strategy for  $P_3$ , on the other hand, is any measurable mapping from  $\mathbb{R}$  into  $\mathbb{R}$ . This, then, completes the description of the strategy spaces  $T_1$ ,  $T_2$  and  $T_3$  (for  $P_1$ ,  $P_2$  and  $P_3$ , respectively), where we suppress the subindices denoting the corresponding stages since each player acts only once. Now, let  $\{ \gamma_1 \in T_1, \gamma_2 \in T_2, \gamma_3 \in T_3 \}$  denote any noncooperative (Nash) equilibrium solution for this three-person nonzero-sum dynamic game. Since  $L_i$  is strictly convex in  $u_i$  ( $i = 1, 2$ ), a set of necessary and sufficient conditions for  $\gamma_1$  and



72 to be in equilibrium (with  $73 \in T_3$  fixed) is obtained by differentiation of  $L_1$  with respect to  $u_1$  ( $i = 1, 2$ ), thus leading to  $X_2(-y_3) + 2'y_1(x_2, x_1) + y_2(x_2, x_1) = 0$ ,  $-S_2(73) - 71(\wedge^2, a; i) + 72(S_2, \wedge^i) = 0$ ,

NASH AND SADDLE-POINT EQUILIBRIA 293 where Solving for  $71(5; 2, :e_1)$  and  $*y_2(x_2, x_1)$  from the foregoing pair of equations, we obtain  $2 \text{ ll}\{x_2, x_1\} = -x_2$ , (6.33a)  $72(x_2, x_1) = \wedge^2 x_2$ , (6.33b) which are the side conditions on the equilibrium strategies  $71$  and  $72$ , and which depend on the equilibrium strategy  $73$  of  $P_3$ . Besides these side conditions, the Nash equilibrium strategies of  $P_1$  and  $P_2$  have no other natural constraints imposed on them. To put it in other words, every Nash equilibrium strategy for  $P_1$  will be a closed-loop representation (cf. Def. 5.12) of the open-loop value (6.33a), and every Nash equilibrium strategy for  $P_2$  will be a closed-loop representation of (6.33b). To complete the solution of the problem, we now proceed to stage 1. Since  $\{71, 72, 73\}$  constitutes an equilibrium triple, with  $\{u_1 = j_1(x_2, X_i), u_2 = \sim/2(x_2, X_i)\}$  substituted into  $L_3$ , the resulting cost functional of  $P_3$  (denoted as  $L_3$ ) should attain a minimum at  $u_3 = 73(a; 1)$ ; and since  $71$  and  $72$  are twice continuously differentiable in their arguments, this requirement can be expressed in terms of the relations  $A_i^3(73(a; i)) = X_3(1 + 71_2 + 72_j - 2\wedge^2 + 73(x_i) = Q_j \wedge^i L_3(73(x_1)) = (1 + 1\wedge^2 + 11? + x_3(j_1 2 X_2 + 72 2 I_2) - 2(72 2) 2 - 272 2 a; 272 + 1 > 0$ , 59 where we have suppressed the arguments of the strategies. Now, by utilizing the side conditions (6.33a)-(6.33b) in the above set of relations, we arrive at a simpler set of relations which are, respectively,  $-[*! + 73(*i)][1 + 112(*2, X_i)] + 73(x_0 = 0$ , (6.34a)  $2 \text{ II} + 7*2(*2, *i) + i12\{x_2, xx\}\} 2 - x^\wedge 1 x^\wedge xx) - 2[722(i_2, z_i)] 2 + 1 > 0$ , (6.34b) 59 Here, we could of course also have nonstrict inequality (i.e.,  $>$ ) in which case we also have to look at higher-order derivatives of  $L_3$ . We avoid this by restricting our analysis at the outset only to those equilibrium triples  $\{71, 72, 73\}$  which lead to an  $L_3$  that is locally strictly convex at the solution point.

294 T. BAŞAR AND G. J. OLSDER where  $x_2$  is again denned as  $x_2 = x_1 + j_3(x_1)$ . (6.34c) These are the relations which should be satisfied by an equilibrium triple, in addition to (6.33a) and (6.33b). The following proposition summarizes the result. Proposition 6.4 Any triple  $\{71^* \in T_1 \wedge^2 * \in r_2, 73^* \in T_3\}$  that satisfies (6.33a)-(6.33b) and (6.3^a)-(6.3J^b), and also possesses the additional feature that (6.34a) with  $71 = 71^*$  and  $j_2 = j_2^*$  admits

a unique solution  $z_3 = z_3^*$ , constitutes a Nash equilibrium solution for the nonzero-sum dynamic game described by (6.31)-(6.32). Proof. This result follows from the derivation outlined prior to the statement of the proposition. Uniqueness of the solution of (6.34a) for each pair  $\{z_1^*, z_2^*\}$  is imposed in order to ensure that the resulting  $z_3^*$  is indeed a globally minimizing solution for  $I$ . D We now claim that there exists an uncountable number of triplets that satisfy the requirements of Prop. 6.4. To justify this claim, and to obtain a set of explicit solutions, we consider the class of  $z_1$  and  $z_2$  described as  $z_1(x_2, a; 1) = -x_2 + p[x_2 - x_2(-y_3)]$ ,  $z_2(x_2, x_1) = -x_2 + q[x_2 - x_2(\wedge)]$ , where  $p$  and  $q$  are free parameters. These structural forms for  $z_1$  and  $z_2$  clearly satisfy the side conditions (6.33a) and (6.33b), respectively. With these choices, (6.34a) can be solved uniquely (for each  $p, q$ ) to give  $z_3^*(1) = -[(2 + 6p)/(11 + 6p)]x_1 - 11/6$ , with the existence condition (6.34b) reading  $(1 + p) + 2pq - q^2 + 1 > 0$ . (6-35) The scalar  $x_2$  is then given by  $x_2 = [9/(11 + 6p)]x_1$ . Hence, we have the following. Proposition 6.5 The set of strategies  $z_1^*(x_2, a; 1) = -x_2 + p\{x_2 - [9/(11 + 6p)]x_1\}$ ,  $z_2^*(x_2, x_1) = -x_2 + q\{x_2 - [9/(11 + 6p)]x_1\}$ ,  $z_3^*(x_1, x_2) = -\{(2 + 6p)/(11 + 6p)\}x_1$

NASH AND SADDLE-POINT EQUILIBRIA 295 constitutes a Nash equilibrium solution for the dynamic game described by (6.31)- (6.32), for all values of the parameters  $p$  and  $q$  satisfying (6.35) and with  $p \wedge -11/6$ . The corresponding equilibrium costs of the players are  $J_u = 2[6/(11 + 6p)]^2(x_1)^2$ ,  $J_2 = -J_3^* = -[(22 + 24p + 36p^2)/(11 + 6p)^2](x_1)^2$ . Several remarks and observations are in order here, concerning the Nash equilibrium solutions presented above. (1) The nonzero-sum dynamic game of this subsection admits uncountably many Nash equilibrium solutions, each one leading to a different equilibrium cost triple. (2) Within the class of linear strategies, Prop. 6.5 provides the complete solution to the problem, which is parameterized by  $p$  and  $q$ . (3) The equilibrium strategy of  $P_3$ , as well as the equilibrium cost values of all three players, depend only on  $p$  (not on  $q$ ), whereas the existence condition (6.35) involves both  $p$  and  $q$ . There is indeed an explanation for this: the equilibrium strategies of  $P_1$  and  $P_2$  are in fact representations of the open-loop values (6.33a)-(6.33b) on appropriate trajectories. By choosing a specific representation of (6.33a),  $P_1$  influences the cost functional of  $P_3$  and thereby the optimization problem faced by him. Hence, for each different representation of (6.33a),  $P_3$  ends up, in general, with a different solution to his minimization problem, which directly contributes to nonuniqueness of Nash

equilibria. For P2, on the other hand, even though he may act analogously—i.e., choose different representations of (6.33b)—these different representations do not lead to different minimizing solutions for P3 (but instead affect only the existence of a minimizing solution) since  $L_2 = -L_3$ , i.e., P2 and P3 have completely conflicting goals (see Thm. 6.9, later in this section, and also the next remark for further clarification). Consequently,  $J_3^*(x_1)$  is independent of  $q$ , but the existence condition explicitly depends upon  $q$ . (This is true also for nonlinear representations, as it can be seen from (6.34a) and (6.34b).) (4) If  $P_i$  has access to only  $x^i$  (and not to  $X_i$ ), then, necessarily,  $p = 0$ , and both  $P_i$  and P3 have unique equilibrium strategies which are  $\{x^i, J_3^*(x_1) = -(2/11)c_i\}$ . (This is true also within the class of nonlinear strategies.) Furthermore, the equilibrium cost values are also unique (simply set  $p = 0$  in  $J_1^*, i = 1, 2, 3$ , in Prop. 6.5). However, the existence condition (6.35) still depends on  $q$ , since it now reduces to  $q^2 < 11/9$ . The reason for this is that P2 still has the freedom of employing different representations of (6.33b), which affects existence of the equilibrium solution but not the actual equilibrium state trajectory, since P2 and P3 are basically playing a zero-sum game (in which case the equilibrium (i.e., saddle-point) solutions are interchangeable). (See Thm. 6.9, later in this section; also recall the feature discussed in Remark 6.7 earlier.)

296 T. BAŞAR AND G. J. OLSDER (5) By setting  $p = q = 0$  in Prop. 6.5, we obtain the unique feedback Nash equilibrium solution (cf. Corollary 6.1 and Remark 6.4) of the dynamic game under consideration (which exists since  $p = 0 < 11/6$ , and (6.35) is satisfied). (6) Among the uncountable number of Nash equilibrium solutions presented in Prop. 6.5, there exists a subsequence of strategies which brings  $P_i$ 's Nash cost arbitrarily close to zero which is the lowest possible value  $L_1$  can attain. Note, however, that the corresponding cost for P3 approaches  $(x^1)^2$  which is unfavorable to him. Before concluding this subsection, it is worthy to note that the linear equilibrium solutions presented in Prop. 6.5 are not the only ones that the dynamic game under consideration admits, since (6.34a) will also admit nonlinear solutions. To obtain an explicit nonlinear equilibrium solution, we may start with a nonlinear representation of (6.33a), for instance,  $2 \cdot 11 \{x_2, x_1\} = -x_2 + p[x_2 - x_2^{(1)}]^2$ , substitute it into (6.34a) and solve for a corresponding  $J_3(x_1)$ , checking at the same time satisfaction of the second-order condition (6.34b). For such a derivation (of nonlinear Nash solutions in a linear-quadratic game) the reader

is referred to Başar (1974). See also Problem 5, Section 6.8. 6.3.2 General results on informationally nonunique equilibrium solutions Section 3.5 has already displayed existence of "informationally nonunique" Nash equilibria in finite multi-act nonzero-sum dynamic games, which was mainly due to the fact that an increase in information to one or more players leaves the Nash equilibrium obtained under the original information pattern unchanged, but it also creates new equilibria (cf. Prop. 3.10). In infinite games, the underlying reason for occurrence of informationally nonunique Nash equilibria is essentially the same (though much more intricate), and a counterpart of Prop. 3.10 can be verified. Toward this end we first introduce the notion of "informational inferior" in such dynamic games. Definition 6.3 Let I and II be two N-person K-stage infinite dynamic games which admit precisely the same extensive form description except the underlying information structure (and, of course, also the strategy spaces whose descriptions depend on the information structure). Let  $r_i \setminus$  (respectively,  $r_i^{\wedge}$ ) denote the information pattern of  $P_i$  in the game I (respectively, II), and let the inclusion relation  $r_i \setminus \subset r_i^{\wedge}$  imply that whatever  $P_i$  knows at each stage of game I he also knows at the corresponding stages of game II, but not necessarily vice versa. Then, I is informationally inferior to II if  $r_i \setminus \subset r_i^{\wedge}$  for all  $i \in N$ , with strict inclusion for at least one  $i$ .

NASH AND SADDLE-POINT EQUILIBRIA 297 Proposition 6.6 Let I and II be two N-person K-stage infinite dynamic games as introduced in Def. 6.3, so that I is informationally inferior to II. Furthermore, let the strategy spaces of the players in the two games be compatible with the given information patterns and the constraints (if any) imposed on the controls, so that  $r_i \setminus \subset r_i^{\wedge}$  implies  $Y_i \subset T_{ij}$ ,  $i \in N$ . Then, (i) any Nash equilibrium solution for I is also a Nash equilibrium solution for II, (ii) if  $\{x_1^*, \dots, x_N^*\}$  is a Nash equilibrium solution for II such that  $x_i^* \in F_i$  for all  $i \in N$ , then it is also a Nash equilibrium solution for I. Proof. Let  $\{x_i^*/ \in N\}$  constitute a Nash equilibrium solution for I. Then, by definition,  $J_1(x_1^*, x_2^*, \dots, x_N^*) < J_1(x_1, x_2^*, \dots, x_N^*), \forall x_1 \in T_{11}$ ; therefore,  $P_1$  minimizes  $J_1(-, x_2^*, \dots, x_N^*)$  over  $T_{1j}$ , with the corresponding solution being  $x_1^* \in T_{1j}$ . Now consider minimization of the same expression over  $T_{1j} \supset T_{1j}$  which reflects an increase in deterministic information concerning the values of state. But, since we have a deterministic optimization problem, the minimum value of  $J_1(x_1, x_2^*, \dots, x_N^*)$  does not change with an increase in information (see Section 5.6). Hence,  $\min J_1(x_1, x_2^*, \dots, x_N^*) = J_1(x_1^*, x_2^*, \dots, x_N^*)$ ; and furthermore, since  $x_1^* \in T_{1j}$ , we have the inequality  $J_1(x_1^*, x_2^*, \dots, x_N^*) < J_1(x_1, x_2^*, \dots, x_N^*)$ .

$\langle J_1(7_1, 7_2^*, \dots, 7_i V^*), V_1 \rangle$ . Since PI was an arbitrary player in this discussion, it follows in general that  $w, \dots, *, \dots, *$   $\langle J_i(7_1^*, \dots, 7_{i-1}^*, y, 7_{i+1}^*, \dots, 7_V^*) \rangle$   $v_1 \in \text{rfr}(i \in n)$  which verifies (i) of Prop. 6.6. Proof of (ii) is along similar lines. D Since there corresponds at least one informationally inferior game (viz. a game with an open-loop information structure) to every multi-stage game with CLPS information, the foregoing result clearly provides one set of reasons for existence of "informationally nonunique" Nash equilibria in infinite dynamic games (as Prop. 3.10 did for finite dynamic games). However, this is not yet the whole story, as it does not explain occurrence of uncountably many equilibria in such games. What is really responsible for this is the existence of uncountably many representations of a strategy under dynamic information (cf. Section 5.6). To elucidate somewhat further, consider the scalar three-person dynamic game of subsection 6.3.1. We have already seen that, for each fixed equilibrium strategy  $\gamma^3$  of P3, the equilibrium strategies of P1 and P2 have unique open-loop

298 T. BAŞAR AND G. J. OLSDER values given by (6.33a) and (6.33b), respectively, but they are otherwise free. We also know from Section 5.6 that there exist infinitely many closed-loop representations of such open-loop policies; and since each one has a different structure, this leads to infinitely many equilibrium strategies for P3, and consequently to a plethora of Nash equilibria. The foregoing discussion is valid on a much broader scale (not only for the specific three-person game treated in subsection 6.3.1), it provides a general guideline for derivation of informationally nonunique Nash equilibria, thus leading to the algorithm given below after introducing the concept of "stagewise equilibrium".

Definition 6.4 For an  $N$ -person  $K$ -stage infinite dynamic game, a set of strategies  $\{\gamma_i^k; i \in N\}$  satisfying the following inequalities for all  $i \in N, k \in K$ , is a stagewise (Nash) equilibrium solution:

$$J_i(7_1^k; \dots; 7_i^k; \dots; 7_A^k) \leq J_i(7_1^k; \dots; 7_i^k, \dots, 7_i^{k-1}, 7_j^k, \dots, 7_j^{k-1}; \dots; 7_A^k). \quad (6.36)$$

Remark 6.9 Every Nash equilibrium solution (cf. the set of inequalities (6.3)) is a stagewise Nash equilibrium solution, but not vice versa.  $\square$

An algorithm to obtain informationally nonunique Nash equilibria in infinite dynamic games

First determine the entire class of stagewise equilibrium solutions. (1) Starting at the last stage of the game ( $k = K$ ), fix the AT-tuples of strategies at every other stage  $k < K$ , and solve basically an AT-person static game at stage  $K$  (which is defined by (6.36) with  $k = K$ ) to determine the open-loop values of the corresponding stagewise equilibrium strategies at  $k = K$  as

functions of the strategies applied previously. Furthermore, determine the equivalence class of representations of these AT-tuples of open-loop values, again as functions of the previous strategies and so that they are compatible with the underlying information structure of the problem. (2) Now consider inequalities (6.36) for  $k = K - 1$  and adopt a specific member of the equivalence class determined at step 1 as a strategy AT-tuple applied at stage  $K$ ; furthermore fix the AT-tuples of strategies at every stage  $k < K - 1$ , and solve basically an AT-person static game to determine the open-loop values of the corresponding stagewise equilibrium strategies at stage  $K - 1$  as functions of the strategies applied previously. Repeat this for every member of the equivalence class determined at step 1. Each will, in general, result in a different set of open-loop values. Now, for each set of open-loop values thus determined, find the corresponding equivalence class of representations which are also compatible with the underlying

NASH AND SADDLE-POINT EQUILIBRIA 299 information structure and which depend on the strategies applied at earlier stages. (K) Now, finally, consider the set of inequalities (6.36) for  $k = 1$ , and adopt specific (and compatible) members of the equivalence classes determined at steps  $1, \dots, K - 1$  as strategies applied at stages  $K, \dots, 2$ ; solve the resulting AT-person static game to determine the corresponding stagewise equilibrium strategies at stage 1, which will in general also require solution of some implicit equations. Repeat this for every compatible  $(K - 1)$ -tuple of members of the equivalence classes constructed at steps  $1, \dots, K - 1$ . This then completes the construction of the entire class of stagewise equilibrium solutions of the given dynamic game. Finally, one has to check to see which of these stagewise equilibrium solutions also constitute Nash equilibria, by referring to the set of inequalities (6.3). Remark 6.10 In obtaining a set of informationally nonunique equilibrium solutions for the scalar example of subsection 6.3.1, we have actually followed the steps of the foregoing algorithm; but there we did not have to check the last condition since every stagewise equilibrium solution was also a Nash equilibrium solution because every player acted only once.  $\square$  The foregoing algorithm leads to an uncountable number of informationally nonunique Nash equilibria in deterministic multi-stage infinite nonzero-sum games wherein at least one player has dynamic information. Furthermore, these informationally nonunique Nash equilibria are, in general, not interchangeable — thereby leading to infinitely many different equilibrium cost  $N$ -tuples. The

reason why we use the words "in general" in the preceding sentence is because there exist some extreme cases of nonzero-sum games which do not exhibit such features. One such class is the so-called team problems for which  $L_1 = L_2 = \dots = L_N$ , i.e., there is a common objective function. The Nash equilibrium solution corresponds in this case to a person-by-person optimal solution and it becomes a team solution under some further restrictions (see the discussion given in Section 4.6). Deterministic team problems are no different from optimal control problems discussed in Section 5.5, and in particular the discussion of Section 5.6 leads to the conclusion that informational nonuniqueness of equilibria does not create any major difficulty in team problems, since all these equilibrium solutions are different representations of the same  $N$ -tuple of strategies which is associated with the global minimum of a single objective functional. Hence, for the special case of team problems, informational nonuniqueness does not lead to different equilibrium cost  $A$ -tuples. Let us now consider the other extreme case—the class of two-person zero-sum infinite dynamic games—in which case  $L_1 = -L_2$ . For such dynamic games, informational nonuniqueness does not lead to different (saddle-point)

300 T. BAŞAR AND G. J. OLSDER equilibrium cost pairs either, since multiple saddle points are interchangeable; however, as we have observed in subsection 6.3.1, existence of a saddle point will depend on the specific pair of representations adopted. Therefore, contrary to team solutions, every representation of a pair of saddle-point strategies is not necessarily a saddle-point strategy pair. The next theorem makes this statement precise, and it also provides a strengthened version of Prop. 6.6 for two-person zero-sum dynamic games. Before giving the theorem, it will be convenient first to introduce the notion of a strongly unique saddle point (say,  $\{71^*, 72^*\}$ ) on a given product strategy space  $T_1 \times T_2$ , as one where the pair  $\{71^*, 72^*\}$  is unique as a saddle-point solution, and in addition  $71^* \in T_1$  is the unique response to  $72^*$  and  $72^* \in T_2$  is the unique response to  $71^*$ . Then, we have the following theorem.

Theorem 6.9 Let  $I$  be a two-person zero-sum  $K$ -stage dynamic game with closed-loop perfect state information pattern. (i) If  $I$  admits a strongly unique feedback saddle-point solution, and also an open-loop saddle-point solution, then the latter is unique. (ii) If  $I$  admits a strongly unique feedback saddle-point solution  $\{71_f, 72_f\}$  with a corresponding state trajectory  $\{x_{k+1}; k \in K\}$ , and  $\{71^*, 72^*\}$  denotes some other closed-loop saddle-point solution with

corresponding state trajectory  $\{x_{k+1}; k \in K\}$ , we have  $x_{k+1} = x_{k+1}^k, k \in K$ ,  $y_k^*(x_1, \dots, x_{k-1}, x_{k+1}) = f_k(x_k, u_k), k \in K, k = 1, 2$ . (ii) If  $\Gamma$  admits a strongly unique open-loop saddle-point solution  $\{\gamma_1^0, \gamma_2^0\}$  with a corresponding state trajectory  $\{x_{k+1}; k \in K\}$ , and if  $\{\gamma_1^*, \gamma_2^*\}$  denotes some other closed-loop saddle-point solution with corresponding state trajectory  $\{x_{k+1}; k \in K\}$  we have  $x_{k+1} = x_{k+1}^k, k \in K, k = 1, 2$ . Proof, (i) Let  $\{\gamma_1^f, \gamma_2^f\}$  be the strongly unique feedback saddle-point solution and  $\{\gamma_1^0, \gamma_2^0\}$  be any open-loop saddle-point solution. It follows from Prop. 6.6 that  $\{\gamma_1^f, \gamma_2^f\}$  and  $\{\gamma_1^0, \gamma_2^0\}$  are also closed-loop saddle-point solutions. Furthermore, because of the ordered interchangeability property of saddle points,  $\{\gamma_1^0, \gamma_2^f\}$  and  $\{\gamma_1^f, \gamma_2^0\}$  are also closed-loop saddle-point strategy pairs,  $J(\gamma_1^0, \gamma_2^f) < J(\gamma_1^f, \gamma_2^0) < J(\gamma_1^f, \gamma_2^f), \forall \gamma_1 \in T_1, \gamma_2 \in T_2$ , (i) and an analogous pair of inequalities for  $\{\gamma_1^f, \gamma_2^0\}$ , where  $T_1$  denotes the closed-loop strategy space of  $P_1$ . Now, let us consider the RHS inequality of (i), which is equivalent to  $J(\gamma_1^0, \gamma_2^f) = \min_{\gamma_2 \in T_2} J(\gamma_1^0, \gamma_2)$ .

**NASH AND SADDLE-POINT EQUILIBRIA 301** This is a standard optimum control problem in discrete time with a cost functional  $J(x_k, u_k) = \sum_{k=0}^{N-1} g(x_{k+1}, u_k, x_k) + g_N(x_N, u_N, x_N)$ ,  $u_k = \Pi(-)$ , and a state equation  $x_{k+1} = A(x_k, u_k) = f_k(x_k, u_k, \gamma_1^f(x_k))$ . Globally optimal solution of this problem over  $T_1$  can be obtained (if it exists) via dynamic programming and as a feedback strategy. But by hypothesis (since  $\{\gamma_1^f, \gamma_2^f\}$  is a strongly unique saddle-point pair), there exists a unique feedback strategy (namely,  $\gamma_1^f$ ) that solves this problem. Since every closed-loop strategy admits a unique open-loop representation (cf. Section 5.6), it follows that there exists a unique open-loop solution to that optimum control problem, which is given by  $\{x_{k+1}^k = f_k(x_k, u_k, \gamma_1^f(x_k)), k \in K\}$ , where  $\{x_{k+1}^k; k \in K\}$  denotes the state trajectory corresponding to  $\{\gamma_1^f, \gamma_2^f\}$ . Therefore,  $\gamma_1^0$  is unique whenever  $\gamma_1^f$  is. It can likewise be shown, by making use of the LHS inequality of (i), that  $\gamma_2^0(x_1) = \gamma_2^f(x_1)$ , and hence  $\gamma_2^0$  is unique whenever  $\gamma_2^f$  is unique. This then completes verification of part (i). We note that a converse statement cannot be made here (i.e., strong uniqueness of open-loop saddle point may not imply strong uniqueness of feedback saddle point), because there could be cost-irrelevant states on which the feedback solution can be chosen to be nonunique. (ii) By Prop. 6.6 and the ordered interchangeability property of saddle points,  $\{\gamma_1^*, \gamma_2^f\}$  and  $\{\gamma_1^f, \gamma_2^*\}$  are also saddle-point strategy pairs; therefore  $J(\gamma_1^*, \gamma_2^f) = \min_{\gamma_2 \in T_2} J(\gamma_1^*, \gamma_2)$ . This equality defines an optimum control problem in discrete time, and as in the



proof of part (i), every globally optimal solution of this optimization problem can be obtained via dynamic programming; and by hypothesis,  $\gamma_1^f$  is the unique feedback strategy that renders the global minimum. Hence, every solution in  $T_1$  has the same open-loop representation as  $\gamma_1^f$ , that is,  $\{\gamma_1^{f^*}(x^k) = \gamma_1^f(x^k) \mid k \in K\}$ . Similarly, this time by starting with  $\{\gamma_1^f, \gamma_2^*\}$ ,  $\gamma_2^*(x^k) = \gamma_2^f(x^k) \mid k \in K$ . Therefore, the pairs of strategies  $\{\gamma_1^f, \gamma_2^f\}$  and  $\{\gamma_1^*, \gamma_2^*\}$  have the same open-loop representations, and this necessarily implies that the corresponding state trajectories are the same. (iii) The idea of the proof here is the same as that of part (ii). By Prop. 6.6 and the ordered interchangeability property of multiple saddle points,  $\{\gamma_1^*, \gamma_2^o\}$  and  $\{\gamma_1^o, \gamma_2^*\}$  are also saddle-point strategy pairs; the former of these implies that  $J(\gamma_1^*, \gamma_2^o) = \min_{\gamma_1, \gamma_2} J(\gamma_1, \gamma_2)$

302 T. BAŞAR AND G. J. OLSDER which defines an optimization problem whose open-loop solution is known to be unique (namely,  $\gamma_1^o$ ). Therefore, every other solution of this optimization problem in  $T_1$  has to be a closed-loop representation of  $\gamma_1^o$  on the trajectory  $\{x^k \mid k \in K\}$ , i.e.,  $\gamma_1^*(x^k, \dots, x^o, x^i) = \gamma_1^o(x^k) \mid k \in K$ . Similarly,  $\gamma_2^*(x^k, \dots, x^o, x^i) = \gamma_2^o(x^k) \mid k \in K$ . Furthermore, because of this equivalence between the open-loop representations, we have  $x^{k+1} = x^{k+1}$ ,  $k \in K$ . **Remark 6.11** The "uniqueness" assumptions of parts (ii) and (iii) of Thm. 6.9 serve to simplify the statement of the theorem, but do not bring in any loss of generality of conceptual nature. "Uniqueness" can be dispensed with provided that the optimal responses to each strategy in the set of saddle points is also a saddle-point strategy. To make this qualification more precise, and to indicate the modification to be incorporated in (ii) and (iii) of Thm. 6.9 in case of nonunique open-loop or feedback saddle points, suppose that there exist two open-loop saddle-point solutions, say  $\{\gamma_1^o, \gamma_2^o\}$  and  $\{\gamma_1^{oo}, \gamma_2^{oo}\}$ . Then, because of the ordered interchangeability property,  $\{\gamma_1^o, \gamma_2^{oo}\}$  and  $\{\gamma_1^{oo}, \gamma_2^o\}$  will also be saddle-point solutions. Now suppose that  $\gamma_2^o$  and  $\gamma_2^{oo}$  are the only two optimal responses to  $\gamma_1^o$  as well as to  $\gamma_1^{oo}$ , and likewise  $\gamma_1^o$  and  $\gamma_1^{oo}$  are the only two optimal responses to  $\gamma_2^o$  and  $\gamma_2^{oo}$ . Then, if  $\{\gamma_1^*, \gamma_2^*\}$  denotes some closed-loop saddle-point solution, the open-loop representation of this pair of strategies will be equivalent to one of the four open-loop saddle-point strategy pairs. The saddle-point value is, of course, the same for all these different saddle-point solutions.  $\square$  Theorem 6.9, together with Remark 6.11, provides a complete characterization of the saddle-point solutions of two-person zero-

sum discrete-time dynamic games, if their open-loop and/or feedback saddle-point solutions are known; every other saddle point can be obtained as their representation under the given information pattern. Furthermore, whenever they both are strongly unique, the open-loop representation of the feedback saddle-point solution is necessarily the open-loop saddle-point solution. This does not imply, however, that if a feedback saddle point exists, an open-loop saddle point necessarily has to exist (and vice versa); we have already seen in subsection 6.2.2 (cf. Prop. 6.2) that the former statement is not true in annequadratic zero-sum games. In conclusion, a nonzero-sum multi-stage infinite game with dynamic information will, in general, admit uncountably many (informationally nonunique) Nash equilibria which are not interchangeable, unless it is a team problem or a zero-sum game. In Başar and Seluz (1976) it has actually been shown within the context of a class of dynamic two-stage nonzero-sum games that this is in fact a strict property, i.e., unless the nonzero-sum game can be converted into an equivalent team problem or a zero-sum game, it admits uncountably many Nash equilibria. One possible way of removing informational nonuniqueness in the nonco-operative equilibria of nonzero-sum dynamic games is to further restrict the equilibrium solution concept by requiring it to be a feedback Nash equilibrium (cf. Def. 6.2) which has been discussed thoroughly in subsection 6.2.2. However,

NASH AND SADDLE-POINT EQUILIBRIA 303 it should be noted that this restriction (which is compatible with a delayed commitment mode of play or strong time consistency) makes sense only if all players have access to the current value of the state (e.g., the CLPS, MPS or FB information pattern). If at least one of the players has access to delayed or open-loop information, the feedback equilibrium solution cannot be defined, and one then has to resort to some other method for eliminating informational nonuniqueness. One such alternative is to formulate the nonzero-sum game problem in a stochastic framework, which is the subject of the next section.

### 6.4 Stochastic Nonzero-Sum Games with Deterministic Information Patterns

We have already seen in the previous sections that one way of eliminating "informational nonuniqueness" in the Nash solution of dynamic nonzero-sum games is to further restrict the equilibrium concept by requiring satisfaction of a "feedback" property (cf. Def. 6.2) which is, however, valid only under the CLPS and MPS information patterns. Yet another alternative to eliminate informational nonuniqueness is to formulate the dynamic game in a stochastic

framework, in accordance with Def. 5.4; for, under an appropriate stochastic formulation, every strategy has a unique representation (cf. Prop. 5.5) and the statement of Prop. 6.6 is no longer valid. This section is, therefore, devoted to an investigation of existence, uniqueness and derivation of Nash equilibrium in stochastic dynamic games with deterministic information patterns. The class of AT-person If-stage discrete-time stochastic infinite dynamic games to be treated in this section will be as described in Def. 5.4, but with the state equation (5.11) replaced by  $x_{k+1} = f_k\{x_k, u_k, \dots, u^0\} + g_k$ ,  $k \in K$ , (6.37) where  $\{g_k, \dots, g_K\}$  is a set of statistically independent Gaussian random vectors with values in  $R^n$ , and  $\text{cov}(g_k) > 0$ ,  $\forall k \in K$ . Furthermore, the cost functionals of the players are taken to be stage-additive, i.e.,  $\sum_{i \in N} J_i(u^1, \dots, u^N) = \sum_{i \in N} J_i(x_k, u^1, \dots, u^N, x_k)$ ,  $i \in N$ . (6.38) Under the CLPS information pattern, the following theorem now provides a set of necessary and sufficient conditions for any Nash equilibrium solution of such a stochastic dynamic game to satisfy the following Theorem 6.10 Every Nash equilibrium solution of an N-person K-stage stochastic infinite dynamic game described by (6.37)-(6.38) and with the CLPS information pattern (for all players) comprises only feedback strategies; and for an N-tuple  $\{J_i^*(a; f_c), f_c \in K; i \in N\}$  to constitute such a solution, it is

304 T. BAŞAR AND G. J. OLSDER necessary and sufficient that the following recursive relations are satisfied:  $\forall k, x) = \min_{u^k} E \{ g_i \{ f^k \{ x, u^i \} + g_k, 1^k \{ x \}, \dots, 1^i \{ x \}, u^i \} + V_i(k+1, f^k \{ x, u^i \} + g_k) \}$  (6.39)  $J_i^k(x, x) + V_i \{ k+1, f_c(x, u^i(x)) + g_k \}$ ;  $V_i(K+1, x) = 0$ ,  $i \in N$ , where  $f^k \{ x, u^i \} = f_k(x, 1^k \{ x \}, \dots, 1^i \{ x \}, u^i, 1^k \{ x \}, \dots, 1^N \{ x \})$  (6.40) and  $E g_k$  denotes the expectation operation with respect to the statistics of  $g_k$ . Any solution obtained as above is strongly time consistent, and the corresponding expected Nash equilibrium cost for  $P_i$  is  $V^i(1, x_i)$ . Proof The first step is to verify the statement of the theorem for stagewise equilibrium solution (cf. Def. 6.4); hence, start with inequalities (6.36) at stage  $k = K$ , which read  $J_i^k(x, x) + V_i \{ k+1, f_c(x, u^i(x)) + g_k \} < W_i$ ;  $\dots$ ;  $J_i^k(x, x) + V_i \{ k+1, f_c(x, u^i(x)) + g_k \} < W_i$ , where  $J_i^k(x, x) = \sum_{j \in N} J_j^k(x, x) + V_j \{ k+1, f_c(x, u^j(x)) + g_k \}$  and  $0^k = \{g_i, \dots, g_k\}$ ,  $k \in K$ . Because of the stage-additive nature of  $J_i$ , these inequalities equivalently describe an AT-person static game defined by  $\min_{u^k} E \{ g_k \{ f^k \{ x, u^k \} + g_k, y^k \} \}$  (i) where  $x^k$  ( $k \in K$ ) is recursively defined by  $x^k = f_c(x^{k-1}, u^k) + g_k$ . (ii) Now, since  $g_k$  is statistically independent of  $\{g_{k-1}, \dots, g_1\}$ , (i) can equivalently be written as  $E \{ g_k \{ f^k \{ x^k, u^k \} + g_k, y^k \} \}$

$E_{k-1} \{ \mathbb{E}_k [ f_k(x^k) ] \} + \mathbb{E}_k [ \mathbb{E}_{k-1} [ \mathbb{E}_k [ f_k(x^k) ] ] ]$ . (ii)  $\{ \mathbb{E}_k [ f_k(x^k) ] \}_{k=1}^K$ .

**NASH AND SADDLE-POINT EQUILIBRIA** 305 =  $\{ \mathbb{E}_k [ f_k(x^k) ] \}_{k=1}^K$ . which implies that the minimizing  $u^k$  will be a function of  $x^k$ ; furthermore (in contrast with the deterministic version)  $x^k$  cannot be expressed in terms of  $x^{k-1}, \dots, x^1$ . Xi SHCG there is a noise term in (ii) which directly contributes to additional errors in case of such a substitution (mainly because every strategy has a unique representation when the dynamics are given by (ii)). Therefore, under the CLPS information pattern, any stagewise Nash equilibrium strategy for  $P_i$  at stage  $k$  will only be a function of  $x^k$ , i.e., a feedback strategy  $u^k = u^k(x^k)$  which further implies that the choice of such a strategy is independent of all the optimal (or otherwise) strategies employed by the players at previous stages. Finally note that, at stage  $k = K$ , the minimizing solutions of (iii) coincide with those of (6.39) since the minimization problems are equivalent. Let us now consider inequalities (6.36) at stage  $k = K - 1$ , i.e.,  $\mathbb{E}_k [ \mathbb{E}_{k-1} [ \mathbb{E}_k [ f_k(x^k) ] ] ] + \mathbb{E}_k [ \mathbb{E}_{k-1} [ \mathbb{E}_k [ f_k(x^k) ] ] ]$ ,  $\forall k \in \{1, \dots, K-1\}$ , and these can further be written as follows because of the stage-additive nature of  $L$  and since  $\{ u^i \}_{i \in N}$  have already been determined at stage  $K$  as feedback strategies (independent of all the previous strategies):  $\mathbb{E}_k [ \mathbb{E}_{k-1} [ \mathbb{E}_k [ f_k(x^k) ] ] ] + \mathbb{E}_k [ \mathbb{E}_{k-1} [ \mathbb{E}_k [ f_k(x^k) ] ] ]$ . In writing down (iv), we have also made use of the statistical independence property of  $\{ \mathbb{E}_k [ f_k(x^k) ] \}_{k=1}^K$ ,  $\mathbb{E}_k [ f_k(x^k) ]$  and  $\mathbb{E}_k [ f_k(x^k) ]$ . Through a reasoning similar to the one employed at stage  $k = K$ , we readily conclude that any  $\{ u^i \}_{i \in N}$  that satisfies (iv) will have to be a feedback strategy, independent of all the past strategies employed and the past values of the state; therefore  $u^k = u^k(x^k)$  ( $k \in N$ ). This is precisely what (6.39) says for  $k = K - 1$ . Proceeding in this manner, the theorem can readily be verified for stage-wise equilibrium; that is, every set of strategies  $\{ u^i \}_{i \in N}$  satisfying (6.39) constitutes a stagewise equilibrium solution for the stochastic game under consideration, and conversely every stagewise equilibrium solution of the stochastic dynamic game is comprised of feedback strategies which satisfy (6.39). To complete the last phase of verification of the theorem, we first observe that every stagewise equilibrium solution determined in the foregoing derivation is also a feedback Nash equilibrium

solution (satisfying (3.28)) since the construction of stagewise equilibrium strategies at stage  $k = I$  did not depend on

306 T. BAŞAR AND G. J. OLSDER the strategies employed at stages  $k < I$  (note the italicized statements above). Now, since every feedback Nash equilibrium under CLPS information pattern is also a Nash equilibrium (this is a trivial extension of Prop. 6.1 to stochastic dynamic games), and furthermore since every Nash equilibrium solution is a stagewise equilibrium solution (see Remark 6.9), it readily follows that the statement of Thm. 6.10 is valid also for the Nash equilibrium solution. The strong time consistency of the solution (cf. Def. 5.14) follows readily from the constructive nature of the derivation given above. Also, since  $V_i(k, x)$  is the expected cost-to-go function for  $P_i$ ,  $\forall t(1, x_1)$  indeed follows as the expected Nash equilibrium cost for him.  $\square$  Remark 6.12 First, since every strategy admits a unique representation under the state dynamics description (6.37) (with  $\text{cov}(\eta_k) > 0$ ,  $k \in K$ ), the solution presented in Thm. 6.10 clearly does not exhibit any "informational nonuniqueness". Therefore if nonunique Nash equilibria exist in a stochastic dynamic game with CLPS information pattern, this is only due to the structure of  $f_k$  and  $g_k$  and not a consequence of the dynamic nature of the information pattern as in the deterministic case. Second, the Nash equilibrium solution presented in Thm. 6.10 will, in general, depend on the statistical properties (i.e., the mean and covariance) of the Gaussian noise term  $\eta_k$  ( $k \in K$ ). For the special case of affine-quadratic stochastic dynamic games, however, no such dependence on the covariance exists as it is shown in the sequel.  $\square$  As an immediate application of Thm. 6.10, let us consider the class of affine-quadratic stochastic dynamic games wherein  $f_k$  and  $g_k$  are as defined by (6.5a)-(6.5b). Furthermore, assume that  $E[\eta_k] = 0$ ,  $\forall k \in K$ . Because of this latter assumption, and since  $f_k$  is affine and  $g_k$  is quadratic, the minimization problem in (6.39) becomes independent of the covariance of  $\eta_k$ , as far as the optimum strategies are concerned, and a comparison with (6.16) readily leads to the conclusion that they should admit the same solution, which is affine in the current value of the state. Therefore, we arrive at the following corollary which we state here without any further verification. Corollary 6.4 An  $N$ -person affine-quadratic stochastic dynamic game as formulated above, and with  $Q_i^{-1} > 0$  ( $i \in N$ ,  $k \in K$ ),  $R_{ik} > 0$  ( $i, j \in N, j \neq i, k \in K$ ) and CLPS information pattern, admits a unique Nash equilibrium solution if, and only if, (6.17a) and (6.17d) admit unique solution sets  $\{P_k^*, i \in N, k \in K\}$  and  $\{u_k^*, i \in N, k \in K\}$  in which case the

equilibrium strategies are given by  $l_i^*(x_k) = -P_k l_i^* X_k - a_k^*$  ( $f \in K, i \in N$ ); (6.41) that is, the unique solution coincides with the unique strongly time-consistent feedback Nash equilibrium solution of the deterministic version of the game (cf. Corollary 6.1). We therefore observe that an appropriate stochastic formulation in affine-quadratic dynamic games with CLPS information pattern serves to remove informational nonuniqueness of equilibria, since it leads to a unique equilibrium

NASH AND SADDLE-POINT EQUILIBRIA 307 which is one of the uncountably many equilibria of the original deterministic game. This particular equilibrium solution (which is also a strongly time-consistent feedback equilibrium solution in this case) may be viewed as the unique robust solution of the deterministic game, which is insensitive to zero-mean random perturbations in the state equation. As an illustration, let us consider the three-person dynamic game example of subsection 6.3.1, and, out of the uncountably many Nash equilibrium solutions obtained there, let us attempt to select the one which features such a property. The new (stochastic) game is now the one which has zero-mean random perturbations in (6.31). More specifically, the state equation is  $\dot{x}_3 = x_2 + u_1 + u_2 + \omega_2$ ,  $\dot{I}^2 = X_i + U_3 + \theta_i$ ,  $J$  where  $E[\omega_2] = E\{\theta_i\} = 0$ ; the variables  $\omega_2, \theta_i, x_i$  are Gaussian, statistically independent, and  $E[(\omega_2)^2] > 0$ ,  $E[(\theta_i)^2] > 0$ . The unique Nash equilibrium solution of this stochastic game is  $\{y_1^*(x_2) = -(2/3)x_2, y_2^*(x_2) = (1/3)x_2, y_3^*(\hat{i}) = -(2/11)a_i\}$  which is also the feedback Nash equilibrium solution of the associated deterministic game. If the underlying information pattern of a stochastic dynamic game is not CLPS (or more precisely, if every player does not have access to the current value of the state), the statement of Thm. 6.10 is not valid. It is, in fact, not possible to obtain a general theorem which provides the Nash equilibrium solution under all different types of deterministic information patterns, and therefore one has to solve each such game separately. One common salient feature of all these solutions is that they are all free of any "informational nonuniqueness", and any nonuniqueness in equilibria (if it exists) is only due to the structure of the cost functionals and state equation. In particular, for affine-quadratic stochastic dynamic games, it can be shown that the Nash equilibrium solution will in general be unique under every deterministic information pattern. One such class of problems, viz. AT-person linear-quadratic stochastic dynamic games wherein some players have CLPS information and others OL information, have been considered in Başar (1975)

and a unique Nash equilibrium solution has been obtained in explicit form. This solution has the property that for the players with CLPS information the equilibrium strategies are dependent linearly not only on the current value of the state, but also on the initial state. The following example now illustrates the nature of this solution and also serves to provide a comparative study of Nash equilibria under four different information patterns. Example 6.1 Consider the stochastic version of the three-person nonzero-sum game of subsection 6.3.1. Under four different information structures, namely, when (i)  $P_1$  and  $P_2$  have access to CLPS information and  $P_3$  has OL information (this case is also covered by Corollary 6.4), (ii)  $P_1$  has CLPS information and  $P_2, P_3$  have OL information, (iii)  $P_2$  has CLPS information and  $P_1, P_3$  have OL information, (iv) all three players have OL information. In these

308 T. BAŞAR AND G. J. OLSDER stochastic games the state dynamics are described by  $\dot{x}_1 = x_2 + u_1 + u_2 + \xi_1$ ,  $\dot{x}_2 = x_1 + u_3 + \xi_2$ , where  $\{\xi_k, x_k\}$  is a set of statistically independent Gaussian random variables with  $E[\xi_k] = 0$ ,  $E[\xi_k^2] > 0$ ,  $k = 1, 2$ . The cost functionals are again as given by (6.32), but now their expected values determine the costs incurred to the players. (i) Under the first information pattern (CLPS for  $P_1$  and  $P_2$ , OL for  $P_3$ ), the unique solution also follows from Corollary 6.4 and is given by (as discussed right after Corollary 6.4)  $J_1^*(x_1) = -3x_1^2$ ,  $J_2^*(x_2) = -x_2^2$ ,  $J_3^*(x_1, x_2) = -x_1^2 - x_2^2$ . The corresponding expected values of the cost functionals are  $J_1^* = -3E[x_1^2] + E[\xi_1^2]$ ,  $J_2^* = -E[x_2^2] + E[\xi_2^2]$ ,  $J_3^* = -E[x_1^2] - E[x_2^2] + E[\xi_1^2] + E[\xi_2^2]$ . (ii) Under the second information pattern (CLPS for  $P_1$ , OL for  $P_2$  and  $P_3$ ), a direct derivation or utilization of the results of Başar (1975) leads to the unique Nash equilibrium solution  $J_1^*(x_1) = -x_1^2$ ,  $J_2^*(x_1, x_2) = -x_1^2 - x_2^2$ ,  $J_3^*(x_1, x_2) = -x_1^2 - x_2^2$ . Note that this unique solution of the stochastic game under consideration corresponds to one of the uncountably many equilibrium solutions presented in Prop. 6.5 for its deterministic version—simply set  $p = 1/6$ ,  $q = 1/3$ . The corresponding expected values of the cost functionals are  $J_1^* = -E[x_1^2] + E[\xi_1^2]$ ,  $J_2^* = -E[x_2^2] + E[\xi_2^2]$ ,  $J_3^* = -E[x_1^2] - E[x_2^2] + E[\xi_1^2] + E[\xi_2^2]$ . (iii) Under the third information pattern (CLPS for  $P_2$ , OL for  $P_1$  and  $P_3$ ), direct derivation leads to a finite solution set  $J_1^*(x_1) = -x_1^2$ ,  $J_2^*(x_1, x_2) = -x_1^2 - x_2^2$ ,  $J_3^*(x_1, x_2) = -x_1^2 - x_2^2$ . This Gaussian distribution may be replaced by any probability distribution that assigns positive probability mass to every open subset of  $R^n$  (see also Remark 5.11), without affecting the results to be obtained in the sequel.

NASH AND SADDLE-POINT EQUILIBRIA 309 which also corresponds to one of the solution sets of Prop. 6.5, this time obtained by setting  $p = 2/3$ ,  $q = 2/3$ . The corresponding expected values of the cost functionals are  $J_1^* = -E\{(\xi_1)^2\} + 4B[(\xi_1)^2] + E\{(\xi_2)^2\}$ , (6.44a)  $J_2^* = -j_3 = -E[(\xi_1)^2] - 2E\{(\xi_1)^2\} - E[(\xi_2)^2]$ . (6.44b) (iv) Under the fourth information pattern (OL for all three players), the unique Nash equilibrium solution is  $u^i = -\xi_i$ ,  $v^i = \xi_i$ ,  $w^i = \xi_i$  which can also be obtained from Prop. 6.5 by letting  $p = 2/3$ ,  $q = -1/3$ . The corresponding expected cost values are  $J_U = E\{(\xi_1)^2\} + E\{(\xi_2)^2\} + E\{(\xi_3)^2\}$  (6.45a)  $J_2 = -j_3 = -E[(\xi_1)^2] - E[(\xi_2)^2]$ . (6.45b) A comparison of the expected equilibrium cost values of the players attained under the four different information patterns now reveals some interesting features of the Nash equilibrium solution. (1) Comparing (6.43a) with (6.45a), we observe that when the other players' information patterns are OL, an increase in information to P1 (that is, going from OL to CLPS) could improve his performance (if  $E[(\xi_1)^2]$  is sufficiently large) or be detrimental (if  $E[(\xi_2)^2]$  is relatively small). (2) When P2 has CLPS information, a comparison of (6.42a) and (6.44a) leads to a similar trend in the effect of an increase of P1's information to his equilibrium performance. (3) When P1 has CLPS information, an increase in information to P2 leads to degradation in  $J_2^*$  (simply compare (6.42b) with (6.43b)). (4) When P1 has OL information, an increase in information to P2 leads to improvement in his performance (compare (6.44b) with (6.45b)). These results clearly indicate that, in nonzero-sum games and under the Nash equilibrium solution concept, an increase in information to one player does not necessarily lead to improvement in his equilibrium performance, and at times it could even be detrimental. D

310 T. BAŞAR AND G. J. OLSDER 6.5 Open-Loop and Feedback Nash and Saddle-Point Equilibria of Differential Games This section presents results which are counterparts of those given in Section 6.2, for differential games that fit the framework of Def. 5.5; that is, we consider  $N$ -person quantitative dynamic games defined in continuous time, and described by the state equation  $\dot{x}(t) = f(t, x(t), u_1(t), \dots, u_N(t))$ ;  $x(0) = x_0$  (6.46a) and cost functional  $U(u_1, \dots, u_N) = \int_0^T g_i(t, x(t), u_1(t), \dots, u_N(t)) dt + q_i(x(T))$   $i \in N$ . (6.46b) Here,  $[0, T]$  denotes the fixed prescribed duration of the game,  $x_0$  is the initial state known by all players,  $x(t) \in \mathbb{R}^n$  and  $u^*(i) \in S^* \subset \mathbb{R}^{m_i}$  ( $i \in N$ )  $\forall i \in [0, T]$ . Furthermore,  $f$  and  $g^* \in \mathbb{R}^*$  ( $i \in N$ ) satisfy the conditions of Thm. 5.1 so that, for a given information structure (as one of (i), (ii) or (iv) of Def. 5.6), the state equation



admits a unique solution for every corresponding AT-tuple  $\{7^* \in T1; i \in N\}$ . Nash equilibrium solution under a given information structure is again defined as one satisfying the set of inequalities (6.3) where  $J1$  denotes the cost functional of  $Pi$  for the differential game in strategic (normal) form. 6.5.1 Open-loop Nash equilibria The results to be described in this subsection are counterparts of those given in Section 6.2.1; and in order to display this relationship explicitly, we shall present them here in the same order as their counterparts in subsection 6.2.1. We therefore first have the counterpart of Thm. 6.1 in the continuous time. Theorem 6.11 For an N-person differential game of prescribed fixed duration  $[0, T]$ , let (i)  $f(x, u, t)$  be continuously differentiable on  $R^n$ ,  $V \in [0, T]$ , (ii)  $U^i(x, u, t)$  and  $q^i(x, t)$  be continuously differentiable on  $R^n$ ,  $V \in [0, T]$ ,  $i \in N$ . Then, if  $\{u^i(x, t) = u^i(x, t); i \in N\}$  provides an open-loop Nash equilibrium solution, and  $\{x^*(t), 0 < t < T\}$  is the corresponding state trajectory, there exist N costate functions  $p^i(x, t) : [0, T] \rightarrow R^n, i \in N$ , such that the following relations are satisfied:  $x^*(t) = f(t, x^*(t), u^1(t), \dots, u^N(t))$ ,  $x^*(0) = x_0$ , (6.47a)  $Y^*(t, x_0) = u^i(x, t) = \arg \min_{u^i} [ \int_0^t p^i(x, u, t) \cdot u^i + \dots ]$  (6.47b)  $\dots, U^{i-1}(x, t), U^{i+1}(x, t), \dots, U^N(x, t)$  (6.47c)  $f(T) = q^i(x^*(T))$ ,  $i \in N$ , where  $H^i(t, p^i(x, u, t), u^1, \dots, u^N) = g^i(t, x, u^1, \dots, u^N) + p^i(x, u, t) \cdot f(t, x, u^1, \dots, u^N)$ ,  $t \in [0, T]$ ,  $i \in N$ . Every such Nash equilibrium solution is weakly time consistent. Proof. The proof follows the same lines as in the proof of Thm. 6.1, but now the minimum principle for continuous-time control systems (i.e., Thm. 5.4) is used instead of Thm. 5.5. D Remark 6.13 One class of differential games for which the necessity condition of Thm. 6.11 is satisfied is that with weakly coupled players (a la Remark 4.1), that is one with the following state equation and cost functions (taking  $N=2$ , without any loss of generality):  $\dot{x}_i(t) = f_i(t, x_i(t), u_1(t)) + e f_{12}(t, x_2(t), y, x_i(t)) = x_{i0}$ , (6.48a)  $J^i(x, u) = \int_0^T [ q^i(t, x_i(t), u^i(t)) + e^{ii}(t, x_i(t), x_j(t)) ] dt + q^i(x_i(T)) + e q^i(x_j(T))$ ;  $j \neq i$ ,  $i, j = 1, 2$ , where  $e$  is a sufficiently small scalar. Under some appropriate convexity (on  $g^i$ 's) and differentiability conditions, it can be shown (see Srikanth and Başar (1992)) that there exists an  $\epsilon_0 > 0$  such that for all  $e \in (0, \epsilon_0)$ , the differential game admits a unique open-loop Nash equilibrium solution that is stable with respect to Gauss-Seidel or Jacobi iterations (see Def. 4.6 for terminology). This solution can be obtained by

NASH AND SADDLE-POINT EQUILIBRIA 311

expanding the state and control corresponding to the open-loop Nash equilibrium solution in power series in terms of  $\epsilon$ ,  $x^*(t; \epsilon) = X^*(t; \epsilon) + \epsilon X^1(t; \epsilon) + \epsilon^2 X^2(t; \epsilon) + \dots$ ,  $u^*(t; \epsilon) = u^0(t) + \epsilon u^1(t) + \epsilon^2 u^2(t) + \dots$ , substituting these into (6.47a)-(6.47c), along with a similar expansion for  $p^1(t)$ , and solving for the different terms  $X^k$  and  $u^k$ ,  $k = 0, 1, \dots$ , iteratively. It turns out that  $u^k$  ( $i = 1, 2$ ) are the (open-loop) optimal controls associated with the decoupled optimal control problems:  $J_i = \int_0^T f_i(t, X_i, u^0(t)) dt + \phi_i(X_i(T))$ ,  $X_i(0) = x_i^0$ ,  $f_i(\epsilon^*) = \int_0^T \dot{X}_i^T \epsilon^* \dot{X}_i dt + \phi_i(X_i(T))$ ,  $\epsilon^* = 1, 2$ ,  $J_0$

312 T. BAŞAR AND G. J. OLSDER and  $x^*$  is the corresponding state trajectory, with  $x^* = (x^1, x^2, \dots)^T$ . For  $k > 1$ ,  $u^k_1$  and  $u^k_2$  are obtained by solving some appropriate linear-quadratic optimal control problems (see Srikant and Başar(1991)). Hence this approach decomposes the original two-player differential game into two nonlinear optimal control problems (the zeroth-order problems) and a sequence of iteratively constructed linear-quadratic control problems. Halting this iteration at the  $k$ th step yields an  $\epsilon^k$ -approximate open-loop Nash equilibrium solution.  $\square$  As indicated earlier, Thm. 6.11 provides a set of necessary conditions for the open-loop Nash equilibrium solution to satisfy, and therefore it can be used to generate candidate solutions. For the special class of "affine-quadratic" differential games, however, a unique candidate solution can be obtained in explicit terms, which can further be shown to be an open-loop Nash equilibrium solution under certain convexity restrictions on the cost functional of the players.

**Definition 6.5** An  $N$ -person differential game of fixed prescribed duration is of the affine-quadratic type if  $S_i = R_i m^*$  ( $i \in N$ ) and  $f(t, x, u^1, \dots, u^N) = A(t)x + \sum_{i=1}^N B_i(t)u^i + c(t)$ ,  $Q = -x^T Q x$ , where  $A(-)$ ,  $B_i(-)$ ,  $Q_i(-)$ ,  $R_i(-)$  are matrices of appropriate dimensions,  $c(-)$  is an  $n$ -dimensional vector, all defined on  $[0, T]$ , and with continuous entries ( $i, j \in N$ ). Furthermore  $Q_i, Q_i^*$  are symmetric, and  $R_i^*(-) > 0$  ( $i \in N$ ). An affine-quadratic game is of the linear-quadratic type if  $c = 0$ .

**Theorem 6.12** For an  $N$ -person affine-quadratic differential game with  $Q_i(-) > 0$ ,  $Q_i^* > 0$  ( $i \in N$ ), let there exist a solution set  $\{M_i^0; i \in N\}$  to the coupled matrix Riccati differential equations  $M_i + M_i A + A^T M_i + Q_i - M_i \sum_{j=1}^N B_j U_j^{-1} U_j^T M_j = 0$ ;  $i \in N$  (6.49)  $A f^*(T) = Q$  ( $i \in N$ ). Then, the differential game admits an open-loop Nash equilibrium solution given by  $y^*(t, x_0) = \epsilon^*(t) = -R_i^{-1}(t) [B_i^T(t) M_i(t) x(t) + m_i^T(t)]$  ( $i \in N$ ), (6.50) where  $\{m_i^*(t), i \in N\}$  solve uniquely the set of linear differential equations:  $\dot{m}_i + A^T m_i + M_i (c - \sum_{j=1}^N B_j R_j^{-1} B_j^T m_j) = 0$ ;  $J \in N$  (6.51)  $m_i(T) = 0$  ( $i \in N$ ),

NASH AND SADDLE-POINT EQUILIBRIA 313 and  $x^*(\cdot)$  denotes the (Nash) equilibrium state trajectory, generated by  $x^*(t) = x_0 + \int_0^t f(s, a(s)) ds$ ,  $dJ(x, u) = F(t, x, u)$ ;  $F(t, x, u) = A - YB^{-1}B' M^{-1} M'(t)$ ,  $v(t) = c(t) - YB^{-1}B' M^{-1} M'(t)$ . Proof- For the affine-quadratic differential game, and in view of the additional restrictions  $Q_1(\cdot) > 0$ ,  $Q_2 > 0$ ,  $L(u_1, \dots, u_N)$  is a strictly convex function of  $u$  for all permissible control functions  $u^j(\cdot)$  ( $j = 1, \dots, N$ ) and for all  $x_0 \in \mathbb{R}^n$ . This then implies that Thm. 6.11 is also a sufficiency result and every solution set of the first-order conditions provides an open-loop Nash solution. Hence, we now show that the solution given in Thm. 6.12 indeed satisfies the first-order conditions. First note that the Hamiltonian is  $H(t, p, x, u) = p'(Ax + c + B'u) - \frac{1}{2} u' S u + \sum_{i=1}^N p_i' (A_i x + c_i + B_i u_i)$  whose minimization with respect to  $u(t) \in \mathbb{R}^m$  yields the unique relation  $u^*(i; t) = -(S^{-1}(B_i' p_i + c_i))$ ,  $i \in N$ . (i) Furthermore, the costate equations are  $\dot{p}_i = -Q_i x - A_i' p_i$ ,  $p_i(T) = Q_i x(T)$  ( $i \in N$ ), and the optimal state trajectory is given by  $x^* = Ax^* + c - Y^* B^{-1} \{R^{-1} B^{-1} p_i, x^*(0) = x_0$ . (ii) This set of differential equations constitutes a two-point boundary value problem. Now, substituting  $p_i = M_i x^* + m_i$  ( $i \in N$ ) into the costate equations, we arrive at the conclusion that if  $M_i$  ( $i \in N$ ) and  $m_i$  ( $i \in N$ ) satisfy (6.49) and (6.51), respectively, this indeed solves the two-point boundary value problem, along with  $x^*$ . The expressions for the open-loop Nash strategies follow readily from (i) by substituting  $p_i = M_i x^* + m_i$ , and likewise the associated state trajectory (6.52) follows from (ii).  $Q$

Remark 6.14 It is possible to obtain the result of Thm. 6.12 in a somewhat different way, by directly utilizing Prop. 5.2; see Problem 10 in section 6.8.  $\square$

314 T. BAŞAR AND G. J. OLSDER Theorem 6.12 provides a set of sufficiency conditions for the Nash equilibrium solution to exist. It will now be shown that a Nash equilibrium may exist even if (6.49) does not admit a solution. For simplicity of presentation assume now that  $c(t) = 0$ . The differential equations for  $p_i(t)$  and  $x^*(t)$  can be written jointly as  $S \dot{x} + A' x + \sum_{i=1}^N Q_i p_i = 0$ ,  $p_i(T) = Q_i x(T)$ ,  $i = 1, \dots, N$  where  $S^* = B^* R^{-1} B^*$ ,  $i = 1, 2, \dots, N$ . The boundary conditions of (6.53)  $1, 2, \dots, N$ . Symbolically (6.53) will be written as  $Q_1 z = -Gz$ , with  $G = S^{-1} A'$  and with  $Q_2 \{x(0) + Q_2^{-1} \sum_{i=1}^N p_i(0) + Q_2^{-1} \sum_{i=1}^N p_i(T)\} = 0$ , (6.54) (6.55)  $\int_0^T (x^0 + Q_2^{-1} \sum_{i=1}^N p_i^0 + Q_2^{-1} \sum_{i=1}^N p_i^T) e^{-G(T-t)} dt = 0$  (6.56) From (6.55) we immediately get that a unique open-loop Nash equilibrium solution exists if and only if the equation  $\{Pe^{GT} + Q\}e^{-GT}z(0) = \{x^0 + \dots + 0\}$  (6.57) is uniquely solvable in  $z(0)$  for every  $x_0$ . The matrix  $G$

consists of  $(N + 1) \times (N+1)$  blocks. The same block structure will be applied to  $eGt$  and these blocks will be indicated by  $W_{ij}(t)$ ,  $i, j = 1, 2, \dots, N + 1$ . If we define  $N H(T) \pm W_n(T) + Y_t W^{+i}(T) Q_k f$ , (6.58)  $k=1$  then the solvability of (6.57) is equivalent to the invertibility of  $H(T)$ . The following example now shows that an open-loop Nash equilibrium solution can exist even if the coupled set of Riccati differential equations (6.49) does not have a solution.

**NASH AND SADDLE-POINT EQUILIBRIA** 315 Example 6.2 Consider the two-person ( $N = 2$ ) linear-quadratic game with  $A = \begin{pmatrix} 0 & 1 \\ -5/22 & \end{pmatrix}$ ,  $B_1 = \begin{pmatrix} 0 & 1 \end{pmatrix}$ ,  $B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$ ,  $R_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ,  $R_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . Choose  $T = 0.1$ . Then numerical calculation leads to  $W_{11}(T) = \begin{pmatrix} 1.1155 & -0.0051 \\ -0.0051 & 1.0230 \end{pmatrix}$ ,  $W_{12}(T) = \begin{pmatrix} -0.1007 & -0.1047 \\ -0.1047 & -0.0051 \end{pmatrix}$ ,  $W_{21}(T) = \begin{pmatrix} 0.0964 & 0.2002 \\ 0.2002 & 0.1005 \end{pmatrix}$ ,  $W_{22}(T) = \begin{pmatrix} 0.0002 & 0.0002 \\ 0.0002 & 0.0002 \end{pmatrix}$ . Here the notation  $\{W_{ij}\}$  refers to the  $j$ 'th element of the  $2 \times 2$  matrix  $W_i(T)$ . Clearly both  $Q_1$  and  $Q_2^*$  are positive definite, whereas  $H(T) = W_n(T) + W_{12}(T)Q_1 + W_{13}(T)Q_2 = \begin{pmatrix} -752 & 6945 \end{pmatrix}$  is not invertible. The conclusion therefore is that an open-loop Nash equilibrium does not exist and the corresponding set of Riccati differential equations does not have a solution. Now we modify the problem statement in one respect, viz. instead of  $T = 0.1$  now take  $T = 0.11$ . Numerical calculation then shows that  $H(T)$  is now invertible. However, the solution of the set of Riccati differential equations, when integrated backward and starting at  $T = 0.11$ , blows up (i.e., becomes  $\infty$ ) at  $t = 0.01$  (note that the solution of the Riccati equations is shift-invariant). Thus it has been shown that for  $T = 0.11$  an open-loop Nash equilibrium solution exists despite the fact that the Riccati differential equations do not have a solution on the interval  $[0, 0.11]$ .  $\square$

316 T. BAŞAR AND G. J. OLSDER What the preceding example has shown is that for a linear-quadratic game defined on a fixed interval  $[0, T]$ , existence of a Nash equilibrium does not imply existence of a solution to the set of Riccati equations (6.49). An equivalence between these two conditions exists, however, if  $T$  is a variable, as shown in (Engwerda, 1998). What can be shown is that the following three statements are equivalent, where  $\tau$  is a positive real number. • For each  $T \in (0, \tau]$  an open-loop Nash equilibrium solution exists for the linear-quadratic game defined on  $[0, T]$ , •  $H(T)$  is invertible for all  $T \in [0, \tau]$ . • The set of Riccati differential equations (6.49)

has a solution on  $[0, T]$ . Open-loop saddle points of zero-sum differential games For the class of two-person zero-sum differential games, we rewrite the cost functional (to be minimized by  $P_1$  and maximized by  $P_2$ ) as  $L(u, u_2) = \int_0^T g(t, x, u_1, u_2) dt + q(x(T))$ , and obtain the following set of necessary conditions for the saddle-point strategies to satisfy. Theorem 6.13 For a two-person zero-sum differential game of prescribed fixed duration  $[0, T]$ , let (i)  $f(t, x, u_1, u_2)$  be continuously differentiable on  $\mathbb{R}^n$ ,  $\forall t \in [0, T]$ , (ii)  $g(t, x, u_1, u_2)$  and  $q(\cdot)$  be continuously differentiable on  $\mathbb{R}^n$ ,  $\forall t \in [0, T]$ . Then,  $\{u_1^*(t); u_2^*(t); x^*(t); p(t)\}$  provides an open-loop saddle-point solution, and  $\{x^*(t), 0 < t < T\}$  is the corresponding state trajectory, there exists a costate function  $p(\cdot) : [0, T] \rightarrow \mathbb{R}^n$  such that the following relations are satisfied:  $x^*(t) = f(t, x^*(t), u_1^*(t), u_2^*(t))$ ;  $x^*(0) = x_0$ , (6.59a)  $H(t, p, x^*, u_1, u_2) < H(t, p, x^*, u_1, u_2^*) < H(t, p, x^*, u_1, u_2)$ , (6.59b)  $p'(t) = -\frac{\partial H(t, p(t), x^*(t), u_1(t), u_2^*(t))}{\partial p}$ ;  $p'(T) = \frac{\partial q(x^*(T))}{\partial x}$ , (6.59c) where  $H(t, p, x, u_1, u_2) = g(t, x, u_1, u_2) + p' f(t, x, u_1, u_2)$ ,  $t \in [0, T]$ . (6.59d)

NASH AND SADDLE-POINT EQUILIBRIA 317 Proof. This result follows directly from Thm. 6.11 by taking  $N = 2$ ,  $g_1 = -g_2 = g$ ,  $q_1 = -q_2 = q$ , and by noting that  $p_1(\cdot) = -p_2(\cdot) = p(\cdot)$  and  $\#(\cdot) = -H_2(\cdot) \frac{\partial H(\cdot)}{\partial p}$ . 0 For the special case of two-person zero-sum affine-quadratic differential games described by the state equation  $\dot{x} = A(t)x + B_1(t)u_1 + B_2(t)u_2 + c(t)$ ;  $x(0) = x_0$  (6.60a) and the objective functional  $L(u, u_2) = \int_0^T (x'(t)Q(t)x(t) + u_1'(t)R_1(t)u_1(t) - u_2'(t)R_2(t)u_2(t)) dt + x'(T)Q_f x(T)$ , we now first impose additional restrictions on the parameters so that  $L(u, u_2)$  is strictly concave in  $u_2$ . Lemma 6.4 For the affine-quadratic two-person zero-sum differential game introduced above, the objective functional  $L(u, u_2)$  is strictly concave in  $u_2$  for all permissible  $u_1(\cdot)$  if, and only if, there exists a unique bounded symmetric solution  $S(\cdot)$  to the matrix Riccati equation  $\dot{S} + A'S + SA + Q + SB_2B_2'S = 0$ ;  $S(T) = Q_f$ , (6.61) on the interval  $[0, T]$ . Proof. A reasoning similar to the one employed in the proof of Lemma 6.1 leads to the observation that the requirement of strict concavity is equivalent to the condition of existence of a unique solution to the optimal control problem  $\min_u \int_0^T \{u'(t)u(t) - y'(t)Q(t)y(t)\} dt - y'(T)Q_f y(T)$  subject to  $\dot{y} = A(t)y + B_2(t)u$ ;  $y(0) = x_0$ . The lemma then follows, in view of this observation, from Remark 5.8. Q Theorem 6.14 For the two-person affine-quadratic zero-sum differential game described by (6.60a)-(6.60b) and with  $Q_f > 0$ ,  $Q(\cdot) > 0$ , let there exist a unique bounded symmetric solution to the matrix Riccati differential equation (6.61) on  $[0, T]$ . Then, (i) there exists a unique

bounded symmetric solution  $M(\cdot)$  to the (generalized) matrix Riccati differential equation  $M + A'M + MA + Q - M\{B_1B_1' - B_2B_2'\}M = 0$ ;  $M\{T\} = Q_f$ , (6.62a)

318 T. BAŞAR AND G. J. OLSDER on  $[0, T]$ , and again on  $[0, T]$  a unique bounded solution  $m(\cdot)$  to the linear vector differential equation  $m + [A - \{B_1B_1' - B_2B_2'\}M]m + Mc = 0$ ;  $m\{T\} = 0$ , (6.62b) (ii) the game admits a unique open-loop saddle-point solution given by  $j_1^*(t, x_0) = u_1^*(t) = -B_1'(t)[M(t)x^*(t) + m(t)]$ ,  $j_2^*(t, x_0) = u_2^*(t) = B_2'(t)[M(t)x^*(t) + m(t)]$ ,  $\{b - M\}$  where  $x^*(\cdot)$  denotes the associated state trajectory, generated by  $x^*(t) = x^*(t, 0)x_0 + \int_0^t [c(\tau) - (B_1B_1' - B_2B_2')m(\tau)]d\tau + F(t)$ ,  $F(t) = I$  (6.64)  $F = A - \{B_1B_1' - B_2B_2'\}M$ . Proof. The steps involved in the proof of this theorem are similar to those used in the proof of Thm. 6.4: we first establish existence of a unique saddle-point solution to the affine-quadratic differential game under consideration, and we then verify (i) and (ii) above in view of this property. For the former step we shall proceed rather casually, since a rigorous verification requires some prior knowledge of functional analysis which is beyond the scope of this book. The interested reader can, however, refer to Balakrishnan (1976) for a rigorous exposition of the underlying mathematics, or to Bensoussan (1971) and Lukes and Russell (1971) for a framework (Hilbert space setting) wherein similar mathematical techniques are used at a rigorous level but in a somewhat different context. 61 Let  $U_i$ , the space of permissible control functions  $u_i(\cdot)$ , be a complete (Banach) space equipped with an inner product  $\int_0^T (u, v) dt = \int_0^T u(t)'v(t)dt$ , for  $u, v \in U_i$ ,  $i = 1, 2$ . Then, since  $L(u_1, u_2)$  is quadratic in  $x$ ,  $u_1, u_2$ , and  $x$  is linear in  $u_1$  and  $u_2$ , the cost  $L(u_1, u_2)$  can be written as  $L(u_1, u_2) = (u_1, L_1u_1) + (u_2, L_2u_2) + (u_1, L_{12}u_2) + (u_2, L_{21}u_1) + r_3$ , where  $L_1 : U_1 \rightarrow U_1$ ,  $L_2 : U_2 \rightarrow U_2$ ,  $L_{12} : U_2 \rightarrow U_1$  are bounded linear operators,  $r_1 \in U_1$ ,  $r_2 \in U_2$  and  $r_3 \in \mathbb{R}$ . Moreover, since  $Q(\cdot) > 0$ ,  $Q_f > 0$ ,  $L_x$  is a strongly positive operator (written as  $L_x > 0$ ), and furthermore under 61 For an alternative rigorous verification of the result of this theorem, see Başar and Bernhard (1995).

NASH AND SADDLE-POINT EQUILIBRIA 319 the condition of unique solvability of (6.61) (cf. Lemma 6.4)  $L_2 > 0$ . Now, if  $L(u_1, u_2)$  admits a unique saddle point in open-loop policies, it is necessary and sufficient that it satisfy the following two relations (obtained by differentiation of (i))  $L_1u_1 + L_{12}u_2 + r_1 = 0$ ,  $-L_{21}u_1 + L_2u_2 + r_2 = 0$ , where  $L_2$  denotes the adjoint of  $L_{12}$ . Solving for

$u_2 \in U_2$  from the second one (uniquely), and substituting it into the first one, we obtain the single relation  $(U + L^*L^*u_1 + n - L^*L^*S = 0$  (ii) which admits a unique solution, since  $L^*L^*L^*L^* > 0$  and is thereby invertible. (This latter feature follows since  $L^* > 0$  and  $L^*L^*L^*L^*$  is a nonnegative self-adjoint compact operator; see Balakrishnan (1976), for details.) We have thus established existence of a unique open-loop saddle-point strategy for  $P_1$  under the condition of Lemma 6.4. This readily implies existence of a unique open-loop saddle-point strategy for  $P_2$  since  $L^*$  is invertible. Thus completing the proof of existence of a unique saddle point, we now turn to verification of (6.62a)-(6.62b) and (6.63). At the outset, let us first note that the unique saddle-point solution should necessarily satisfy relations (6.59a)-(6.59d) which can be rewritten for the affine-quadratic game as  $i^* = A(t)x^* + B_1^*u_1^* + B_2(t)u_2^*(t) + c(t); x^*(0) = x_0, u_1^*(t) = -B_1(t)'p(t); u_2^*(t) = B_2(t)'p(t),$  (iii)  $p(t) = -Q(t)x^*(t) - A(t)'p(t); p(T) = Qf^*(T).$  Since  $r^1(-)$  and  $r^2(-)$  introduced earlier are affine functions of  $x_0$ , the unique solution  $u_1^*$  of (ii) is necessarily an affine function of  $x_0$ , and so is  $u_2^*$ . Therefore, the dependence of  $p(t)$  on  $x_0$  in (iii) should be affine. This readily implies that we can write, without any loss of generality,  $p(t) = M(t)x^*(t) + m(t)$ , where  $M(-)$  is some bounded matrix-valued function of dimension  $(n \times n)$  and  $m(-)$  is some bounded vector-valued function of dimension  $n$ , both with continuously differentiable elements. Now, substitution of this relation into (iii), and requiring (iii) to hold true for all  $x_0 \in R^n$ , leads to (6.62a)-(6.62b), which in turn yields (6.63).

**D Remark 6.15** The statement of Thm. 6.14 is valid even if the "nonnegative-definiteness" condition on  $Q(-)$  does not hold, provided that some other appropriate conditions are imposed on the game parameters to ensure that  $L$  defined by (6.60b) is strictly convex in  $u_1$ . One such condition, which is in fact tight, can be obtained directly from Lemma 6.4, by simply replacing  $Qf$  by  $-Q$ ,  $Q(-)$  by  $-Q(-)$  and  $B_2$  by  $B_1$ , which is the existence of a unique bounded symmetric solution  $S_1(-)$  to the matrix Riccati differential equation  $S_1' + A'S_1 + S_1A + Q - SB_1B_1'S = 0; S(T) = Qf$ , (6.65) on the interval  $[0, T]$ .

320 T. BAŞAR AND G. J. OLSDER 6.5.2 Closed-loop no-memory and feedback Nash equilibria Under the memoryless perfect state information pattern, the following theorem provides a set of necessary conditions for any closed-loop no-memory Nash equilibrium solution to satisfy. **Theorem 6.15** For an  $N$ -person differential game of prescribed fixed duration  $[0, T]$ , let (i)  $f_i^*(t, \bullet)$  be continuously differentiable on  $R^n \times S_1 \times \dots \times S_N, \forall t \in [0, T]$ , (ii)  $g'(t, \bullet)$

and  $q_i(-)$  be continuously differentiable on  $R^n \times S^1 \times \dots \times S^N$  and  $R^n$ , respectively,  $\forall i \in [0, T], i \in N$ . Then, if  $\{j_i^*(t, x, x_0); i \in N\}$  provides a closed-loop no-memory Nash equilibrium solution such that  $\gamma_i^*(\cdot, \cdot, \cdot; e_0)$  is continuously differentiable on  $R^n$  ( $\forall t \in [0, T], i \in N$ ), and  $\{x^*(t), 0 < t < T\}$  is the corresponding state trajectory, there exist  $N$  costate functions  $p^i(-) : [0, T] \rightarrow R^n, i \in N$ , such that the following relations are satisfied:  $x^*(t) = f(t, x^*(t), u^1(t), \dots, u^N(t)); x^*(0) = x_0$ , (6.66a)  $Y^* \{t, x \mid x_0\} = \langle \cdot \rangle^*(*) = \arg \min_t \{t, p^i(t), x^*(t), u^1(t), \dots, u^N(t)\}$  (6.66b)  $\langle \cdot \rangle^{*-1}(t, \langle \cdot \rangle^*, \langle \cdot \rangle^{i+1}(\cdot), \dots, \cdot, \cdot / (*t)), P^{*i}(t) = - \hat{H}_i(t, p_i(t), x^*, \gamma_1^*(\cdot, a^*, a^*), \dots, \gamma_i^{i-1}(\cdot, a^*, a^*), \langle \cdot \rangle^{i*}, \gamma_{i+1}^*(\cdot, x^*, x_0), \dots, \gamma_N^*(\cdot, \cdot, a^*), \cdot)$ , (6.66c)  $\dot{p}^i(t) = -k_{q_i} \{x^* \mid T\}' \quad i \in N$ ; where if  $\cdot^*$  is as defined in Thm. 6.11. Proof. Let us consider the  $i$ th inequality of (6.3), which fixes all players' strategies (except the  $i$ th one) at  $\gamma_j = \gamma_j^* \quad (j \neq i, j \in N)$  and constitutes an optimal control problem for  $P_i$ . Therefore, relations (6.66a)-(6.66c) directly follow from Thm. 5.4 (the minimum principle for continuous-time systems), as do relations (6.47a)-(6.47c) of Thm. 6.11. We should note, however, that the partial derivative (with respect to  $x$ ) in the costate equations (6.66c) receives contribution also from dependence of the remaining  $N - 1$  players' strategies on the current value of  $x$ —a feature which was clearly absent in the costate equations of Thm. 6.11. D The set of relations (6.66a)-(6.66c) in general admits an uncountable number of solutions, which correspond to "informationally nonunique" Nash equilibrium solutions of the differential game under memoryless perfect state information pattern. All of these are weakly time consistent, one of which is the open-loop

NASH AND SADDLE-POINT EQUILIBRIA 321 solution given in Thm. 6.11. The analysis of subsection 6.3.1 can in fact be duplicated in the continuous time to produce explicitly a plethora of informationally nonunique Nash equilibria under dynamic information; and the reader is referred to Başar (1977b) for one such explicit derivation. In order to eliminate informational nonuniqueness in the derivation of Nash equilibria under dynamic information (specifically under the MPS and CLPS patterns) we refine the Nash solution concept further, by requiring it to satisfy conditions similar to those of Def. 6.2, but now in the continuous time. Such a consideration leads to the concept of a "feedback Nash equilibrium solution" which is introduced below. Definition 6.6 For an  $N$ -person differential game of prescribed fixed duration  $[0, T]$  and with memoryless perfect state (MPS) or closed-loop perfect state (CLPS) information pattern, an  $N$ -tuple of strategies  $\{f_i^* \in r^*; i \in N\}$  constitutes a



feedback Nash equilibrium solution if there exist functionals  $V^*(v)$  defined on  $[0, T] \times \mathbb{R}^n$  and satisfying the following relations for each  $i \in N$ :  $V^*(t, x) = \int_t^T g_i(s, x^*(s), u_1^*(s, V^*(s)), \dots, u_N^*(s, V^*(s))) ds + q(x^*(T))$  (6.67) where, on the interval  $[t, T]$ ,  $\dot{x}(s) = f(s, x(s), u_1(s, V^*(s)), \dots, u_N(s, V^*(s)))$ ;  $x_i(t) = x_i$ , (6.68)  $V^*(s) = f(s, x^*(s), u_1^*(s, V^*(s)), \dots, u_N^*(s, V^*(s)))$ ;  $X^*(t) = X$ , and  $t, s$ , stands for either the data set  $\{x(s), x_0\}$  or  $\{x(a), a < s\}$ , depending on whether the information pattern is MPS or CLPS. One salient feature of the concept of feedback Nash equilibrium solution introduced above is that if an AT-tuple  $\{u_i^*/i \in N\}$  provides a feedback Nash equilibrium solution (FNES) to an AT-person differential game with duration  $[0, T]$ , its restriction to the time interval  $[t, T]$  provides an FNES to the same differential game defined on the shorter time interval  $[t, T]$ , with the initial state taken as  $x(t)$ , and this being so for all  $0 < t < T$ ; hence, an FNES is strongly time consistent. An immediate consequence of this observation is that, under either MPS or CLPS information pattern, feedback Nash equilibrium strategies Here TI is chosen to be compatible with the associated information pattern.

322 T. BAŞAR AND G. J. OLSDER will depend only on the time variable and the current value of the state, but not on memory (including the initial state  $x_0$ ). The following proposition now verifies that an FNES is indeed a Nash equilibrium solution. Proposition 6.7 Under the memoryless (respectively, closed-loop) perfect state information pattern, every feedback Nash equilibrium solution of an N-person differential game is a closed-loop no-memory (respectively, closed-loop) Nash equilibrium solution (but not vice versa). Proof. Note that  $V_i(t, x)$  is the value function (cf. subsection 5.5.2) associated with the optimal control problem of minimizing over  $Y \in P$  the function  $J^i(u, \dots, u_{i-1}^*, u_{i+1}^*, \dots, u_N^*)$  (see in particular, equation (5.24a)). Therefore (6.67), together with the given terminal boundary condition, implies satisfaction of the  $i$ th inequality of (6.3). Since  $i \in N$  was arbitrary, and the foregoing argument is valid under both the MPS and CLPS information patterns, the result stands proven. Q If the value functions  $V_i$  ( $i \in N$ ) are continuously differentiable in both arguments, then N partial differential equations, related to the Hamilton-Jacobi-Bellman equation of optimal control (cf. subsection 5.5.2), replace (6.67). Theorem 6.16 For an N-person differential game of prescribed fixed duration  $[0, T]$ , and under either MPS or

CLPS information pattern, an  $N$ -tuple of strategies  $\{j^i \in V^i; i \in N\}$  provides a feedback Nash equilibrium solution if there exist functions  $V^i : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}, i \in N$ , satisfying the partial differential equations  $dV^i(t, x) + g^i(t, x, u^i) - \sum_{j \neq i} \frac{\partial V^i}{\partial x_j} (t, x) f_j(t, x, u^j) = g^i(x), i \in N, \forall t \in [0, T], x \in \mathbb{R}^n$  (6.69) where  $r(t, x, u^i) = g^i(t, x, u^i) - \sum_{j \neq i} \frac{\partial V^i}{\partial x_j} (t, x) f_j(t, x, u^j)$ , where  $r(t, x, u^i) = g^i(t, x, u^i) - \sum_{j \neq i} \frac{\partial V^i}{\partial x_j} (t, x) f_j(t, x, u^j)$ ,  $V^i(t, x) = g^i(t, x, u^i) - \sum_{j \neq i} \frac{\partial V^i}{\partial x_j} (t, x) f_j(t, x, u^j)$ . Every such equilibrium solution is strongly time consistent, and the corresponding Nash equilibrium cost for  $P_i$  is  $V^i(0, x_0)$ . Proof. In view of the argument employed in the proof of Prop. 6.7, this result follows by basically tracing the standard technique of obtaining (5.25) from (5.24a).  $\square$

**NASH AND SADDLE-POINT EQUILIBRIA 323** For the class of  $N$ -person affine-quadratic differential games (cf. Def. 6.5), it is possible to obtain an explicit solution for (6.69), which is quadratic in  $x$ . This also readily leads to a set of feedback Nash equilibrium strategies which are expressible in closed-form. **Corollary 6.5** For an  $N$ -person affine-quadratic differential game with  $Q^i \succ 0, Q^i \succ 0, R^i \succ 0 (i, j \in N, j \neq i)$ , let there exist a set of matrix valued functions  $Z^i(t) \succ 0, i \in N$ , satisfying the following  $N$  coupled matrix Riccati differential equations:  $Z^i + Z^i F + F^T Z^i + Y^i Z^j B_j R_j^{-1} R_j B_j^T Z^i + Q^i = 0; Z^i(T) = Q^i$ , where  $F(t) = A(t) - \sum_{j \neq i} B_j^T R_j^{-1} R_j B_j^T Z^j(t)$ . (6.70a)  $Z^i(T) = Q^i$ , where  $F(t) = A(t) - \sum_{j \neq i} B_j^T R_j^{-1} R_j B_j^T Z^j(t)$ . (6.70b)  $i \in N$  Then, under either the MPS or CLPS information pattern, the differential game admits a feedback Nash equilibrium solution, affine in the current value of the state, given by  $Y^i(t, x) = -B^i(t)^{-1} [Z^i(t)x(t) + C(t)] (i \in N)$ , (6.71) where  $Z^i(t) (i \in N)$  are obtained as the unique solution of the coupled linear differential equations  $C + F^T C + \sum_{j \neq i} Z^j B_j R_j^{-1} R_j B_j^T C + Z^i(i = 0; C(T) = 0$ , (6.72a)  $i \in N$  with  $0 \preceq C - \sum_{j \neq i} B_j^T R_j^{-1} R_j B_j^T C$  (6.72b) The corresponding values of the cost functionals are  $J^i = y^i J(0, x_0) = -x_0^T Z^i x_0 + x_0^T C(0) + n^i(0) (t \in N)$ , (6.73a) where  $n^i(t) (i \in N)$  are obtained from  $n^i + (\sum_{j \neq i} C_j B_j R_j^{-1} R_j B_j^T C_i - C_i) = 0; n^i(T) = 0$ . (6.73b) Proof Simply note that, under the condition of solvability of the set of matrix Riccati differential equations (6.70a), the partial differential equation

324 T. BAŞAR AND G. J. OLSDER (6.69) admits a solution in the form  $V^i(t, x) = x^T Z^i(t)x + x^T C(t) + n^i(t) (i \in N)$  with the corresponding minimizing controls being given by (6.71). The "nonnegative definiteness" requirement imposed on  $Z^i(t)$  is a consequence of the fact that  $V^i(t, x) > 0 \forall x \in \mathbb{R}^n, t \in [0, T]$ , this latter feature being due to the eigenvalue restrictions imposed a priori on  $Q^i(t), Q^i$

and  $R^{\{-}$ ). Finally, the corresponding "Nash" values for the cost functionals follow from the definition of  $V\{t,x)$  (see equation (6.67)).

**Remark 6.16** The foregoing corollary provides only one set of feedback Nash equilibrium strategies for the affine-quadratic game under consideration, and it does not attribute any uniqueness feature to this solution set. However, in view of the discrete-time counterpart of this result (cf. Corollary 6.1), one would expect the solution to be unique under the condition that (6.70a) admits a unique solution set; but, in order to verify this, one has to show that it is not possible to come up with other (possibly nonlinear) solutions that satisfy (6.67), and hitherto this has remained an unresolved problem. What has been accomplished, though, is verification of uniqueness of feedback Nash equilibrium when the players are restricted at the outset to affine memoryless strategies (cf. Lukes, 1971).

**Remark 6.17** As in the case of Remark 6.3, the result above extends readily to more general affine-quadratic dynamic games where the cost functions of the players contain additional terms that are linear in  $x$ , that is, with  $g^0$  and  $q^0$  in Def. 6.5 replaced, respectively, by  $g^* = \sum_{i=1}^n [A_i' x + 2f_i^*(\cdot)] + \sum_{i=1}^n l_i^*$  where  $f_i^*(\cdot)$  is a known  $n$ -dimensional vector-valued function, continuous on  $[0, T]$ , and  $l_i^*$  is a fixed  $n$ -dimensional vector, for each  $i \in N$ . Then, the statement of Corollary 6.5 as well as its derivation remains intact, with only the differential equation (6.72a) that generates  $\lambda^*(\cdot)$  now reading  $C + F'C + \sum_{i=1}^n Z_i B_i R_i^{-1} B_i' \lambda^* + \sum_{i=1}^n Z_i (g_i^* + \lambda^* = 0; C(T) = I)$ .

**Remark 6.18** For general nonlinear-nonquadratic differential games wherein the players are weakly coupled through the system equation as well as the cost functions, the features observed in Remark 6.13 for the open-loop Nash solution can also be derived for the feedback Nash solution, but now we have to use the sufficiency result of Thm. 6.16. Again confining ourselves to the two-player case, we take the system equation and the cost functions, as well as the expansion of

NASH AND SADDLE-POINT EQUILIBRIA 325 the state, to be in the same structural form as in Remark 6.13, and only replace the expansion for  $u^*(t; e)$  by a similar expansion for the feedback strategy:  $u^*(t,x;e) = \sum_{k=0}^{\infty} u_k^*(t,x;e)$ . Invoking a similar expansion on  $V^*(t,x;e) = \sum_{k=0}^{\infty} V_k^*(t,x;e)$  it can be shown (see Srikant and Başar, 1991) that  $V_k^*$  ( $i = 1,2$ ), the zeroth order terms, satisfy decoupled Hamilton-Jacobi-Bellman equations (associated with the optimal control problems obtained by setting  $e = 0$ ), and the higher-order terms,  $V_k^*$ ,  $k > 1$ , involve simple cost evaluations subject to some state

equation constraints. Furthermore, the higher-order feedback strategies,  $\gamma^k$ ,  $k > 1$ , are obtained from linear equations. More explicitly, for the zeroth order we have  $-V^0(t, x) = \min_{u_1} \max_{u_2} \{ f(t, x, u_1, u_2) + g^i(t, x, u_1) - V^1(t, x) - f_i(t, x, u_1, u_2) + g^*(t, x, u_1, u_2) \}$ ,  $\frac{d}{dt} V^0(t, x) = q^0(x)$ ,  $i = 1, 2$ . (6.74a) and to first-order  $\gamma^1$  satisfies  $\frac{d}{dt} V^1(t, x) = \min_{u_1} \max_{u_2} \{ f(t, x, u_1, u_2) + g^i(t, x, u_1) - V^2(t, x) - f_i(t, x, u_1, u_2) + g^*(t, x, u_1, u_2) \}$ ,  $\frac{d}{dt} V^1(t, x) = q^1(x)$ ,  $i = 1, 2$ . (6.74b) and  $\gamma^2$  is obtained from the linear equation  $\frac{d}{dt} V^2(t, x) = \min_{u_1} \max_{u_2} \{ f(t, x, u_1, u_2) + g^i(t, x, u_1) - V^3(t, x) - f_i(t, x, u_1, u_2) + g^*(t, x, u_1, u_2) \}$ ,  $\frac{d}{dt} V^2(t, x) = q^2(x)$ ,  $i = 1, 2$ . (6.74c) Note that (6.74a) is the Hamilton-Jacobi-Bellman equation associated with the optimal control problem obtained by setting  $\epsilon = 0$  in the differential game with weakly coupled players, and hence  $\gamma^k$  is the feedback representation

326 T. BAŞAR AND G. J. OLSDER of the open-loop policy  $u^k$  obtained in Remark 6.13, on a common zeroth-order trajectory  $x^k$ . Hence, to zeroth order, there exists a complete equivalence between open-loop and feedback Nash equilibrium solutions in weakly coupled differential games—as in the case of optimal control (see, for example, Bensoussan, 1988). For higher-order terms, however, no such correspondence exists in nonzero-sum differential games, which is one explanation for the nonequivalence of the equilibrium trajectories under different information structures—a feature we have observed throughout this chapter. □ Feedback saddle-point solution of zero-sum differential games We now first specialize the result of Thm. 6.16 to the class of two-person zero-sum differential games of prescribed fixed duration, and obtain a single partial differential equation to replace (6.69). Corollary 6.6 For a two-person zero-sum differential game of prescribed fixed duration  $[0, T]$ , and under either MPS or CLPS information pattern, a pair of strategies  $\{\gamma^i \in F^*; i = 1, 2\}$  provides a feedback saddle-point solution if there exists a function  $V : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  satisfying the partial differential equation  $-\frac{d}{dt} V(t, x) = \min_{u_1} \max_{u_2} \{ f(t, x, u_1, u_2) + g^i(t, x, u_1) - V(t, x) - f_i(t, x, u_1, u_2) + g^*(t, x, u_1, u_2) \}$ ,  $\frac{d}{dt} V(t, x) = q^i(x)$ ,  $i = 1, 2$ . Every such saddle-point solution is strongly time consistent, and the unique saddle-point value of the game is  $V(0, a_0)$ . Proof. This result follows as a special case of Thm. 6.16 by taking  $N = 2$ ,  $g^1 = -g^2$  and  $g^1 = -q^2$ , in which case  $V^1 = -V^2 = V$  and existence of a saddle point is equivalent to interchangeability of the min-max operations. □ Remark 6.19 The partial differential equation (6.75) was first obtained by Isaacs in the early

1950s (see Isaacs, 1954-1957; see also Isaacs, 1975) and is therefore generally referred to as Isaacs equation, and also as Hamilton-Jacobi-Isaacs equation. Furthermore, the requirement for interchangeability of the "min" and "max" operations in (6.75) is also known as the "Isaacs condition". For more details on this, the reader is referred to Chapter 8.  $\square$  We now obtain the solution of the Isaacs equation (6.75) for the case of a two-person zero-sum affine-quadratic differential game described by (6.60a)-(6.60b). Since the boundary condition is quadratic in  $x$  (i.e.,  $V(T, x) = q(x) = x'Qfx$ ), we try out a solution in the form  $V(t, x) = x'Z(t)x + x'f(t) + n(t)$ , where  $Z(-)$

NASH AND SADDLE-POINT EQUILIBRIA 327 is symmetric, with  $Z(T) = Qf$ , and  $f(T) = 0$ ,  $n(T) = 0$ . Substitution of this general quadratic structure in (6.75) yields the following equation:  $-x'Z(t)x - x'f(t) - m(t) \min^i \max^j [x'Z(t) + f(t)] [A + BV + B^2u^2 + c] + x'Q(t)x - u^1u^1 - u^2u^2$ . Since the kernel on the RHS is strictly convex (respectively, strictly concave) in  $u^1$  (respectively,  $u^2$ ), and is further separable, it admits a unique saddle-point solution given by  $u^* = y^0(x) = (-1)'B'(*)'[Z(*)x + at]$ ,  $i = 1, 2$ . If these unique values for  $u^1$  and  $u^2$  are substituted back into the RHS of the Isaacs equation, we finally obtain the following differential equations for  $Z(-)$ ,  $f(\bullet)$  and  $n(-)$ , by also taking into account the end-point conditions, the starting assumption that  $Z$  is symmetric (which is no loss of generality) and the requirement that equation (i) should be satisfied for all  $x \in \mathbb{R}^n$ :  $Z + A'Z + ZA + Q - Z^*B^1 - B^2B^2)'Z = 0$ ;  $Z(T) = Qf$ , (6.76a)  $f + A'C + Zc - Z^*B^1 - B^2B^2)'X = 0$ ;  $C(T) = 0$ , (6.76b)  $n + cX - c, \{B^1B^1 - B^2B^2\}' = 0$ ;  $n(T) = 0$ . (6.76c) The following theorem now even strengthens this result. Theorem 6.17 Let there exist a unique symmetric bounded solution to the generalized matrix Riccati differential equation (6.76a) on the interval  $[0, T]$ . Then, the two-person affine-quadratic zero-sum differential game described by (6.60a)-(6.60b), and under the MPS or CLPS information pattern, admits a unique feedback saddle-point solution given by  $u^i(t, x) = (-1)^i B^i(t)'[Z(t)x(t) + at]$   $t = 1, 2$ , (6.77) where  $f(\bullet)$  is obtained as the unique solution of (6.76b). The state trajectory associated with this unique pair of feedback saddle-point strategies is given by the solution of  $\dot{x}(t) = [A - \{B^1B^1 - B^2B^2\}'Z]x - \{B^1B^1 - B^2B^2\}'X f(t) + c(\bullet)$ ; (6.78a) subject to the initial condition  $x(0) = x_0$ , and the saddle-point value of the game is  $J^*(x_0, u^1, u^2) = x_0'Z(0)x_0 + x_0'a(0) + n(0)$ . (6.78b) If, furthermore,  $Q(-) > 0$  and  $Qf > 0$ , then the solution to (6.76a),  $Z(t)$ , is nonnegative definite for all  $t \in [0, T]$ .

328 T. BAŞAR AND G. J. OLSDER Proof. With the exception of the "uniqueness" of the feedback saddle-point solution, and the nonnegative definiteness of  $Z(\cdot)$ , the result follows as a special case of Corollary 6.6, as explained prior to the statement of the theorem. To verify uniqueness of the solution, one approach is to go back to the Isaacs equation and investigate existence and uniqueness properties of solutions of this partial differential equation. This, however, requires some advance knowledge on the theory of partial differential equations, which is beyond the scope of our coverage in this book. In its stead, we shall choose here an alternative approach which makes direct use of the affine-quadratic structure of the differential game. This approach will provide us with the uniqueness result, as well as the equilibrium strategies (6.77). Toward this end, let us first note that, under the condition of unique solvability of (6.76a), the linear differential equations (6.76b) and (6.76c) admit unique solutions. Furthermore,  $x(T)'Qf x(T)$  can be written as  $x\{T\}'Qf x\{T\} = x\{T\}'Z\{T\}x\{T\} + 2x\{T\}'C,\{T\} + 2m\{T\} = x'0Z\{0\}x0 + 2x'0(\{0\}+2m\{0\}) + \int_0^T \{x(t)'Z(t)x(t) + 2x\{t\}X\{t\} + 2m(t)\} dt$   $\int_0^T d' = x'oZ\{0\}xo + 2x'oC(0) + 2m(0) + \int_0^T \{x(t)'Z(t)x(t) + 2[x(t)'Z(t) + at'] Jo +x(t) + 2x\{t\}'C,\{t\} + 2m(t)\} dt$ , and using (6.76a)-(6.76c), the integrand can equivalently be written as  $x'[-A'Z -ZA-Q + Z\{BxB1' - B2B2'\}Z]x + 2[x\{t\}'Z\{t\} + \xi(*)]'(*) + 2x'[-AX -Zc + Z\{B1B1' - B2B2'\}Q - c'C + \c,'{B1B1' - B2B2'X}$ , where we have suppressed the  $\wedge$ -dependencies of various terms. If we now substitute for  $x(\cdot)$  its expression from (6.60a), we obtain (after canceling out some common terms)  $x\{T\}'Qf x(T) = x'oZ\{0\}xo + [ \{2u1'B1'\{Zx + Q + 2u2'B2'\{Zx + Q Jo - x'Qx + [Zx + Q'iB^1' - B2B2'] [Zx + Q dt + 2a/0C(0) + 2m(0)$ . If this identity is used in (6.60b), the result is the following equivalent expression for  $L\{u\},u2\}$ :  $rp L\{u\}u2) = ^x'0Z\{0\}x0+x'oaO)+m\{0\} + ]- [ lu1 + B1'[Zx + Q\}2 dt * 2 Jo 1 fl 2 B2'[Zx + Q\}2dt$ .

NASH AND SADDLE-POINT EQUILIBRIA 329 Since the first term above is independent of  $u_1$  and  $u_2$ ,  $L\{u_1,u_2\}$  is separable in  $u_1$  and  $u_2$ , and it clearly admits the saddle-point solution given by (6.77). The saddle-point value of  $L\{v\},u_2)$  also readily follows as  $^x'oZ(0)xo+x'o^(0)+m(0)$ . To prove that  $\blacksquare \xi(\cdot) > 0$  whenever  $Q(\cdot) > 0$  and  $Qf > 0$ , follow the line of reasoning used in verifying the last statement of Thm. 6.7 (for the discrete-time counterpart). With  $c(\cdot) = 0$ , the value function is  $V(t,x) = ^x'Z(t)x$ , which is necessarily nonnegative since the maximizer can assure such a lower bound by simply choosing  $u_2(\cdot) = 0$ . D Remark 6.20 As in the case of Nash equilibria (cf.

Remark 6.17), the result above extends naturally to the more general affine-quadratic zero-sum dynamic games which have in the cost function additional linear terms in  $x$ , to be denoted  $x(T)$  for the terminal cost term, and  $x(t)$  for the integrand. Then, the only modification will have to be made in equation (6.76b), for  $\mathcal{L}(\cdot)$ , which will now read:  $\mathcal{L}(x) = x(T)' \mathcal{L}_T x(T) + \int_0^T [x(t)' \mathcal{L}(t) x(t) + u(t)' \mathcal{L}_u(t) u(t) - x(t)' \mathcal{L}_x(t) x(t)] dt$ ;  $\mathcal{C}(T) = \mathcal{L}_T x(T)$ . Otherwise, the statement of Thm. 6.17 remains intact.  $\square$

Remark 6.21 A comparison of Thms. 6.14 and 6.17 reveals one common feature of the open-loop and feedback saddle-point solutions in affine-quadratic differential games with  $Q(\cdot) > 0$ ,  $Q_f > 0$ , which is akin to the one observed for games described in discrete time (cf. Remark 6.7). The two matrix Riccati differential equations (6.62a) and (6.76a) are identical, and so are the two linear differential equations for  $m$  and  $\mathcal{L}$  ((6.62b) and (6.76b), respectively) and by this token the state trajectories generated by the open-loop saddle-point solution and feedback saddle-point solution are identical; and furthermore the open-loop values of the feedback saddle-point strategies are correspondingly the same as the open-loop saddle-point strategies. A comparison of the existence conditions of Thms. 6.14 and 6.17, on the other hand, readily leads to the conclusion that existence of an open-loop saddle-point solution necessarily implies existence of a feedback saddle-point solution (since then (6.62a) (equivalently (6.76a)) admits a unique bounded symmetric solution), but not vice versa. This latter feature is a consequence of the following observation: for the class of affine-quadratic differential games satisfying the restriction  $B(t)'B(t) - B_2(t)'B_2(t) > 0 \forall t \in [0, T]$ , (6.79) the matrix Riccati differential equation (6.76a) admits a unique nonnegative definite solution, whenever  $Q_f > 0$ ,  $Q(\cdot) > 0$  (since then (6.76a) becomes comparable with (5.35a)); however, (6.79) is not sufficient for (6.61) to admit a bounded solution on  $[0, T]$ . This situation is now exemplified in the sequel.  $\square$

Example 6.3 Consider a scalar linear-quadratic two-person zero-sum differential game described by  $\dot{x} = V_2 u_1 - u_2$ ;  $x(0) = x_0, x_0 \in [0, 2]$ ;

330 T. BAŞAR AND G. J. OLSDER  $L\{u\} = \int_0^2 [x(t)]^2 + 1 - \int_0^2 [u(t)]^2 - \int_0^2 u_2(t)^2 dt$ . For these parametric values, the Riccati differential equation (6.76a) can be written as  $Z - Z^2 = 0$ ;  $Z(2) = 1$  which admits the unique solution  $Z(t) = 1 - (2-t)^2$ . Hence, the differential game under consideration admits the unique feedback saddle-point solution  $u_1(t, x) = -[2/(3-t)]x(t)$ ,  $u_2(t, x) = -[V(3-t)]x(t)$ . Let us now investigate the existence of an open-loop saddle-point solution for the same scalar differential game. The Riccati differential equation (6.61) can be

written, for this example, as  $S + S^2 = 0$ ;  $S(T) = 1$  ( $t \in [0, T], T = 2$ ) which admits a bounded solution only if  $T < 1$ , in which case For  $T > 1$ , the Riccati differential equation does not admit a solution, and it displays (in retrograde time) finite escape at  $t = T - 1$ ; i.e., it has a conjugate point at  $t = T - 1$ . In our case, since  $T = 2 > 1$ , the existence condition of Thm. 6.14 is thus not satisfied, and therefore an open-loop saddle-point solution does not exist. This implies, in view of Lemma 6.4, that under the open-loop information structure  $L(u_1, u_2)$  is not concave in  $u_2$ , and hence the upper value of the game is unbounded (since it is scalar and quadratic).  $\square$  Even though feedback saddle-point equilibrium requires less stringent existence conditions than open-loop saddle-point equilibrium in affine-quadratic differential games (as also exemplified in the preceding example), this feature does not necessarily imply (at the outset) that the least stringent conditions for existence of saddle-point equilibrium under closed-loop information structure are required by the feedback saddle-point solution. This also brings up the question of whether the affine-quadratic differential game still admits a saddle-point solution in spite of existence of a conjugate point to the matrix Riccati equation (6.76a) on the interval  $[0, T]$ . In other words, can a saddle point survive a conjugate point? Intuitively, this could happen if at a conjugate point  $t^*$ , while  $Z$  goes off to infinity the corresponding state trajectory  $x^*$  goes to zero at least at the same rate, so that the control actions dictated by the feedback strategies

NASH AND SADDLE-POINT EQUILIBRIA 331 remain (square-) integrable, and that  $Z$  does not change sign in any small open neighborhood of  $t^*$ . To permit for such behavior, the allowable strategies can no longer be continuously differentiable, or even continuous, but their realizations will have to be square-integrable for every square-integrable open-loop policy of the opponent player. Bernhard (1979) has shown, by working in this more general class of strategies, that a saddle point may survive a conjugate point, if the conjugate point is of an "even" type. To give a brief description of this result, let us first note that in the absence of a conjugate point, the solution of (6.76a) can be decomposed as  $Z(t) = Y(t)X^{-1}(t)$ , (6.80) where  $X$  and  $Y$  are  $n \times n$  matrices satisfying the linear matrix differential equations  $\dot{X} = AX - \{B'B_1' - B_2B_2'\}Y$ ,  $X(T) = I$ , (6.81a)  $\dot{Y} = -QX - A'Y$ ,  $Y(T) = Qf$ . (6.81b) Clearly, being linear, these equations admit unique solutions regardless of the length of the time interval, and hence the conjugate point of  $Z$  is the first instant (in retrograde time) when  $X$  becomes singular. Let us adopt the convention that at



points where  $X$  is singular,  $X^{-1}$  in (6.80) is replaced by its pseudo-inverse  $X^{\dagger}$ . This then extends the solution of (6.76a) uniquely beyond a conjugate point. The following result, which is from (Bernhard, 1979), now provides the necessity part of Thm. 6.17. **Theorem 6.18** Consider the affine-quadratic zero-sum differential game described by (6.60a)-(6.60b), with  $Q(-) > 0$ ,  $Q_f > 0$ , and under the CLPS information structure. The game admits a saddle point (with finite value) if, and only if, the following four conditions are satisfied: (i)  $x_0 \in \text{Range}(X(0))$ , (ii) rank of  $X(t)$  is piecewise constant, (iii)  $\text{Range}(B(t)) \subset \text{Range}(X(t))$ ,  $\forall t \in [0, T]$ , except, perhaps, at isolated points, (iv)  $Z(t) > 0 \forall t \in [0, T]$ , under the extended definition of a solution provided by (6.80) and (6.81a)-(6.81b). Then, the pair of feedback strategies (6.77) provides a saddle-point solution, and the saddle-point value is given by (6.78b). If the conditions above are not satisfied, then the upper value of the game becomes unbounded. The following example (again due to Bernhard (1979)) illustrates the foregoing result, and demonstrates the occurrence of an "even" conjugate point.

332 T. BAŞAR AND G. J. OLSDER **Example 6.4** Consider the time-varying scalar linear-quadratic differential game described by  $\dot{x} = (2-t)u_1 + tu_2$ ,  $x(0)=1$ ,  $L\{u_1, u_2\} = 1 - [x(t)]^2 + \int_0^t \{u_1^2(t) - u_2^2(t)\} dt$ . The Riccati equation (6.76a) in this case is  $Z - 4(1-t)Z^2 = 0$ ,  $Z(2) = \pm 1$ , which has a conjugate point at  $t = 1$ . Using the construction (6.80), however, we can obtain an extended solution as  $Z(t) = 2(1-t)^{-1} e^{2\int_t^1 (1-t) dt}$ . This shows that the conjugate point  $t^* = 1$  is of the even type (i.e.,  $Z(t)$  does not change sign in an open neighborhood of  $t^*$ ). Using this solution, the strategies (6.77) can be written as  $u_1^*(t, x) = -x$ ,  $u_2^*(t, x) = 2(1-t)^{-1} x$  with the corresponding state trajectory satisfying  $\dot{x} = 1 - tx$ ,  $x^*(0) = 1$ . Note that, for  $t < 1$ ,  $x^*(t) = (1-t)^{-1}$ , and hence open-loop representations of both  $u_1$  and  $u_2$  remain bounded as  $t \rightarrow 1$ , and so does the value function. Note also that all four conditions of Thm. 6.18 are satisfied in this case, with  $X(t) = (1-t)^{-2}$ .  $\square$  Another perspective on the relationship between existence of conjugate points and saddle points in zero-sum affine-quadratic games can be obtained by studying the dependence of the solution on a scalar parameter. Of particular importance (as we shall later see in Section 6.6) is the case when the parameter is chosen as the weighting on the control of the maximizer. Consider, therefore, the modified cost (in place of (6.60b))  $L\{u_1, u_2\} = -\int_0^T (x'(t)Qx(t) + u_1(t)V(t) - Au_2(t)V(t)) dt + \int_0^T (6.82) + x'(T)Qfx(T)$ , where  $A$  is a scalar positive parameter,  $Q(-) > 0$  and  $Q_f > 0$ . Let the state

dynamics be again described by (6.60a). Then, the counterpart of the generalized Riccati differential equation (6.76a) is  $Z + A'Z + ZA + Q - Z^{\wedge}B1' - \sim B2B2' \backslash z = 0; Z(T) = Qf$ , (6.83)

NASH AND SADDLE-POINT EQUILIBRIA 333 whose solution over the interval  $[0, T]$  we will denote by  $Z(-)$ . Note that the limiting case  $A = \infty$  yields the Riccati equation (5.35a) associated with the one-player game (with the maximizer being absent), which admits a nonnegative definite solution. Hence, by continuity, for every fixed  $T$  there exists some range of values of  $A$  for which (6.83) does not have any conjugate point on  $[0, T]$ . What can in fact be shown (see Başar and Bernhard (1995)) is that there exists some threshold value of  $A$ , say  $A^*$ , such that (i) for all  $A > A^*$ , (6.83) has no conjugate points (and hence the game has a finite value under CLPS information), (ii) for all  $A < A^*$ , (6.83) has a conjugate point, and (iii) for all  $A < A^*$ , the upper value of the differential game is unbounded (regardless of the underlying information structure). We summarize this result in the following theorem. Theorem 6.19 Consider the parameterized affine-quadratic zero-sum differential game described by (6.60a) and (6.82), with  $Q(-) > 0, Qf > 0$ . There exists a critical value  $A^* > 0$  for  $A$  such that (i) For each  $A > A^*$ , (6.83) admits a unique nonnegative definite solution  $Zx(-)$  (which is positive definite if either  $Qf > 0$  or  $Q(-) > 0$ ), and the game has a finite saddle-point value, given by (6.78b) with  $Z$  replaced by  $Zx$ , and  $B2B2'$  in (6.76b) and (6.76c) replaced by  $\backslash B2B2'$ . (ii) For  $A < A^*$ , (6.83) has a conjugate point. (iii) For  $A < A^*$ , the upper value of the game is unbounded, even if the minimizer has access to full information (on the control action of the maximizer). (iv) For  $A > A^*$ , the unique feedback saddle-point strategies of the players are  $7r(t, x) = -B1'[Zx\{t\}x + Ut); j2''(t, x) = jB2'[Zx(t)x + <;x(t)\}$ ; where  $Ca(-)$  is given by (6.76b) with  $Z\backslash$  replacing  $Z$  and  $jB2B2$  replacing  $B2B2'$ . Proof. A proof of this result for the linear case (with  $c = 0$ ) can be found in Chapter 9 of Başar and Bernhard (1995), which readily extends to the affine case in view of Thm. 6.17. D 6.5.3 Linear-quadratic differential games on an infinite time horizon In this subsection, we will study the limiting cases of the Nash equilibrium solution of the linear-quadratic differential game as the time horizon becomes unbounded—first for the open-loop case and then for the closed-loop feedback case.

334 T. BAŞAR AND G. J. OLSDER The open-loop case Here, we will study the limiting case of Thm. 6.12 as the time horizon becomes unbounded. The

starting point will be Def. 6.5, with the assumptions that  $N = 2$  and all matrices involved are time-invariant, and  $Q \setminus = Q_2, \dots = 0$ . Moreover, we will restrict ourselves to the linear-quadratic case, i.e.,  $c = 0$ . The so-called algebraic Riccati equations (compare with (6.49))  $M_1 X A + A' M_1 X + Q_1 - N^* M_1 - M_1 X S_2 M_2 = 0$ , (6.84)  $M_2 A + A' M_2 + Q_2 - M_2 S_2 M_2 - M_2 S_1 M_1 = 0$ , (6.85) where, as before,  $S_i = B_i \{R_i\}^{-1} B_i^T$ ,  $i = 1, 2$ , play an important role here. Under certain technical conditions solutions of (6.49), with  $N = 2$ , will approach the solutions of (6.84) and (6.85) as  $T \rightarrow \infty$ . Since the dependence of the solutions of (6.49) on the final time  $T$  will be emphasized, these solutions will be written as  $M^*(t, T)$ . We will elaborate on the relationship between the solutions of these equations and the matrix  $G$ : (6.86) Define  $G_{mv} = \{K \setminus G K \subset \mathbb{R}^n\}$ , i.e.,  $G_{mv}$  is the set of all  $G$ -invariant subspaces of  $\mathbb{R}^n$ . It is well known that this set consists of a finite number of (different) elements if and only if all eigenvalues of  $G$  have geometric multiplicity one (otherwise there is an infinite number of different elements). Next define the subset  $G_{sub}$  of  $G_{mv}$  such that  $w \in G_{sub} \iff \exists T > 0$  such that  $\{e^{G(t-T)} w \mid 0 \leq t \leq T\} \subset \mathbb{R}^n$ , where  $\exists$  refers to "image of". Note that the elements in  $G_{sub}$  are the images of suitably chosen matrices of size  $3n \times n$ . In Engwerda (1998) the following theorem has been proved. Theorem 6.20 The set of coupled algebraic Riccati equations (6.84) and (6.85) have a real solution  $(M_1, M_2)$  if, and only if,  $M_i = X_i X_i^{-1}$ ,  $i = 1, 2$  for matrices  $X_1, X_2$  such that  $X_i X_i^{-1} C = 0 \setminus X_i \in \mathbb{R}^n$ . Moreover, if the control functions  $u_i^*(t) = -(R_i^{-1})^* B_i^* M_i^*(t, Q) x(t)$  are used in the system equation  $\dot{x} = Ax + \sum_{i=1}^2 B_i u_i$ , the eigenvalues of the closed-loop matrix  $A - \sum_{i=1}^2 B_i M_i^*$  coincide with the eigenvalues of  $G \setminus K$ , which is the restriction of  $G$  to the subspace  $K$ .

NASH AND SADDLE-POINT EQUILIBRIA 335 It follows that every element of  $G_{sub}$  defines one solution of (6.84) and (6.85). If the geometric multiplicity of all eigenvalues of  $G$  equals one, then the set  $G_{sub}$  has a finite number of different elements and consequently (6.84) and (6.85) have at most a finite number of real solutions. Equations (6.84) and (6.85) do not have a real solution if and only if  $G_{sub}$  is empty. An immediate consequence of Thm. 6.20 is the following corollary. Corollary 6.7 The algebraic Riccati equations (6.84) and (6.85) have a set of stabilizing solutions (i.e., the closed-loop matrix  $A - \sum_{i=1}^2 B_i M_i^*$  has all its eigenvalues in the left-half plane) if, and only if a  $G$ -invariant subspace  $K \in G_{sub}$  exists such that  $\text{Re } \lambda > 0$  for all  $\lambda \in \sigma(G \setminus K)$  (the symbol  $\sigma$  refers to the spectrum of its argument). We now introduce

a useful concept, which will lead to a technical condition needed in the main result. Definition 6.7 The  $3n \times 3n$  matrix  $G$  is called dichotomically separable if sub-spaces  $V_1$  and  $V_2$ , with  $\dim V_1 = n$  and  $\dim V_2 = 2n$ , exist such that  $GV_i \in V_i \forall i = 1, 2$ ,  $V_1 \oplus V_2 = \mathbb{R}^{3n}$  and  $RA > R/x$  for all  $A \in \langle r(G \setminus V_1), \setminus x \in a \{G \setminus V_t\}$ . It can be shown that dichotomic separability implies that  $G_{sub}$  is nonempty, but the reverse implication is not necessarily true. Theorem 6.21 Assume that  $H(T)$  is invertible for all finite  $T$ ,  $G$  is dichotomically separable and  $\text{span } I Q) ] \oplus V_2 = \mathbb{R}^{3n}$ . Then,  $M^*(0, T) \rightarrow X_j X$ ,  $i = 1, 2$  if  $T \rightarrow \infty$ , where  $X$  and  $X_i$  are defined as  $V_i = \text{span}(X^{\wedge} X_2)'$ . Corollary 6.7 and Thm. 6.21 immediately lead to the following corollary. Corollary 6.8 // the time horizon of the linear quadratic differential game tends to infinity, then the unique open-loop Nash equilibrium converges to  $u^*(t) = -\{R^T M\}^{-1} B^T M(0, t) x_0$ ,  $i = 1, 2$ , which leads to a stable system, if the following conditions are satisfied: • The conditions of Thm. 6.21 are fulfilled; •  $\forall t > 0$  for all  $X \in \langle T \{G \setminus V_1\}$ . For the cost function of each player to remain bounded, we assume that  $Q_i > 0$ ,  $i = 1, 2$ , and under the equilibrium solution the state of the system converges to zero as  $t$  tends to infinity.

336 T. BAŞAR AND G. J. OLSDER Theorem 6.22 Let  $M_i$ ,  $i = 1, 2$ , be solutions of (6.84) and (6.85) satisfying the additional constraint that the eigenvalues of the resulting system matrix  $A - \sum_{i=1}^2 S_i M_i$  are all in the left-half plane. Then the controls  $u_i^*(t) = -\{R_i I\}^{-1} B_i^T M_i(t) x_0$ ,  $i = 1, 2$ , form an open-loop Nash equilibrium solution. Proof. Let  $L_i(u_1, u_2; T)$  denote the cost function of  $P_i$  on the time interval  $[0, T]$ . It can then be shown that the optimization problem  $\min \lim_{T \rightarrow \infty} L_i(u_1, u_2^*; T)$ ,  $u_1 \in \mathbb{R}^{2n}$  admits a solution given by  $u_1^*(t) = -\{R_1 I\}^{-1} B_1^T M_1(t) x_0$ . A similar reasoning holds for the optimization problem  $\min \lim_{T \rightarrow \infty} L_2(u_1^*, u_2; T)$ , leading to the conclusion that  $(u_1^*, u_2^*)$  forms an open-loop Nash equilibrium solution. D The following example now demonstrates the somewhat surprising fact that an infinite-horizon differential game may admit a Nash equilibrium solution in open-loop policies, while its finite-horizon version (with a particular terminal state weighting) may not. Example 6.5 Reconsider Example 6.2 and replace the  $A$ -matrix by  $-CD$ . The eigenvalues of the related  $G$ -matrix are  $\{1.8810, 0.1883, -1.7966\}$ . Further numerical calculation shows that the algebraic Riccati equations have three stabilizing solutions. Hence, according to Thm. 6.22, the infinite-horizon problem has at least three open-loop Nash equilibrium solutions. On the other hand, the

matrices  $Q_i, i = 1, 2$ , can suitably be chosen such that the matrix  $G(t)$  fails to be invertible for certain  $t$ , e.g.,  $t = 0.1$  as in Example 6.2. Thus the infinite-horizon problem can have solutions whereas the finite-horizon problem fails to have a solution.  $\square$  One might conjecture that if the finite-horizon game has a unique solution for all finite  $T$ , then this solution converges to a solution of the infinite-horizon game. This is not true, as shown in Engwerda (1998), which presents an example in which the limiting solution does lead to a solution of the coupled algebraic Riccati equations, but fails to be a stabilizing solution. The closed-loop feedback case We now study the limiting cases of Corollary 6.5 and Thm. 6.17 (as well as Thm. 6.19) as the time horizon becomes unbounded (that is,  $T \rightarrow \infty$ ), when

NASH AND SADDLE-POINT EQUILIBRIA 337  $c = 0, Q_f = 0$ , and all matrices are time-invariant. This will constitute the counterparts of the discrete-time results of subsection 6.2.3 in the continuous time. First we consider the convex-quadratic Nash differential game covered by Corollary 6.5, and note that any asymptotically convergent (as  $T \rightarrow \infty$ ) limit of the solution of (6.70a) will have to satisfy the coupled algebraic Riccati equations:  $Z_i F + F' Z_i + \sum_j B_j R_j^{-1} R_{ii}^{-1} B_j' z_j + Q_i = 0$ , (6.87a)  $J \in \mathbb{N}$  where  $F$  is given by (6.70b), with all matrices replaced by their time-invariant counterparts.<sup>63</sup> The limiting stationary feedback (Nash) strategies are (from (6-71))  $u_i(x) = -R_{ii}^{-1} B_i' z_i, i \in \mathbb{N}$ . (6.87b) Furthermore, introduce for each  $i \in \mathbb{N}$  the matrices  $F_x = A - \sum_j B_j R_j^{-1} B_j' Z_j$ ;  $Q_i^0 = Q_i + \sum_j B_j R_j^{-1} B_j' Z_j$ .  $j \in \mathbb{N}$  Then the following result is the natural counterpart of Prop. 6.3 in the continuous time. Proposition 6.8 Let there exist an  $N$ -tuple of matrices  $\{Z_i, i \in \mathbb{N}\}$  satisfying (6.87a), and further satisfying the conditions that for each  $i \in \mathbb{N}$  the pair  $(F_i, B_i)$  is stabilizable and the pair  $(F_i^0, Q_i)$  is detectable.<sup>64</sup> Then, (i) the  $N$ -tuple of stationary feedback policies (6.87b) provides a Nash equilibrium solution for the linear-quadratic nonzero-sum differential game of this subsection, leading to the infinite-horizon Nash equilibrium cost  $\int_0^\infty x' Z_i^0 x$  for  $P_i$ , (ii) the resulting system dynamics, described by  $\dot{x} = F_x x, x(0) = x_0$  are asymptotically stable. Again, as in the discrete-time case, there could exist a Nash equilibrium for the infinite-horizon game (as covered by Prop. 6.8) which is not obtainable as the limit of any Nash equilibrium of the time-truncated finite-horizon version. Furthermore, we do not have any direct conditions (on the system parameters) that would guarantee satisfaction of the indirect conditions of Prop. 6.8 for any  $i$ .<sup>63</sup>Note that here we are using the same

notation as in the discrete-time case, but this should not create any confusion.  
 64 See subsection 5.5.2 for the terminology.

338 T. BAŞAR AND G. J. OLSDER sufficiently general class of nonzero-sum differential games. For zero-sum games, however, more explicit results can be obtained as will be discussed next. Consider the class of zero-sum differential games described by (6.60a)-(6.60b), with  $c = 0$ ,  $Q_f = 0$ , all matrices time-invariant,  $Q > 0$ , and  $T = \infty$ . As in subsection 6.2.3, we would expect the feedback saddle-point solution for this game to be in the form  $J(x) = -B_1'Zx$ ; (6.88a)  $u_2(x) = B_2'Zx$ , (6.88b) where  $Z$  is the limiting solution of (6.76a) which (if it exists) should satisfy the continuous-time generalized algebraic Riccati equation (GARE)  $A'Z + ZA + Q - Z\{B_1B_1' - B_2B_2'\}Z = 0$ . (6.89) Denote the solution of the generalized Riccati differential equation (6.76a) by  $Z(t;T)$ , where dependence on the terminal time is explicitly recognized. Now, the first question that would be of interest is: If  $Z(t; T)$  is bounded for all  $T > 0$  (i.e., (6.76a) does not have any conjugate points on  $[0, T]$  for any  $T$ ) and  $\lim_{T \rightarrow \infty} Z(t, T) = Z$  exists, does this necessarily imply that the pair (6.88a)-(6.88b) is in saddle-point equilibrium? The answer is "no", as the following (counter) example, taken from (Mageirou, 1976) clearly illustrates. Example 6.6 Consider the scalar state dynamics  $\dot{x} = x + u_1 + u_2$ ,  $x(0) = x_0$  and the objective function  $J = \int_0^{\infty} L(u_1, u_2) dt$ . The generalized Riccati differential equation (6.76a) associated with this game is  $Z + 2Z + 1 - Z^2 = 0$ ;  $Z(T;T) = 0$  which admits the unique solution  $Z(t;T) = \sqrt{6} \tanh[\sqrt{6}(T-t) + \tanh^{-1}(\sqrt{6}Z(T;T))]$ . This function has a well-defined limit, as  $T \rightarrow \infty$ , which is  $t$ -independent:  $Z = 2 + \sqrt{6}$ . Hence the policies (6.88a)-(6.88b) are  $u_1(x) = -(2 + \sqrt{6})x$ ;  $u_2(x) = \sqrt{6}x$ .

NASH AND SADDLE-POINT EQUILIBRIA 339 Note that the closed-loop system under these policies is  $\dot{x} = -\sqrt{6}x$  which is asymptotically stable. Also, the system only under  $u_1$  is  $\dot{x} = -(1-\sqrt{6})x + u_2$  which is bounded input-bounded state stable. Now it is easy to see (using Prop. 5.4) that if  $u_1 = u_1^*$  the policy for  $P_2$  that maximizes  $J$  is  $u_2^*$ . However, when  $u_2 = u_2^*$   $u_1^*$  is not a minimizing solution for  $P_1$ . The underlying minimization problem is  $J = \int_0^{\infty} (4 + \frac{1}{\epsilon} |x| + u_1) dt \rightarrow \text{minimum}$ . Since the open-loop system is unstable, and there is a negative cost on state in the cost function, the cost can be driven to  $-\infty$  (pick, for example,  $u_1 = 0$ ). Even though there is no continuity in the minimizer's policy at  $T = \infty$ , nevertheless the value

of the game (even in the absence of a saddle point) is continuous. It can be shown that given an  $\epsilon > 0$ , there exists a  $t_c > 0$  such that by choosing the time-varying feedback policy  $u(t, x) = (VZ^* \text{fo}^* M^*)^{-1} (VZ^* \text{fo}^* M^* - (2 + \sqrt{6})(x_0)^2 - \epsilon(x_0)^2)$ ,  $0 < t < t_c$ ,  $P_2$  can assure that  $\min L(u_1, u_2) > (2 + \sqrt{6})(x_0)^2 - \epsilon(x_0)^2$ , where  $(2 + \sqrt{6})(x_0)^2 = \lim_{T \rightarrow \infty} Z(t; T)$  which is the limit of the value of the finite-horizon game (as  $T \rightarrow \infty$ ).<sup>65</sup> Since  $\epsilon > 0$  can be chosen arbitrarily small, this shows that the infinite-horizon game of this example indeed has a value.  $\square$  What has been demonstrated in the last part of the example above actually turns out to be the case for the general linear-quadratic differential game, under some general conditions, as stated in the following theorem, which is the counterpart of Thm. 6.8, as well as of Lemma 6.3. 't above is chosen such that  $Z(0; t_c) + \epsilon > 2 + \sqrt{6}$ .

340 T. BAŞAR AND G. J. OLSDER Theorem 6.23 Consider the infinite-horizon time-invariant linear-quadratic differential game, where  $Q_f = 0$ ,  $Q > 0$  and the pair  $(A, Q)$  is observable. Then, (i) for each fixed  $t$ , the solution to (6.76a),  $Z(t; T)$ , is nondecreasing in  $T$ ; that is, if (6.76a) has no conjugate point in a given interval  $[0, T]$ ,  $Z(t; T') - Z(t; T'') > 0$ ,  $T > T' > T'' > 0$ ; (ii) every nonnegative definite solution of (6.89) is in fact positive definite, and if there exists a positive definite solution there is a minimal such solution, denoted  $Z^+$ . This matrix has the property that  $Z^+ - Z(t; T) > 0$  for all  $T > 0$ ; (iii) the game has equal upper and lower values if, and only if, the GARE (6.89) admits a positive definite solution, in which case the common value is  $J^{\circ\circ*} = x_0' Z^+ x_0$ ; (iv) the upper value is finite if, and only if, the upper and lower values are equal; (v) if  $Z^+ > 0$  exists, as given above, the controller<sup>1</sup> given by (6.88a), with  $Z$  replaced by  $Z^+$ , attains the finite upper value, in the sense that  $\sup_{J_1} |y| - \frac{1}{2} = x_0' Z^+ x_0$ , and the maximizing feedback solution above is given by (6.88b), again with  $Z = Z^+$ ; (vi) the minimal positive definite solution of (6.89),  $Z^+$ , is feedback stabilizing, i.e., the closed-loop system  $\dot{x} = [A - (B_1 B_1' - B_2 B_2')] Z^+ x$  has no eigenvalues in the right-half plane. If  $Q > 0$ , then  $Z^+$  is the only feedback stabilizing solution of (6.89). Proof. See Mageirou (1976) or Başar and Bernhard (1995).  $\square$  The result of Thm. 6.23 above is not the strongest possible, since the given policies (6.88a)-(6.88b) are not strictly feedback stabilizing (i.e., the closed-loop feedback matrix could have some eigenvalues on the imaginary axis of the complex plane) and furthermore the policies are not necessarily in saddle-point equilibrium (as also observed in Example 6.6). In view of this one would like to obtain conditions under which the policies (6.88a)-(6.88b) would be in saddle-point equilibrium and/or they

would be strictly feedback stabilizing. If we assume that there exists a solution to GARE (6.89) that is strictly feedback stabilizing, then it can be shown (see Jacobson (1977)) that the strategy pair (6.88a)-(6.88b) is in saddle-point equilibrium in the restricted class of feedback policies that are strictly feedback stabilizing and under which  $x(t) \rightarrow 0$  as  $t \rightarrow \infty$  for all  $x_0 \in \mathbb{R}^n$ . This result follows essentially from a completion

NASH AND SADDLE-POINT EQUILIBRIA 341 of squares argument, using the GARE (6.89). An alternative condition, again under the assumption of existence of a strictly feedback stabilizing solution to (6.89), is to require that under 72, the minimization problem faced by PI is well defined; this is equivalent to requiring (using some results of Willems (1971)) that an associated ARE has a negative definite solution (see Mageirou (1976)). But, this is a rather complicated condition, and admits no clean interpretation in terms of the parameters of the game. Yet another set of conditions can be obtained, by parameterizing the differential game, as in the finite-horizon case discussed earlier. Consider the infinite-horizon version of the cost function (6.82), where  $A$  is again a scalar parameter (to be varied). For  $A = \infty$ , the underlying problem is essentially a single player game, which is the same as the linear regulator problem discussed in subsection 5.5.2. We know from Prop. 5.4 that if  $(A, B)$  is stabilizable and  $(A, Q)$  is observable (respectively, detectable), then the associated ARE admits a positive (respectively, nonnegative) definite solution, which is further strictly feedback stabilizing. Since eigenvalues of the feedback matrix in the game case are continuous functions of  $A$ , the result will also hold for some range of finite values for  $A$ . This result is now made precise under also a weaker condition of detectability in the following theorem, which is the infinite-horizon counterpart of Thm. 6.19. Its proof can be found in Başar and Bernhard (1995). Theorem 6.24 For the parameterized infinite-horizon linear-quadratic differential game formulated above, let  $(A, B)$  be stabilizable and  $(A, Q)$  be observable (respectively, detectable). Then, there exists a threshold value  $A^* > 0$  for  $A$  such that the following three properties hold. (i) For every  $A > A^*$ , the GARE  $A'Z + ZA + Q - Z\{B'XB + B'Z - 1\}Z = 0$  (6.90) admits a minimal positive (respectively, nonnegative) definite solution,  $Z$ . (ii) For every  $A > A^*$ , the game has a finite upper value, which is achieved by the pair of strategies  $\gamma_1(x) = -B'Z + X$ ,  $H(x) = \frac{1}{2}z'xz$ . (6.91) Furthermore, the closed-loop matrix  $FX = A - \{B'Z + X - 1\}Z$  is Hurwitz (has all its eigenvalues in the open left-



half plane). (Hi) For  $A < A^*$ , the upper value of the game is unbounded.

342 T. BAŞAR AND G. J. OLSDER 6.6 Applications in Robust Controller Designs:  $H^\infty$ -Optimal Control The results presented in the previous section on zero-sum differential games, and especially on linear-quadratic games, as well as their counterparts in Section 6.3, have important applications in a special class of worst-case controller design problems, known as "jET $^\infty$ -optimal control" problems (see Zames (1981); Francis (1987); Doyle et al. (1989); Başar and Bernhard (1995); Stoorvogel (1992), Green and Limebeer (1995) to cite just a few representative papers and books from this voluminous literature). In these design problems, the (linear) plant has two types of inputs (controlled and disturbance) and two types of outputs (regulated and measured), and the objective is to obtain a controller (that uses the measured output) such that the effect of the disturbance on the regulated output is minimized. Even though both frequency and time domain formulations are possible, the latter is preferable since it allows for transient analysis and can handle time varying as well as nonlinear plants. In this section we first present a general formulation of this worst-case design problem, and then show how the theory of the previous section can directly be used to construct minimax (jET $^\infty$ -optimal) controllers. The presentation will be brief, and be restricted to the continuous-time problem only. Further details, and counterparts of these results in the discrete-time, can be found in the book by Başar and Bernhard (1995). To introduce the jET $^\infty$ -optimal control problem (in the continuous-time), let us adopt the plant dynamics  $\dot{x} = Ax + Bu + Dw$ ,  $x(0) = 0$  (6.92a) to replace the linear state equation (6.60a). Here  $u$  is the controlled input,  $w$  is the disturbance, and  $A$ ,  $B$ ,  $D$  are matrices which could have time-varying entries if the time horizon is not infinite. Let the regulated and measured outputs be given, respectively, by  $z = Hx + Gu$ , (6.92b)  $y = Cx + Ew$ , (6.92c) where  $H$ ,  $G$ ,  $C$ ,  $E$  are again appropriate dimensional matrices, with possibly time-variant entries for the finite-horizon problem. The controlled input is chosen as a function of the present and past values of the measured output  $y$ , a relationship that we will write as  $u(t) = n(t, y(s), s < t)$ ,  $t \geq 0$ , (6.92d) or simply as  $u = n(y)$ . Here  $n$  is the (causal) control policy. If in (6.92c)  $C = I$  and  $E = 0$ , then we have the CLPS information pattern, where the control has access to all components of the state. Now for any given  $x \in M$ , where  $M$  is chosen so that some measurability and smoothness conditions are satisfied, substitution of  $u$  into (6.92a) and (6.92c)

NASH AND SADDLE-POINT EQUILIBRIA 343 leads to a unique mapping from  $w$  to  $z$ , obtained from,  $z = Hx + G/j,(Cx + Ew)$ , (6.93a)  $x = Ax + B\backslash x(Cx + Ew) + Dw$ ,  $x(0) = 0$ . (6.93b) Let us denote this mapping by  $Q^\wedge$ , i.e.,  $z = g^\wedge w$ . (6.93c) Then, the  $jET^{\circ\circ}$ -optimal control problem is to minimize the induced (operator) norm of  $Q^\wedge$ , that is, to solve the optimization problem  $\inf \sup \{ \|S, \gg \|U/HU\}$ , (6-94)  $H \in M$   $w \in \text{-ft}$  where  $\| \cdot \|_z$  and  $\| \cdot \|_w$  denote Hilbert space norms on the regulated output and disturbance spaces, respectively, where the latter space is denoted by  $H$ - More specifically, for the finite-horizon problem, defined on the interval  $[0, T]$ ,  $\text{rp rp } \|z\|^2 z^\wedge'(T) Qfz(T) + \int_0^T z(t)' z(t) dt$ ,  $H|^* = / w(t)' w(t) dt$   $Jo Jo$  and for the infinite-horizon problem  $T = \infty$ , and  $Qf = 0$ . Note that (6.94) defines the upper value of a differential game where the minimizer ( $P_1$ ) is the controller, and the maximizer ( $P_2$ ) is the disturbance. Let the square of this upper value be denoted by  $A$ , and assume for the moment that there exists a control policy,  $\backslash x^*$ , that achieves it. Then, (6.94) can equivalently be expressed as (where we suppress the subindices on different norms) (i)  $\|6M-HI\|^2 < A \|w\|^2$ ,  $\forall w \in W$ , (6.95a) and (ii) there exists no other  $/x \in M$  (say,  $/x$ ), and a corresponding  $A = A < A$ , such that  $\|S/\gg\|^2 < A \|H\|^2$ ,  $\forall w \in H$ . (6.95b) Now, introducing the parameterized (in  $A > 0$ ) family of cost functions  $Lx\{v, W\} = \|G\gg\{w\}\|^2 - nH^2$ , (6-96) (i) and (ii) above become equivalent to the problem of finding the "smallest" value of  $A > 0$  under which the upper value of the associated game with objective function  $L\backslash(n, w)$  is zero, and finding the corresponding controller that achieves this upper value.

344 T. BAŞAR AND G. J. OLSDER We now relate this problem to the ones covered by Thms. 6.17, 6.19, 6.23 and 6.24, when the information pattern is CLPS. Toward this end let us first make the simplifying assumptions that  $H'G = 0$ ,  $G'G = I$ ,  $H'(T)QfG\{T\} = 0$  and  $G'(T)QfG(T) = 0$ . Then,  $Lx$  can be written as (6.82), with  $Qf = H(T)'QfH(T)$ ,  $Q = H'H$ . (6.97) In the finite horizon, we then have precisely the linear-quadratic game covered in Thm. 6.19, with  $B_1$  replaced by  $B$ ,  $B_2$  replaced by  $D$ , and  $x_0 = 0$ . We know from Thm. 6.19, together with Thm. 6.18, that the upper and lower values are equal whenever the upper value is bounded, and that this finite value is zero (since  $x_0 = 0$ ). Furthermore, where this breaks down (that is, when the upper value becomes unbounded) is at the largest value of  $A$  for which the generalized Riccati differential equation (6.83), that is,  $Z + A'Z + ZA + Q - z(bB' - jDD'JZ = 0$ ,  $Z\{T\} = Qf$ , (6.98) has a conjugate point. This value of  $A$  (which we had denoted  $A^*$ ) is precisely the  $A$  introduced earlier as the square of (6.94). Hence the

following result follows as a direct corollary to Thm. 6.19. Corollary 6.9 For the finite-horizon  $H^\infty$ -optimal control problem formulated above, and with CLPS information, let  $A^*$  be defined as in Thm. 6.19, in connection with the Riccati differential equation (6.98). Then, (i)  $\gamma = \gamma^*$ , (ii) for all  $A > A^*$ , the control policy  $Hx(t,x) = -B(t)'Zx(t)$ ,  $t \geq 0$  delivers a performance level of at least  $-\sqrt{A}$ , that is, for all  $w \in H$ ,  $\|II^MII/IIK\| \geq \sqrt{A}$ . For the infinite-horizon problem, the counterpart of Corollary 6.9 is the following corollary to Thm. 6.24. Corollary 6.10 For the infinite-horizon  $H^\infty$ -optimal control problem, and with CLPS information, let  $(A,B)$  be stabilizable and  $(A,Q)$  be observable (respectively, detectable). Let  $A^*$  be the smallest positive scalar with the property that for all  $\gamma > \gamma^*$ , the GARE  $AZ + ZA + Q - z(BB' - tDD')J^2 = 0$  (6.99) admits a minimal positive (respectively, nonnegative) definite solution  $Z$ . Then,

NASH AND SADDLE-POINT EQUILIBRIA 345 (ii) for all  $A > A^*$ , the feedback control policy  $jix(x) = -B'Zxx$ ,  $t > 0$  delivers a performance level of at least  $\sqrt{A}$ , and under it the linear system  $\dot{x} = (A - BB'Z)x + Dw$  is bounded-input bounded-state stable. The following scalar example (taken from Başar and Bernhard (1995)) illustrates the two corollaries above. Example 6.7 Consider the scalar plant with dynamics  $\dot{X} = U + W$ , and  $JET^\infty$ -performance index:  $r \int_0^T [x(t)]^2 + [u(t)]^2 dt + \gamma \int_0^T [w(t)]^2 dt$ . At The associated GRDE is  $Z + 1 - \gamma^{-1} = 0$ ,  $Z(T) = 0$ , (0  $\gg$ ) whose unique solution takes two different forms, depending on whether  $A > 1$  or not. For the former case,  $A > 1$ :  $Z(A(0) = \tanh[\gamma^{-1}(T-0)]$ ,  $\gamma^* = \sqrt{(A-1)/A}$ , and for the latter,  $A < 1$ :  $Zx(t) = \pm \tan[m(T-t)]$ ,  $m = \gamma/(1-A)/A$ , which exists provided that Note that the given solution family is continuous at  $A = 1$ . Now, for fixed  $T$ , the GRDE has a conjugate point in the interval  $[0, T]$  if  $A < A^*(T)$ , and for all  $A < A^*(T)$  the upper value of the related game is unbounded. Hence, the optimum  $JET^\infty$ -performance level (that is, the upper value of the game with performance index (i)) is  $\gamma^*\{T\}$ . For  $A > A^*(T)$ , the controller that guarantees this level of performance is  $H^*x(t,x) = -Zx(t)$ ,  $0 < t < T$ . (Hi)

346 T. BAŞAR AND G. J. OLSDER Note that as  $A \downarrow A^*(T)$ ,  $Z(0) \rightarrow \infty$ , and hence the "optimum" controller is a high gain controller, which may not be desirable. For any  $A > A^*(T)$ , however, the control gain is finite, and one can achieve  $\epsilon$ -optimal performance (that is, one that is within an  $\epsilon$ -neighborhood of  $A^*(T)$ ) using finite gain controllers of the type (iii). For the infinite-horizon case, we simply let  $T \rightarrow \infty$ , leading to  $A^* = 1$ , and the unique stationary

feedback controller -  $t^{\wedge}$  for  $A > 1$ . Again note that the solution ceases to exist as  $A \downarrow A^* = 1$ , and hence an "optimal" controller does not exist, but an  $\epsilon$ -optimal controller does. We now consider the imperfect state measurements case, as originally formulated. For technical reasons we will treat the initial state  $x_0$  as also unknown and as part of the disturbance, and attach weight to it in the cost function. Furthermore, we take the system and measurement disturbances to be independent (which is tantamount to taking  $ED' = 0$ ), and take without any loss of generality  $EE' = I$  and let  $Ew = v$ , which is viewed as a separate disturbance (in addition to  $w$  and  $x_0$ ). Then, the associated parameterized game is one with state dynamics  $\dot{x} = Ax + Bu + Dw$ ,  $x(0) = x_0$ , measurement equation  $y(t) = Cx(t) + v(t)$ ,  $t > 0$ , and objective function (for finite horizon)  $J(x; u, w, v, x_0) = \int_0^T (\|x(t)\|^2 Q + \|u(t)\|^2) dt + \rho \|x_0\|^2 + \int_0^T (\|w(t)\|^2 + \|v(t)\|^2) dt$  (6.100a) (6.100b) (6.100c) For the infinite horizon, we simply let  $T \rightarrow \infty$ ,  $Q \rightarrow 0$ . In view of our earlier discussion in this section, we seek the smallest value of  $A$ , say  $A^*$ , such that for all  $A > A^*$ , and with  $u = \{w, v, x_0\}$ ,  $\inf \sup J(x; u) = 0$ , (6.101) where  $M$  is defined here as the class of all (measurable, smooth) control policies that depend causally on  $y$ . The theory we have developed in this chapter does not directly apply to this problem, since the information structure is neither CLPS nor OL. Similar game theoretic techniques can be used, however, to obtain a complete, very

NASH AND SADDLE-POINT EQUILIBRIA 347 appealing solution to the problem. The result is given below in Thm. 6.25 for the finite horizon, and Thm. 6.26 for the infinite-horizon case. Verifications of these results will be sketchy, since the details are lengthy and fairly technical. A complete treatment can be found in Başar and Bernhard (1995). Before presenting Thm. 6.25, we introduce a second GARE, dual to (6.98), again defined on  $[0, T]$ :  $-\dot{Z} + AZ + EA' + \epsilon > \epsilon' - \epsilon (C'Z - jQ)S = 0$ ;  $Z(0) = Q, *$ . (6.102) We further define  $G(t) = Z[I - A^{-1}EZ]^{-1}$  (6.103) whenever it exists, and introduce two different equations  $\dot{x} = [A - \{BB' - \lambda x DD'\}Z]x + [I - \lambda^{-1}Z]^{-1}Y, C(y - Cx)$ ;  $x(0) = 0$  (6.104a) and  $\dot{z} = (j4 - ECC' + A^{-1}EQ)z - Bu - ECV$ ,  $z(0) = 0$ . (6.104b) Theorem 6.25 Let  $A$  be the set of all  $X > 0$  such that (a) the GRDE (6.98) does not have a conjugate point on  $[0, T]$ , (b) the GRDE (6.102) does not have a conjugate point on  $[0, T]$ , and (c) the matrix  $XI - H(t)Z(t)$  has only positive eigenvalues for all  $t \in [0, T]$ , or equivalently  $\rho(H(t)Z(t)) < A$ ,  $t \in [0, T]$ , (6.105) where  $\rho(-)$  is the spectral radius. For each  $X \in A$ , let a controller  $Hx \in M^e$  defined by  $Hx(t, y) = -B(t)'Zx(t) + x(t)$ ,  $t > 0$ , (6.106a) where  $Z$  is the solution of (6.98), and  $x$  is

generated by (6.102). Then, (i)  $\inf\{A > 0 : A \in A\} = A^*$ , where  $X^*$  is the smallest positive scalar  $X$  such that (6.101) holds; (ii) for each  $X > A^*$ , the controller given by (6.106a) delivers a performance level of at least  $\bullet/X$ ; (iii) for each  $X > A^*$ , the controller (6.106a) can equivalently be written as

$$348 \text{ T. BAŞAR AND G. J. OLSDER } \dot{x}(t,y) = -B(t)'Gx(t) - x(t), t > 0, \quad (6.106b)$$

where  $G$  is given by (6.103) and  $x$  is generated by (6.104b). Proof. We will sketch here a proof of sufficiency, using completion of squares. First note that since  $Q^{-1} > 0$ , the solution of (6.102) is positive definite, thus admitting an inverse. Define the function  $V(t) = x(t)'G\{t\} + \|x(t) - x(t)\|^{-1}\{t\}$  and add the identically zero function  $V(0) - V(T) + \int_0^T V(t) dt$  to  $L$  given by (6.100c). The derivative  $\dot{V}$  involves the derivatives of  $G$ ,  $E^{-1}$  and  $e = x - \hat{x}$ , which can be shown to be given by (through elementary but extensive manipulations)  $\dot{G} + G\{A + A^{-1}EQ\} + (A + A^{-1}EQ)'G + Q - G(BB' - A^{-1}EG'GE)G = 0$ ,  $\dot{E}^{-1} + E^{-1}M + j^4E^{-1} + E^{-1}U'E^{-1} - C'C + A^{-1}Q = 0$ ;  $E^{-1}(0) = Q0$ ,  $e = (A - EC'C)e - EC'w + Dw - A^{-1}EQx$ ;  $e(0) = x0$ . If all these relationships are used in  $L$ , it simplifies to  $Lx(u; w, v, x0) = -e(T)' + [Qf - X^{-1}(T)]^{-1}Qfx(T)\|1. \_1\{T\} \_Qf - 2\|x(T)\|2Q/(A \_E(T)Q) \_1E(T)Q/ + \|U + B'Gxf(o) - X\|v + C\{e - A \^Gx\}\|2 - A\|tu - U'E \^H2$ , where  $\|\bullet\|$  denotes the appropriate Hilbert space norm. Now, invoking condition (c) of the theorem we conclude that  $AE^{-1}T) - Qf > 0$ ,  $Qf(XI - XWQf)^{-1}WQf > 0$ , and hence the first two terms in (o) are nonpositive. Furthermore, since the control given by (6.106b) makes the third (nonnegative) term in (o) zero, we have the bound  $Lx(x; w, v, x0) < 0$ ,  $\forall w, v, x0$ . Its maximum is actually equal to zero, which can be seen by picking in (o),  $w = D^{-1}\xi^{-1}e$ ,  $w = -C(e - A^{-1}EGx)$ ,  $x0 = 0$ . This completes the proof of (ii). To prove (iii), we start with the transformation  $x(t) = \{I - Z(t)Z(t)\}^{-1}x(t)$  and differentiate the RHS with respect to  $t$ . This leads (after some algebra) to the conclusion that  $x$  satisfies (6.104a) whenever  $x$  satisfies (6.104b). D

NASH AND SADDLE-POINT EQUILIBRIA 349 Remark 6.22 The proof given above for sufficiency (that is, parts (ii) and (iii)) is an indirect one, and follows the arguments of Uchida and Fujita (1989). Başar and Bernhard (1995) provide a more constructive proof, for all three parts of the theorem. Other proofs can be found in Limebeer et al. (1989), Khar-gonekar (1991) and Khargonekar, Nagpal and Poolla (1991) to cite just a few. An important observation to be made in the context of Thm. 6.25 is the duality between the two GRDEs (6.98) and (6.102). This duality enables us to deduce the behavior

of the minimax controller  $n$  given by (6.106a) for the time-invariant problem, as  $T \rightarrow \infty$ , by making direct use of the results of subsection 6.5.3, as well as that of Corollary 6.10. This leads to the following counterpart of Corollary 6.10. Theorem 6.26 For the infinite-horizon  $H^\infty$ -optimal control problem with imperfect state measurements, as defined above, let  $(A, B)$  be stabilizable,  $(A, Q)$  be detectable,  $(A, C)$  be detectable, and  $(A, D)$  be stabilizable. Let  $A^\infty$  be the set of all  $A > 0$  such that (6.99) admits a minimal nonnegative definite solution (denoted  $Z_f$ ), the following dual GARE admits a minimal nonnegative definite solution (denoted  $S^A: A\dot{x} + (A' - D'D^{-1}D - T)(c'c - j\dot{q})x = 0$ , (6.107) and the following spectral radius condition is satisfied:  $\rho\{T + xZ^+\} < A$ . (6.108) Then, (i)  $\inf\{A > 0 : A \in A^\infty\} = A^\wedge$ , where  $A^\wedge$  is the smallest positive scalar  $A$  such that the infinite-horizon version of (6.101) holds; (ii) for each  $A > A^\wedge$ , the controller  $px(y) = -B'Z^+(I - A^{-1}E + Z^+)^{-1}x$ , (6.109a) where  $x = (A + A^{-1}E + Q)x + Bu + Z^+C'(y - Cx)$ ;  $x(0) = 0$ , (6.109b) delivers a performance level of at least  $\forall A$ . We close this section by revisiting Example 6.7 for the imperfect state measurements case. Example 6.8 Consider Example 6.7 now with a measurement equation  $y = x + v$ ,

350 T. BAŞAR AND G. J. OLSDER and the  $J_{ET}^\infty$ -performance index compatibly replaced by  $m^2 + h^2)^{1/2}/(\|W\|^2 + \|v\|^2)^{1/2}$ , where we have abused the notation and have taken  $Dw = w$ . Note that since  $x_0 = 0$ , we have taken in (6.100c)  $Qq_1 = 0$ . In view of this, the second GRDE (6.102) reads  $t = Q - 1) S^2 + 1, S(0) = 0$ , which is the dual of (ii) in Example 6.7 (simply the time is reversed), and hence requires the same conjugate-point condition ( $A < AT^2/(AT^2 + t^2)$ ) as the  $Z$  equation. Using the notation of Example 6.7, for  $A > A^*(T)$  we have  $1 - \tan^2 m \leq \tan[m(T - t)] < A^{-4} \Rightarrow -\tan m \leq m$  which is more restrictive than the other two conditions (which are identical):  $A < A^*(T)$ . Hence, in this case the spectral radius condition is binding. For the infinite-horizon case, again from duality existence of  $Z^+$  and  $S^+$  requires the same condition, which is  $A > 1$ . The spectral radius condition is  $Z^+E^+ = -r^+ < A$  &  $A > 2. A - 1$  Hence  $A^\wedge = 2$ , and again the spectral radius condition is binding.  $\square$

### 6.7 Stochastic Differential Games with Deterministic Information Patterns

As it has been discussed in subsection 6.5.2, one possible way of eliminating "informational nonuniqueness" of Nash equilibria in nonzero-sum differential games is to require the equilibrium solution also to satisfy a feedback property (cf. Def. 6.2), which necessarily leads to strongly time

consistent solutions; but this is a valid approach only if the information structure of the problem permits the players to have access to the current value of the state. Yet another approach to eliminate informational nonuniqueness is to formulate the problem in a stochastic framework, and this has been done in Section 6.4 for dynamic games defined in discrete time. A stochastic formulation for quantitative differential games (i.e., dynamic games defined in continuous time) of prescribed duration involves (cf. Section 5.3) a stochastic differential equation 
$$dx_t = \sum_{i=1}^N u_i(t, x_t) dW_t + a(t, x_t) dt, \quad x_t|_{t=0} = x_0, \quad (6.110) \quad 0 \leq t \leq T$$

NASH AND SADDLE-POINT EQUILIBRIA 351 which describes the evolution of the state (and replaces (6.37)), and  $N$  cost functionals 
$$J_i(u^1, \dots, u^N) = \int_0^T f_i(t, x_t, u^1(t), \dots, u^N(t)) dt + q_i(x_T); i \in N. \quad (6.111)$$
 If  $T^*(i \in N)$  denotes the strategy-space of  $P_i$  (compatible with the deterministic information pattern of the game), his expected cost functional, when the differential game is brought to normal form, is given by 
$$J_i(1, \dots, n) = E\{L_i(u^1, \dots, u^N) | \mathcal{F}_t^i, J \in N\} \quad (6.112)$$
 with  $J \in \mathcal{F}_t^i$  ( $i \in N$ ) and  $E[\cdot]$  denoting the expectation operation taken with respect to the statistics of the standard Wiener process  $\{W_t, t \geq 0\}$ . Furthermore,  $F$  and  $a$  are assumed to satisfy the requirements of Thm. 5.2, and in particular  $a$  is assumed to be nonsingular. This section presents counterparts of some of the results of Section 6.4 for stochastic nonzero-sum differential games of the type described above; the analysis, however, will be rather informal since the underlying mathematics (on stochastic differential equations, stochastic calculus, etc.) is beyond the scope of this volume. Appropriate references will be given throughout for the interested reader who requires a more rigorous exposition of our treatment here. First, by way of introduction, let us consider the one-person version of the AT-person stochastic differential game, which is a stochastic optimal control problem of the type described in Chapter 5, for which a sufficiency condition for existence of an optimal control has been given in Prop. 5.7. This condition involves the existence of a solution to a particular partial differential equation (specifically, (5.71)), which corresponds to the Hamilton-Jacobi-Bellman equation (5.25) of dynamic programming, but now we have the extra second term which is due to the stochastic nature of the problem. We should note that if the control was allowed to depend also on the past values of the state, the optimal control would still be a feedback strategy since the underlying system is Markovian (i.e.,  $F$  and  $a$  depend only on  $x_t$  and not on  $\{x_s, s < t\}$  at time  $t$ ). We now

observe an immediate application of Prop. 5.7 to the AT-person nonzero-sum stochastic differential game of this section. Theorem 6.27 For an N-person nonzero-sum stochastic differential game of prescribed fixed duration  $[0, T]$ , as described by (6.110)-(6.111), and under FB, MPS or CLPS information pattern, an N-tuple of feedback strategies  $\{7^{**} e r^{*}; i \in N\}$  provides a Nash equilibrium solution if there exist suitably smooth functions  $W^i : [0, T] \times R^n \rightarrow R, i \in N$ , satisfying the coupled semilinear parabolic partial differential equations  $dt \sum_{j=1}^n dx^j dX^j = \min [ \sum_{i=1}^N W_i - F_i^*(\langle x, \langle i \rangle) + s \langle^*(\langle x, \langle i \rangle) ]$  (6.113) =  $VXW \rangle \blacksquare F^>^*(t, x, Y^*(t, x)) + g^>^*(t, x); W^i(T, x) = q_i(x) (i \in N)$ ,

352 T. BAŞAR AND G. J. OLSDER where  $o_{ki}$  is the  $k$ th element of the symmetric matrix  $ad'$ ,  $F^{TM} \{t, x, u\} = t F \{t, x, u\} + \sum_{i=1}^n r^i \cdot x \cdot t_i^* = g^i(t, x, u)$ ,  $\{U^i\} = \{U_1^i, \dots, U_{i-1}^i, U_i, U_{i+1}^i, \dots, U_N^i\}$ . Proof. This result follows readily from the definition of Nash equilibrium and from Prop. 5.7, since by fixing all players' strategies, except the  $j$ 'th ones, at their equilibrium choices (which are known to be feedback by hypothesis), we arrive at a stochastic optimal control problem of the type covered by Prop. 5.7 and whose optimal solution (if it exists) is a feedback strategy. D Theorem 6.27 actually involves two sets of conditions: (i) existence of minima to the RHS of (6.113), and (ii) existence of "sufficiently smooth" solutions to the set of partial differential equations (6.113). A sufficient condition for the former is the existence of functions  $u^i : [0, T] \times R^n \times R^n \rightarrow S^1$  which satisfy the inequalities  $W^i(t, x, p_i; u_1, \dots, u_i, \dots, u^i) < H^i(t, x, t; u_1, \dots, u_{i-1}, u_i, u_{i+1}, \dots, u_N)$ ,  $\forall t \in [0, T], \forall p_i \in R^n, \forall u^i \in R^n$  (6.114a)  $u_i = 0, u_i = u_i + 10, \dots, u^i = 0$ ,  $\forall t \in [0, T], \forall p_i \in R^n, \forall u^i \in R^n$  (6.114b) If such functions  $u^i(t, x, p_i)$  ( $i \in N$ ) can be found, then we say that the Nash condition holds for the associated differential game.66 In fact, Uchida (1978) has shown, by utilizing some of the results of Davis and Varaiya (1973) and Davis (1973) obtained for stochastic optimal control problems, that under the Nash condition and certain technical restrictions on  $F, a, S, g^i$  and  $q_i$  ( $i \in N$ ) there indeed exists a set of sufficiently smooth solutions to the partial differential equations (6.113). Therefore, the Nash condition is sufficient for a suitably well-defined class of stochastic differential games to admit a Nash equilibrium solution;67 and a set of conditions sufficient for the Nash condition to hold are given in Uchida (1979). For the special case of zero-sum stochastic differential games, i.e., when  $l = -I = L$  and  $N = 2$ , every solution set of (6.113) is given by  $W_1 = -W_2 = W$ , 66Note that such a condition could also have a counterpart in Thm. 6.16, for the deterministic



problem. Uchida (1978) actually proves this result for a more general class of problems for which the state equation is not necessarily Markovian, in which case a more general version replaces (6.113).

NASH AND SADDLE-POINT EQUILIBRIA 353 where  $W$  satisfies

$$-\dot{y}^i = \min_{u_1} \max_{u_2} [V_x W + F(t, x, v, u_1, u_2)] = \max_{u_1} \min_{u_2} [V_x W + F(t, x, u_1, u_2) + g(t, x, u_1, u_2)]; u_2 \in S_2 \quad u_1 \in S_1 \quad W(T, x) = q(x). \quad (6.115)$$

Furthermore, the Nash condition (6.114a) can be re-expressed for the stochastic zero-sum differential game as existence of functions  $u_i^\circ: [0, T] \times R^n \times R^m \rightarrow S_i, i = 1, 2$ , which satisfy  $H(t, x, p; u_1^\circ, u_2^\circ) = \min_{u_1} \max_{u_2} H(t, x, p; u_1, u_2) = \max_{u_1} \min_{u_2} H(t, x, p; u_1, u_2)$ ,  $u_2 \in S_2 \quad u_1 \in S_1$  (6.116a) where if  $(\cdot, x, p; u_1, u_2) = p'F(t, x, p; u_1, u_2) + g(t, x, u_1, u_2)$ . (6.116b) Because of the similarity with the deterministic problem, condition (6.116a) is known as the Isaacs condition. Now, it can actually be shown (see Elliott, 1976) that, if the Isaacs condition holds and if  $S_1$  and  $S_2$  are compact, there exists a suitably smooth solution to the semilinear partial differential equation (6.115) and consequently a saddle-point solution to the stochastic zero-sum differential game under consideration. Therefore, we have the following.

68 Corollary 6.11 For the two-person zero-sum stochastic differential game described as an appropriate special case of (6.110)-(6.111), and satisfying certain technical restrictions, let  $S_1$  and  $S_2$  be compact and let the Isaacs condition (6.116a) hold. Then, there exists a saddle-point solution in feedback strategies, which is  $u_1^*(t, x) = u_1^\circ(t, x, V_x W(t, x))$   $u_2^*(t, x) = u_2^\circ(t, x, V_x W(t, x))$  where  $W(t, x)$  satisfies (6.115) and  $u_i^\circ(t, x, p)$  ( $i \in N$ ) are determined from (6.116a). This solution is strongly time consistent. For the special case of affine-quadratic stochastic differential games, both Thm. 6.27 and Corollary 6.11 permit explicit expressions for the equilibrium strategies. Toward this end, let  $F, g_1$  and  $g_2$  ( $i \in N$ ) be as given in Def. 6.5 (with  $F$  replacing  $f$ ), and let  $d(t, x) = d(t)$  which is nonsingular. Furthermore, assume that  $S_i = R^m$  ( $i \in N$ ). 69 Then,  $u_i^\circ$  from (6.114a) is uniquely given by  $u_i^\circ = -[d^{-1}(t, x) B_i' \text{ityp}]^{-1} (i \in N)$ ,  $;$  68 Elliott (1976) has proven this result for a non-Markovian problem in which case (6.115) is replaced by a more complex relation; here we provide a version of his result which is applicable to Markovian systems. 69 In order to apply Corollary 6.11 directly, we in fact have to restrict ourselves to an appropriate (sufficiently large) compact subset of  $R^m$ .

354 T. BAŞAR AND G. J. OLSDER in view of which (6.113) can be written

as  $dW^* = -V^{-1} [A^T W^* + Q - B^T R^{-1} B^T W^* + C] dt + \sum_{k,j} U_{k,j} dX_j$ ,  $W(T, x) = V^{-1} [Z(T)x + C(T) + m(T)] + C(T)$  ( $i \in N$ ), where  $Z \in \mathbb{R}^{n \times n}$ , satisfy the AT coupled matrix Riccati differential equations (6.70a),  $\ell'$  and  $m'$  satisfy (6.72a) and (6.73b), respectively, and  $\ell'$  satisfies the differential equation  $C = -\text{Tr}[a(t)a^T(t)]$ . Now, since  $J^*$  is given by it readily follows that the Nash equilibrium solution of the stochastic affine-quadratic differential game coincides with that of the deterministic game, as provided in Corollary 6.5. Therefore, we have the following.

**Corollary 6.12** The set of Nash equilibrium strategies (6.71) for the deterministic affine-quadratic differential game with MPS or CLPS information pattern also provides a Nash equilibrium solution for its stochastic counterpart, with a not depending on the state  $x_t$ . A similar analysis and reasoning also leads to the following.

**Corollary 6.13** Under the conditions of Thm. 6.17, the set of saddle-point strategies (6.77) for the deterministic affine-quadratic two-person zero-sum differential game with MPS or CLPS information pattern also provides a saddle-point solution for its stochastic counterpart, provided that  $a$  does not depend on the state  $x_t$ .

**Remark 6.23** Under appropriate convexity restrictions, which hold for the affine-quadratic nonzero-sum game treated here, existence of a solution set to (6.113) is also sufficient for existence of a Nash equilibrium (see Varaiya, 1976);

**NASH AND SADDLE-POINT EQUILIBRIA 355** therefore the solution referred to in Corollary 6.12 is the unique Nash equilibrium solution for the affine-quadratic stochastic differential game within the class of feedback strategies. It should be noted that we do not rule out existence of some other Nash equilibrium solution that depends also on the past values of the state (though we strongly suspect that this should not be the case), since Thm. 6.27 provides a sufficiency condition only within the class of feedback strategies. The reader should note, however, that we cannot have information-ally nonunique equilibrium solutions here, since  $\{w_t, t > 0\}$  is an independent-increment process and  $a$  is nonsingular. For the zero-sum differential game, on the other hand, an analogous uniqueness feature may be attributed to the result of Corollary 6.13, by making use of a sufficiency result of Elliott (1976), which is obtained for compact  $S_1$  and  $S_2$  but can easily be generalized to noncompact action spaces for the affine-quadratic problem.  $\square$

**6.8 Problems 1.** An alternative derivation of the result of Thm. 6.2 makes use of Prop. 5.1. First

show that the  $i$ th inequality of (6.3) dictates an optimization problem to  $P_i$  which is similar to the one covered by Prop. 5.1, with  $c_k$  in Prop. 5.1 replaced by  $c_k + J_2^{eN} Blul^*$  which is only a function of  $X_i$  (and hence can be considered as a constant). Then the open-loop Nash equilibrium solution is given as the solution of the set of relations  $\langle f = -p \xi^{i+1} A_k x_t - p_i U^{i+1} + s_i^{i+1} c_k + s_i^{i+1} Y, B_i \langle f_c - [K_k + t S_k^k + l t S_k l - O f_c i \blacksquare S_i t = Q X + A_k s_k + 1 [i \sim \sim -^{f_c} - P f_c - S f_c + i l^{f_c} - S f_c + i = (2 f_c + i) > 4 = A_k l l - B l p_i S l + 1 \rangle [s l + 1 + S i + l c_k + S l + 1 Y, B \langle h \rangle \xi N s_k + i = 0, k \in K, i \in N$ , where  $x^{*k+1}$  ( $k \in K$ ) denotes the corresponding state trajectory. Now finally prove that the preceding set of relations admits a unique solution as given in Thm. 6.2, under precisely the same existence condition. 2. Consider the general formulation of an AT-person nonzero-sum dynamic game, as described by (6.1) and (6.2), where the players have access to the state with a delay of one time step, i.e., the information available to each player at stage  $k$  is  $\forall k = \{x_{f_c-i}, X_{f_c-2}, \dots, a^? i\}$ ,  $k > 2$ . We wish to obtain a strongly time consistent Nash equilibrium solution for this class of dynamic games.

356 T. BAŞAR AND G. J. OLSDER (i) Obtain the counterpart of Thm. 6.6 under this information structure. (ii) Specialize the result above to affine-quadratic dynamic games, so as to obtain the counterpart of Corollary 6.1 for this one-step delay CLPS pattern. 3. For the class of affine-quadratic zero-sum dynamic games covered by Thm. 6.7, modify the information structure so that now  $P_i$  is allowed to have access to the current and past actions of  $P_2$ , i.e.,  $P_i$ 's information structure is  $T_{j1} = \{X_k, U_k, X_{k-1}, U_{k-1}, \blacksquare \blacksquare \blacksquare, X_i, V\%$ ,  $k > 1$ . Develop the counterpart of the result of Thm. 6.7 for this class of games. In particular, show that (i) there exists a saddle point under conditions that are less stringent (for  $P_i$ ) than the conditions of Thm. 6.7; (ii) the saddle-point policies of the players that are also strongly time consistent are affine functions of the current information (which is  $\{r_{f_c j} U_j\}$  for  $P_i$ , and  $X_k$  for  $P_2$ , at stage  $k$ ); (iii) the corresponding state trajectory is the same as the one generated by the feedback saddle-point solution of Thm. 6.7, whenever the latter exists. 4. Consider the infinite-horizon two-player linear-quadratic dynamic game whose two-dimensional state  $x_k = (x^1, x^2)'$  is described by  $\hat{f}_c + i = 2 x_{f_c} + u_k + e x_k; x_1 = 1, x_{k+1} = -2 x_k + u_k - e x_k; x_1 = 1$ , where  $e > 0$  is a scalar parameter. Let the cost functions be given by  $\hat{\blacksquare} = \xi(4+i)2 + K)2 - i = i, 2. f_c = i$  (i) Find the largest value of  $e > 0$  (say,  $e^*$ ), such that for all  $e < e^*$ , the game admits a stationary feedback Nash equilibrium solution. (ii) Obtain the feedback Nash equilibrium

solution when  $e = e^*/2$ , by directly using (6.28a)-(6.28c), and also by iterating on the solution of the finite-horizon game. Derive a set of nonlinear Nash equilibrium solutions for the dynamic game of subsection 6.3.1. Specifically, obtain a counterpart of the result of Prop. 6.5 by starting with the representations  $J_1(x_2, a; i) = -x_2 + p[x_2 - X_2\{j_3\}]^2$ ,  $J_2(x_2, x_i) = -x_2 + q[x_2 - X_2\{j_3\}]^2$ ,

**NASH AND SADDLE-POINT EQUILIBRIA 357** and by mimicking the analysis that precedes Prop. 6.5, How does the underlying existence condition compare with (6.35)?

6. Consider the scalar two-person three-stage stochastic dynamic game described by the state equation  $X_4 = X_3 + u + \eta$ ,  $X_3 = X_2 + u + \eta$ ,  $X_2 = x + u + \eta$ , and cost functionals  $J_1 = (x_4)^2 + (u)^2 + (u^t)$ ,  $J_2 = (x_4)^2 + 2(u)^2 + (u)^2$ . Here  $\eta$  denotes a Gaussian random variable with mean zero and variance 1, and both players have access to CLPS information. Show that this dynamic game admits an uncountable number of Nash equilibria, and obtain one such set which is linear in the state information available to both players. Will there be additional Nash equilibria if  $\eta$  has instead a two-point distribution with mean zero and variance 1?

7. For the stochastic dynamic game of the previous problem, prove that the Nash equilibrium solution is unique if the underlying information pattern is, instead, one-step delay CLPS pattern. Obtain the corresponding solution, and identify it as a member of the solution set determined in Problem 6.

8. Consider the class of two-person  $k$ -stage linear-quadratic stochastic dynamic games with cost functional (6.10b) and state equation  $x_{k+i} = A_k x_k + B_k u + B_k u + \eta_k, k \in K$ , where  $\{\eta_k, k \in K\}$  denotes a set of independent Gaussian vectors with mean zero and covariance  $\Sigma_k$ . Obtain the saddle-point solution when (i) both players have access to one-step delayed CLPS information, (ii)  $P_1$  has access to CLPS information and  $P_2$  has access to one-step delayed CLPS information, and (iii)  $P_1$  has access to CLPS information whereas  $P_2$  has only access to the value of  $x$  (i.e., OL information). How do the existence conditions compare in the three cases, among themselves and with (6.26a)-(6.26b)? (Hint: Make use of Thm. 6.9 in order to simplify the derivation of the saddle-point strategies.)

9. Execute the procedure outlined in Remark 6.13 (for two-person games with weakly coupled players) when the state dynamics in (6.48a) are linear and the cost functions in (6.48b) are quadratic, and obtain expressions for the first three terms in the expansions for the open-loop Nash strategies of the

358 T. BAŞAR AND G. J. OLSDER 10. Show that Thm. 6.12 can also be obtained by directly utilizing Prop. 5.2 in a manner which is parallel to the one outlined in Problem 1 for the discrete-time counterpart. 11. This problem addresses a generalization of Problem 1 in Section 5.8. Assume that we now have two companies,  $P_1$  and  $P_2$ , with  $P_i$  having  $x_i(t)$  customers at time  $t$  and making a total profit of  $\int_0^t [c^* x_i(s) - u_i(s)] ds$  up to time  $t$ , where  $c^*$  is a constant. The function  $u_i(t)$ , restricted by  $u_i(t) > 0$ ,  $f_i(t) > 0$ , represents the money put into advertisement by  $P_i$ , which helps to increase the number of customers attracted to that firm, according to  $\dot{x}_i = u_i - u_j$ ,  $x_i(0) = x_{i0} > 0$ ,  $i, j = 1, 2, j \neq i$ . The total number of customers shared between the two companies is a constant, so that we may write, without any loss of generality, the constraint relation  $x_1(0) + x_2(0) = 1$ .  $\forall t \in [0, T]$ , where  $T$  denotes the end of the time period when each firm wishes to maximize its total profit (i.e.,  $f_i(T)$ ). Obtain the open-loop Nash equilibrium solution of this two-person nonzero-sum differential game wherein the decision variable for  $P_i$  is  $u_i$ . (For details see Olsder, 1976). 12. Consider a variation on the previous problem so that now the state equation is described by  $\dot{x} = X_2 U_1 - X_1 U_2$ ,  $\dot{x}_2 = X_1 U_2 - X_2 U_1$ , and the decision variables satisfy the constraints  $u_i(t) \leq 1$ ,  $t = 1, 2$ . The objective for  $P_i$  is still maximization of  $f_i(T)$ . (i) Obtain the open-loop Nash equilibrium solution. (ii) Obtain the feedback Nash equilibrium solution when both players have access to the current values of  $x_1$  and  $x_2$ .

NASH AND SADDLE-POINT EQUILIBRIA 359 13. Obtain the complete set of informationally nonunique linear Nash equilibrium solutions of the scalar linear-quadratic two-person nonzero-sum differential game described by the state dynamics  $\dot{x} = 2u_1 - u_2$ ,  $x(0) = 1$ ,  $t \in [0, 2]$  and cost functionals  $J_i = \int_0^2 [x^2 + f_i'(t)]^2 dt$ ,  $i = 1, 2$ ,  $J_0$  and with the underlying information structure being MPS for both players. Which of these would still constitute a Nash equilibrium when (i)  $P_2$  has, instead, access to OL information, (ii) both players have access to OL information? (Note that, in this case, OL information is in fact "no information" since the initial state is a known number.) 14. Prove that the statement of Thm. 6.9 is also valid for two-person zero-sum differential games with prescribed fixed duration. 15. Repeat Problem 9 for the CLNM information structure and using the feedback Nash equilibrium solution. Here Remark 6.18 replaces Remark 6.13, and Corollary 6.5 replaces Thm. 6.12. 16. Generalize the result of Thm. 6.17 to the case when the cost function has, under the integral, the additional term  $u_i(t)' K(t) u_i(t)$ , where  $K(-)$  is a matrix of

appropriate dimensions and with continuous entries. Is it possible for the differential game to have a bounded lower value even if the conditions for existence of a saddle point fail? 17. This problem uses the notation and terminology introduced in Section 6.6. We wish to design optimal controllers for the following two-dimensional system, under different information patterns, using the  $H^\infty$  criterion:  $\dot{x}_1 = x_2 + w$ ;  $\dot{x}_2 = -x_1 + x_2 + u$ ;  $z = x_1 + x_2$ ;  $y = x_1 + v$ . Here,  $x_1$  and  $x_2$  denote the two states, whose initial values are completely unknown,  $u$  is the scalar control variable,  $y$  is the measured output,  $z$  is the regulated output,  $w$  is the scalar system disturbance and  $v$  is the measurement disturbance. With this system, we associate the performance index  $J = \int_0^T [z(t)^2 + u(t)^2] dt + \frac{1}{2} [x_1(0)^2 + x_2(0)^2] + \int_0^T [w(t)^2 + v(t)^2] dt$  where  $T$  is the terminal time.

360 T. BAŞAR AND G. J. OLSDER (i) Let  $T = \infty$ , and the controller have access to perfect state information. Denote the optimum (minimax) performance level for this problem by  $A^*$ . Obtain the value of  $A^*$ , and a controller that will ensure a performance level no worse than  $A = A^* + 0.01$ . (ii) Repeat the above now under the original imperfect state information. (iii) Now take the time horizon to be finite, and  $T = 1$ . Furthermore take the initial state to be known to be zero. Obtain (approximately) the  $A^*$  for this problem under perfect state measurements by numerically solving the corresponding generalized Riccati differential equation for different values of  $A$ . 18. The following two-dimensional system description depends on a parameter  $\epsilon$ :  $\dot{x}_1 = x_1 + 2x_2 + 2u + w$ ;  $\dot{x}_2 = -x_1 + x_2 + \epsilon u + 3w$ . Given the cost function:  $J = \int_0^T [2x_2^2 + 2x_1x_2 + x_1^2 + u^2] dt + \frac{1}{2} [x_1^2 + x_2^2]$ , obtain the  $J$ -optimal performance level ( $A^*(\epsilon)$ ), under perfect state measurements, for  $\epsilon = 1, 0.1, 0.01, 0.001, 0.0001$ , to the nearest four decimal places. Solution: 2.4332, 2.7843, 2.9724, 2.9972, 2.9997. 19. Repeat Problem 18 for the following system and cost function:  $\dot{x}_1 = x_1 + x_2 + 3u + 2w$ ;  $\dot{x}_2 = -x_1 + x_2 + 2u + w$ . We now have a measurement equation  $y = 2x_1 + x_2 + 3v$ , where  $v$  is the measurement noise. Given the cost function,  $J = \int_0^T [2x_1^2 + 2x_1x_2 + 3x_2^2 + u^2] dt + \frac{1}{2} [x_1^2 + x_2^2]$  obtain the  $J$ -optimal performance level ( $A^*(\epsilon)$ ), under imperfect state measurements, for  $\epsilon = 1, 0.1, 0.01, 0.001, 0.0001$ , to the nearest four decimal places, and determine a controller (for each case) that assures a

performance no worse than  $A^*(e) + 0.01$ .

NASH AND SADDLE-POINT EQUILIBRIA 361 Partial Solution:  $A^* = 55.1993, 17.3221, 9.9331, 9.1197, 9.0308$ . 21. Consider the differential game of Problem 13, but with the state equation replaced by  $\dot{x} = 2u_1 - u_2 + a$ , where it is a random variable taking the values  $+1$  and  $-1$  with equal probability  $1/2$ . Under the CLPS information pattern for both players, does this differential game, which also incorporates a chance move, admit a unique Nash equilibrium solution? In case the reply is "yes", obtain the unique solution; in case it is "no", obtain the complete set of Nash equilibria that are linear in the current value of the state. 22. Obtain the saddle-point solution of the class of two-person zero-sum stochastic differential games described by the state equation  $d\mathbf{x} = [\mathbf{A}(t)\mathbf{x} + \mathbf{B}^1 u_1 + \mathbf{B}^2 u_2]dt + a(t)d\mathbf{w}$  and cost functional (6.60b), where  $P_1$  has access to CLPS information and  $P_2$  has OL information. Here,  $\mathbf{A}(t)$  is a nonsingular matrix,  $\{\mathbf{w}, t > 0\}$  is the standard Wiener process, and the initial state vector  $\mathbf{x}_0$  is a known quantity. (For details see Başar (1977b)).

6.9 Notes Section 6.2. Results pertaining to Nash equilibria in infinite dynamic games first appeared in the works of Case (1967, 1969) and Starr and Ho (1969a, b) but in the continuous time (i.e., for differential games). Here we present counterparts of these results in discrete time, and with special attention paid to affine-quadratic (nonzero-sum and zero-sum) games, and existence conditions for equilibria under open-loop and memoryless perfect state information patterns. A possible extension of these results would be to the class of game problems in which the players also have the option of whether they should make a perfect state measurement or not, if it is costly to do so; for a discussion of such problems see Olsder (1977b). For applications of some of these results in macroeconomics, see Pindyck (1977), Kydland (1975) and Başar, Turnovsky and d'Orey (1986). The last reference also studies, within the context of a two-country macroeconomic model, the computation and stabilizability of stationary feedback Nash equilibria using the solutions of a sequence of time-truncated dynamic games.

Section 6.3. Existence of uncountably many informationally nonunique Nash equilibria in dynamic games under closed-loop information was first displayed in Başar (1974), Başar (1976a) for discrete-time games, and in Başar (1977b) for differential games. The underlying idea in this section therefore follows these references. The algorithm of subsection 6.3.2 is from Başar (1977d), and Thm. 6.9 was first presented in Başar (1977a); some further discussion on the relation between

existence of a saddle point and the underlying information pattern in zero-sum dynamic games can be found in Başar (1976b)

362 T. BAŞAR AND G. J. OLSDER and Witsenhausen (1971b), where the latter reference also discusses the impact of information on the upper and lower values of a game when a saddle point does not exist. Section 6.4. Elimination of informational nonuniqueness in Nash equilibria through a stochastic formulation was first discussed in Başar (1976a), and further elaborated on in Başar(1975, 1979b, 1989b); this section follows basically the lines of those references. The fact that an increase in information could be detrimental for a player under Nash equilibria was first pointed out in Başar and Ho (1974), but within the context of stochastic games with noisy (imperfect) observation. For some results on the equilibrium solution of linear-quadratic stochastic games with noisy observation, see Başar(1978), Başar and Mintz (1972, 1973). Section 6.5. Derivation of open-loop and feedback Nash equilibria in nonzero-sum differential games was first discussed in Case (1967, 1969) and Starr and Ho (1969a, b), where the latter two references also display the differences (in value) between these two types of equilibria. Derivation of open-loop saddle-point equilibria, however, predates this development, and was first discussed in Berkovitz (1964) and Ho et al. (1965), with the latter devoted to a particular class of linear-quadratic differential games of the pursuit-evasion type, whose results were later put into rigorous framework in Schmitendorf (1970); see also Halanay (1968) which extends the results of Ho et al. (1965) to differential games with delays in the state equation. For an excellent exposition on the status of nonzero-sum and zero-sum differential game theory in the early 1970s, see the survey paper by Ho (1970). Some selective references which display the advances on Nash and saddle-point equilibria of deterministic differential games since the early 1970s are (in addition to those already cited in the text) Bensoussan (1974), Ho, Luh and Olsder (1980), Leitmann (1974, 1977), Scalzo (1974), Vidyasagar (1976) and related papers in the edited volumes Blaquiere (1973), Hagedorn et al. (1977), Ho and Mitter (1976), Kuhn and Szego (1971), Başar(1986b), Başar and Bernhard (1989), Hamalainen and Ehtamo (1991), Başar and Haurie (1994). For applications of this theory to problems in economics, see Case (1971, 1979), Clemhout et al. (1973), Pohjola (1986), Jorgensen (1986), Dockner and Feichtinger (1986) and Leitmann and Wan, Jr. (1979), where the last one discusses a worst-case approach toward



stabilization of uncertain systems, that utilizes the zero-sum differential game theory. Tolwinski (1978b) provides a rather complete account of the results on existence of open-loop Nash equilibria in nonzero-sum differential games, and Tolwinski (1978a) discusses a possible computational scheme for the numerical evaluation of the Nash equilibrium solution; see also Varaiya (1967, 1970), Rosen (1965) and Friedman (1971) for existence results concerning noncooperative equilibria. For the special class of linear-quadratic games, Papavassilopoulos and Cruz, Jr. (1979b) and Papavassilopoulos et al. (1979) discuss existence of a unique solution to the coupled set of Riccati equations (6.70a), which plays an important role in the characterization of feedback Nash equilibria; for computational algorithms for some specially structured games, see Başar (1991a), Srikant and Başar (1991), Gajic, Petkovski and Shen (1990). Papavassilopoulos and Olsder (1984) demonstrate that an infinite-horizon linear-quadratic differential game might have multiple Nash equilibria even though every time-truncated version of it has a unique feedback Nash equilibrium. Eisele (1982) was one of the first to point out that linear-quadratic differential games might admit multiple open-loop Nash equilibria, whose results were later extended by Engwerda (1998) who showed that an

NASH AND SADDLE-POINT EQUILIBRIA 363 open-loop Nash equilibrium solution may exist for a linear quadratic game even if the corresponding differential Riccati equations may not have a solution. The latter reference also studied the limiting behavior of finite time horizon open-loop solutions of linear-quadratic games as the final time approaches infinite. For a discussion of non-cooperative equilibria in differential games with state dynamics described by partial differential equations, see Roxin (1977). Section 6.6. The theory briefly presented here has extensions to other types of information structures, such as sampled-data (both perfect state and imperfect state), and delayed measurements, as well as to nonlinear systems; see Başar and Bernhard (1995). It is also possible to study the robustness (or sensitivity) of the  $H^\infty$ -optimal controllers to unmodeled fast plant dynamics, within the framework of singular perturbations; see Pan and Başar (1993, 1994a,b). These papers also develop, as a byproduct, a complete theory for the saddle-point equilibria of singularly perturbed linear-quadratic differential games under perfect and imperfect state measurements. It is also possible to combine Nash and saddle-point equilibria to develop a counterpart of  $H^\infty$ -optimal control theory for multiple controller systems under multiple worst-case design criteria (Başar

1992). Section 6.7. For some explicit results in the case of stochastic quadratic differential games with mixed information patterns, see Başar (1977b,c, 1979a, 1980a). Extensions of these results to stochastic differential games with noisy observation meet with formidable difficulties, some of which have been resolved in the case of linear-quadratic zero-sum games (see Willman (1969), Bagchi and Olsder (1981) and Başar (1981c), where the latter reference deals with the case of identical noisy measurements for the players). For a distributed algorithm for the computation of Nash equilibria in linear-quadratic stochastic differential games, see Başar and Li (1989).

This page intentionally left blank

## Chapter 7 Stackelberg Equilibria of Infinite Dynamic Games 7.1 Introduction

This chapter is devoted to derivation of the Stackelberg solution in infinite dynamic games with fixed prescribed duration. The chapter starts with a treatment of dynamic games defined in discrete time and with the number of players restricted to two. Continuous-time counterparts of most of these results and possible extensions to many-player games are treated in the latter part. The next two sections, i.e., Sections 7.2 and 7.3, deal, respectively, with the Stackelberg solution under open-loop information and the feedback Stackelberg solution under CLPS information. These solutions are obtained using two standard techniques of optimal control theory, viz. the minimum principle and dynamic programming, respectively. Derivation of the (global) Stackelberg solution under the CLPS information pattern, however, requires a much more subtle analysis since all standard techniques and approaches of optimal control theory fail to provide the solution. Therefore, Section 7.4 is devoted exclusively to this topic and to elucidation of an indirect approach toward derivation of the closed-loop Stackelberg solution. This indirect approach is first introduced within the context of two scalar examples (cf. subsections 7.4.1 and 7.4.2), and then its application is extended to the class of two-person linear-quadratic dynamic games in subsection 7.4.3. There is a clear relationship with the theory of "incentives", and this is the subject of subsection 7.4.4. Section 7.5 discusses derivation of the Stackelberg solution in stochastic dynamic games with deterministic information patterns, and primarily for the open-loop information (cf. subsection 7.5.1) and for the CLPS information structure under the "feedback Stackelberg" solution (cf. subsection 7.5.2). Derivation of the (global) Stackelberg solution under the CLPS

information pattern is a challenging topic, which is briefly discussed also in subsection 7.5.1. Stochastic incentive problems are discussed in subsection 7.5.2, as the counterpart of 365

366 T. BAŞAR AND G. J. OLSDER subsection 7.4.4 in the stochastic case. Section 7.6 treats continuous-time games, first presenting the Stackelberg solution under open-loop information for the leader (subsection 7.6.1), and then deriving the feedback Stackelberg solution (defined as a natural extension of the discrete-time FB Stackelberg solution concept) under CLPS information (subsection 7.6.2).

7.2 Open-Loop Stackelberg Solution of Two-Person Dynamic Games in Discrete Time In this section we discuss the Stackelberg solution for the class of two-person deterministic discrete-time infinite dynamic games of prescribed fixed duration (cf. Def. 5.1) when the underlying information structure is open-loop (for both players) and  $P_1$  acts as the leader. Hence, abiding by our standard terminology and notation, the state evolves according to  $X_{k+1} = A_k X_k + B_k u_k + w_k$ ,  $k \in K$ , (7.1) where  $X_k \in X = \mathbb{R}^n$ ,  $x_0$  is specified a priori, and  $u_k \in U_k \subset \mathbb{R}^m$ ,  $i = 1, 2$ ; and the stage-additive cost functional for  $P_i$  is introduced as  $J_i(u_1, u_2) = \sum_{k \in K} g_k(x_k, u_k, v_k)$ ,  $i = 1, 2$ . (7.2) Because of the open-loop information structure, the game is already in normal form, and therefore Defs 3.26, 3.27 and 3.28 readily introduce the Stackelberg equilibrium solution with  $P_1$  acting as the leader (since finiteness of the game was not essential in those definitions). One method of obtaining the Stackelberg solution for this class of games is to make use of the theory of Section 4.5, since these games can be viewed as static ones under the open-loop information structure; such an approach also readily leads to conditions for existence of a Stackelberg solution. Toward this end, we first note that, by recursive substitution of (7.1) into (7.2), the cost functional  $J_i(u_1, u_2)$  can be made to depend only on the control vectors  $u_1, u_2$  (which are of dimensions  $m_1 K$  and  $m_2 K$ , respectively) and the initial state  $x_0$ , which is known a priori. Now, if (i)  $J_i^*$  is continuous on  $U_1 \times U_2$  ( $i = 1, 2$ ), (ii)  $J_2(u_1, u_2)$  is strictly convex on  $U_2$  for all  $u_1 \in U_1$ , (iii)  $U_i$  is a closed and bounded (thereby compact) subset of  $\mathbb{R}^{m_i}$  ( $i = 1, 2$ ), then it follows from Corollary 4.4 (more generally from Thm. 4.8) that a Stackelberg equilibrium solution exists for the open-loop infinite game. A brute-force

HIERARCHICAL (STACKELBERG) EQUILIBRIA 367 method to obtain the corresponding solution would be first to determine the unique reaction curve of

the follower by minimizing  $J_2(u_1, u_2)$  over  $u_2 \in U_2$  for every fixed  $u_1 \in U_1$ , which is a meaningful optimization problem because of assumptions (i)-(iii) above. Denoting this unique mapping by  $T_2 : U_1 \rightarrow U_2$ , the optimization problem faced by PI (the leader) is then minimization of  $J_1(u_1, T_2 u_1)$  over  $U_1$ , thus yielding a Stackelberg strategy for the leader in this open-loop game. Despite its straightforwardness, such a derivation is not practical, especially if the number of stages in the game is large—which directly contributes to the dimensions of the control vectors  $u_1, u_2$ . An alternative derivation, which parallels that of Thm. 6.1 in the case of open-loop Nash equilibria, utilizes techniques of optimal control theory. Toward this end, we first determine the unique optimal response of the follower to every announced strategy  $\gamma_1 \in T_1$  of the leader. Since the underlying information pattern is open-loop, the optimization problem faced by the follower is (for each fixed  $u_1 \in U_1$ )  $\min_{u_2} L_2(u_1, u_2)$  subject to  $X_{k+i} = f_k(x_k, u_{jt}, u)$ ,  $x_i$  given. This is a standard optimal control problem whose solution exists and is unique under conditions (i)-(iii), and which can further be obtained by utilizing Thm. 5.5. Lemma 7.1 In addition to conditions (i)-(iii) assume that  $f_k(x, u, u)$  is continuously differentiable on  $\mathbb{R}^n$ , and  $f_k(-, u, -)$  is convex on  $\mathbb{R}^n \times U_2$ , ( $k \in K$ ), ( $v$ )  $J_2(u_1, u_2)$  is continuously differentiable on  $\mathbb{R}^n \times \mathbb{R}^n$ , ( $k \in K$ ). Then, to any announced strategy  $\gamma_1 = \gamma_1 \in T_1$  of the leader, there exists a unique optimal response of the follower (to be denoted by  $\hat{u}_2(x)$ ) satisfying the following relations:  $X_{k+1} = f_k(X_k, u_1, u_2)$ ,  $X_i = X_i$ ,  $u_2^k = \arg \min_{u_2} H(p_{k+1}, u_1, u_2, x_k)$ , (7.3a) (7.3b)  $\sum_{k \in K} \int_{x_k} f_k(x_k, u_1, u_2) dx_k + I^*(x_k)$  (7.3c)  $\nabla H(p_{k+1}, u_1, u_2, x_k) = 0$ , (7.3c) Except for very special cases, conditions (iv) and (v) given below, and conditions (vi), (vii), (ix) and (x) given later, are required for satisfaction of condition (i) (i.e., they are, in general, implicit in condition (i)); but we nevertheless rewrite them here since they will be needed explicitly in some of the results obtained in the sequel.

368 T. BAŞAR AND G. J. OLSDER where  $H(p_{k+1}, u_1, u_2, x_k) = \int_{x_k} (A^T p_{k+1} - f_k(x_k, u_1, u_2)) dx_k + I^*(x_k)$  (7.3d)  $+ \sum_{k \in K} \int_{x_k} f_k(x_k, u_1, u_2) dx_k$ . Here,  $\{p_i, - \cdot, - \cdot, p_{k+1}\}$  is a sequence of  $n$ -dimensional costate vectors associated with this optimal control problem. If we further assume (TM)  $f_k(x_k, u_1, u_2)$  is continuously differentiable on  $U_k$  ( $k \in K$ ), (vii)  $\hat{u}_2(x_{k+1}, u_1, -, X_k)$  is continuously differentiable on  $U_k$  ( $k \in K$ ), (viii)  $u_2$  in Lemma 7.1 is an inner-point solution for every  $u_1 \in U_1$ , then (7.3b) can equivalently be written as  $\nabla_{u_2} H(p_{k+1}, u_1, u_2, x_k) = 0$

( $k \in K$ ). (7.4) Now, to obtain the Stackelberg strategy of the leader, we have to minimize  $J^*(u^*, u^2)$ , in view of the unique optimal response of the follower to be determined from (7.4) in conjunction with (7.3a) and (7.3c). Therefore,  $P_i$  is faced with the optimal control problem  $\min L^*(u^*, u^2)$  (7.5a) subject to  $X_{k+1} = f_k(X_k, U_k, u^k)$ ,  $P_k = F_k(x_k, u^k, u^k, p_{k+1})$ ,  $\forall u^k \in H_k(P_{k+1}, u^k, x_k) = 0$   $x_i$  given,  $P_{K+1} = 0$ , ( $k \in K$ ), (7.5b) (7.5c) (7.5d) where  $H_k$  is defined by (7.3d) and  $d F_k = -Q^k - f_k(x_k, U_k, u^k)$ ,  $P_{k+1} + d T T - 9k(f_k(x_k, u^k, u^k) - u^k, u^k)$   $ax_k$  (7.5e) Note that, in this optimal control problem the constraint equations involve boundary conditions at both end points, and hence Thm. 5.5 is not applicable; but a standard result of nonlinear programming (cf. Canon et al., 1970) can be utilized to yield the following conclusion. Theorem 7.1 In addition to conditions (i)-(viii) stated in this section, assume that (be)  $f_k(x_k, -, u^k)$ ,  $g^k(x_{k+1}, -, v^k, x_k)$  are continuously differentiable on  $U_k$ ,  $k \in K$ ,

**HIERARCHICAL (STACKELBERG) EQUILIBRIA** 369 (x)  $g_l(-, \bullet, \bullet, \bullet)$  is continuously differentiable on  $R^n \times U^k \times U^0 \times R^n$ ,  $k \in K$ , (xi)  $f_k(-, u^k, \bullet)$  is twice continuously differentiable on  $R^n \times I^k$ , and  $g^k(-, u^k, \bullet, \bullet)$  is twice continuously differentiable on  $R^n \times U_k \times R^n$ ,  $k \in K$ . Then, if  $\{f_k^*(x_i) - u_k^* \in U_k, k \in K\}$  denotes an open-loop Stackelberg equilibrium strategy for the leader in the dynamic game formulated in this section, there exist finite vector sequences  $\{A_i, \dots, A^k\}$ ,  $\{v_i, \dots, v_k\}$ ,  $\{v^k, \dots, v^k\}$  that satisfy the following relations:  $4+i = f_k\{x_i, u^k, u^k\}$ ,  $x^k = x_1$ , (7.6a)  $\forall u^k \in H_k\{k, v^k, v^k, p_{k+1}, u^k, u^k, x^k\} = 0$ , (7.6b)  $\forall u^k \in H_k\{k, v^k, v^k, p_{k+1}, u^k, u^k, x^k\} = 0$ , (7.6c)  $K-i = H_k(x_k, v^k, v^k, p_{k+1}, u^k, u^k, x^k)$ ,  $X_K = 0$ , (7.6d)  $l_i^k = g^k\{x_i, u^k, u^k, x^k\}$ ,  $H_i = 0$ , (7.6e)  $\forall u^k \in H_k\{k, v^k, v^k, p_{k+1}, u^k, u^k, x^k\} = 0$ , (7.6f)  $p_l = F_k(x_l, u^k, u^k, p_{l+1})$ ,  $p_{K+1}^k = 0$ , (7.6g) where  $W^k = \{x_i, u^k, u^k, x^k\}$  is defined by (7.3d),  $F_k$  by (7.5e), and  $U_k$  denotes the interior of  $U_k$ . Furthermore,  $\{u^k, k \in K\}$  is the corresponding unique open-loop Stackelberg strategy of the follower (P2), and  $\{x^k, k \in K\}$  is the state trajectory associated with the Stackelberg solution. Proof. As discussed prior to the statement of the theorem, the leader ( $P_i$ ) is faced with the optimal control problem of minimizing  $L^*(u^1, u^2)$  over  $U^1$  and subject to (7.5b)-(7.5d). Since the problem is formulated in discrete time, it is equivalent to a finite dimensional nonlinear programming problem the "Lagrangian" of which is  $L = \sum_{k=1}^K \{g^k(x_{k+1}, u^k, u^k, x_k) + X^k [f_k(x_k, u^k, u^k) - x_{k+1}] + v^k [-f_k(x_k, u^k, u^k) - p_{k+1}] + u^k [-f_k(x_k, u^k, u^k) - p_{k+1}]\}$ , where  $X_k: v^k, v^k$  ( $k \in K$ ) denote appropriate Lagrange-multiplier vectors.

Now, if  $\{u_k^*, k \in K\}$  is a minimizing solution and  $\{x_{k+1}, p_{k+1}, u_k^*; k \in K\}$  are the

370 T. BAŞAR AND G. J. OLSDER corresponding values of the other variables so that the constraints (7.5b)-(7.5d) are satisfied, then it is necessary that (see, e.g., Canon et al., 1970, p. 51)  $V_{,i}L = 0$ ,  $V_{u^i}L = 0$ ,  $V_{x^i}L = 0$ ,  $V_{p^i}L = 0$ , ( $i \in K$ ) wherefrom (7.6b)-(7.6e) follow. **D Remark 7.1** If the follower (P2) has, instead, access to closed-loop perfect state (CLPS) information, his optimal response (cf. Lemma 7.1) will be any closed-loop representation of the open-loop policy  $\{u_k^*, k \in K\}$ ; however, since the constraints (7.5b)-(7.5d) are basically open-loop relations, these different representations do not lead to different optimization problems for the leader (P1). Therefore, the solution presented in Thm. 7.1 also constitutes a Stackelberg equilibrium for the case in which the follower has access to CLPS information (with the leader still having OL information). Under this latter information pattern, we may still talk about a unique open-loop Stackelberg strategy for the leader; whereas for the follower the corresponding optimal response strategy will definitely not be unique, since any closed-loop representation of  $u^*$  given in Thm. 7.1 will constitute an optimal strategy for P2. **□ Remark 7.2** The open-loop Stackelberg (OLS) solution is not weakly time consistent (a la Def. 5.13), since the OLS solution of any subgame does not necessarily coincide with the OLS solution of the full game, not even on the equilibrium trajectory. **□ Linear-quadratic games** We now specialize the result of Thm. 7.1 to the class of linear-quadratic dynamic games (cf. Def. 6.1) wherein  $f_k(x_k, u_k) = A_k x_k + B_k u_k$ , (7.7a)  $g_k(x_k, u_k) = x_k^T Q_k x_k + x_k^T R_k u_k + u_k^T U_k u_k$ ,  $j = 1, 2$ . (7.7b) We further assume that  $Q_{k+1} > 0$ ,  $R_j > 0$ , and  $U_k = R_k$   $\forall k \in K$ ,  $i = 1, 2$ . Under these assumptions, it readily follows that the optimal response of the follower to any announced control sequence of the leader is unique and affine, and moreover the minimum of  $L_1$  under the constraint imposed by this affine response curve is attained uniquely by an open-loop control sequence (for P1) that is linear in  $x$ . **7.1** This unique solution can be obtained explicitly by utilizing Thm. 7.1 and in terms of the current (open-loop) value of the state; it should be noted that Thm. 7.1 provides in this case both necessary and **7.1** This follows since  $L_1(u_1, u_2)$  is quadratic and strictly convex on  $R^m \times R^n$ . It should also be noted that the assumptions made on  $Q_{k+1}$  and  $R_j$  are sufficient, but not necessary, for the open-loop Stackelberg solution to exist and be unique; they are made here for

convenience in the derivation to follow.

HIERARCHICAL (STACKELBERG) EQUILIBRIA 371 sufficient conditions because of strict convexity.<sup>72</sup> The result is given below in Corollary 7.1, which is obtained, however, under a condition of invertibility of a matrix related to the parameters of the game. Corollary 7.1 The two-person linear-quadratic dynamic game described by (7.7a)-(7.7b) under the parametric restrictions  $Q_k > 0, R_k > 0$  ( $k \in K, i = 1, 2$ ) admits a unique open-loop Stackelberg solution with  $P_i$  acting as the leader, which is given, under the condition of invertibility of the matrix appearing in the braces in (7.9c), by  $\langle x^* = 11k^*(x_1) = -B1'K1kxi, (7.8a) \langle ** = 11^*(x_1) = \sim B2kK2x^*k (fceK), (7.8b)$  where  $K \setminus = [I + Q1 + 1B1B2k r > Ak + 1\$k + Q1 + 1B1[I + B1'Q1 + 1BirR12B1'Pk + ^k + Q1 + x\{I + B2kB2kQ2k + 1\} - 1AkMk, (7.9a) K2 \pounds Pk + 1\$k (7.9b) *fc ^ \{1 + B2B2kPk + 1 + BkBI'Q2k + 1B2k(I + B2'Q1 + 1B2r1Rk2B2kPk + 1 + B1BI\| + Q1 + 1B1BZ') - 1Ak + 1\} \sim 1 \blacksquare \setminus Ak - B1BI'Q2 + 1(I + B2B2kQ2 + ir1AkMk), (7.9c)$  and  $P_k, A_k, M_k$  are appropriately dimensioned matrices generated by the difference equations  $P_k = A'kPk + 1 + A'kQ2k + 1\$k; PK + 1 - Q2K + 1, (7.10a) A^* = A'k[Q1 + 1\$k + Ak + 1 + Q2k + 1AkMk - Q2k + 1B2Nk\}; AK + 1 - QxK + 1 (7.10b) Mk + 1 = AkMk - B2kNk; Mx = 0 (7.10c) with Nk^(I + B2kQ2k + 1B2) - 1[B2k\{Ak + ^k + Q2k + 1AkMk\} - Rk2B2kK2). (7.10d) Furthermore, the unique state trajectory associated with this pair of strategies satisfies  $x_{1+i} = ^kX^*k, x_1 = xi. (7.10e)$  <sup>72</sup> Note that condition (iii) in the general formulation is also met here since we already know that a unique solution exists with  $U_i$  taken as  $R_m^*$  and therefore it can be replaced by an appropriate compact subset of  $R_{mi}$  ( $i = 1, 2$ ).$

372 T. BA§AR AND G. J. OLSDER Proof. Existence of a unique solution to the linear-quadratic open-loop Stackelberg game has already been verified in the discussion prior to the statement of the corollary. We now apply Thm. 7.1 to obtain the corresponding equilibrium strategies of the players, in the structural form (7.8a)-(7.8b). Toward this end, we first rewrite (7.6b)-(7.6g), under the "linear-quadratic" assumption (7.7a)-(7.7b), respectively, as follows:  $\langle 1^* + BZQ1 + 1xUi + Blk \setminus k + BlkQ1 + 1Aknk + B\} iQ1 + 1B2kvk = 0, (i) R12ul^* + B^*Q1 + 1xk + 1 + B^*k + BtQ1 + 1Akfik + (I + B\%Q1 + 1B1)vk = 0, (ii) A^* \_ ! = A'kQ1 + 1x^*k + 1 + A'kXk + A'kQ2k + x(Aklxk + B2kvk); \setminus K = 0, (Hi) Hk + i = AkHk + B \setminus vk; m = 0, (iv) v1^* + B1'(Q1 + 1x^*k + 1 + pi + 1) = 0, (v) pi = A'k \setminus pi + i + Q1 + ix1 + i]; p^*k + i = o- (\langle *) The coefficient of  $v_k$  in (ii) is clearly positive$

definite; therefore, solving for  $v_k$  from (ii) and substituting this in (i), (iii) and (iv), we obtain, respectively,  $4^* = -Bx_k\{Q^{k+1}-Q^{k+1}Bl\{I + BtQ^{k+1}Bl\}-$   
 $'BtQ^{k+1}\}x^{k+1} \sim Bx_k [I - Q^{k+1}B^2k(I + B^*Q^{k+1}Bl)-^*B^*]k -Bi'Q^{2k+1}[I - Bl(I +$   
 $B^*Q^{k+1}Bl)^{B^*Q^{k+1}}]Ak^{nk} +BlQ^{k+1}xBl\{I + BtQ^{k+1}xBl\}-'R^ut, (vii) U-i$   
 $A^k[Q^{i+i} - Q^{i+i}BUl + BiQ^{i+i}Bl)-'B^*Q^{i+i}\}x^{i+i} +A^k\{I - Q^{i+i}Bl(I +$   
 $B^*Q^{i+i}Bl)-lB^*\}k +A^kQ^{i+i}lI-BtI + B^*Q^{i+i}Bl)-iB^*Q^{i+i}\}Ak^k [U) -$   
 $A^kQ^{i+i}(I + BtQ^{i+i}Bl)-iR^ul^*; XK = 0, W^{i+1} = \{I-Bl\{I + BtQ^{2k+i}Bl\}-$   
 $'B^*Q^{i+1}\}AkHk -B^2k(I + B^2kQ^{i+1}Bl)-1 (ix) iRk^2\{k + B^2k (Q^{i+1}x^{i+1} +$   
 $A^*)\}; m = 0. Now, letting  $p^{*k+1} = (P^{k+1}X -Q^{2k+1})x^{*k+1}, Xk =$   
 $\{^k+i\sim Q^{k+i}\}x^{i+1}, Hk = M^kx^i,$  and substituting (vii) and (v) into the state  
equation, we obtain the relation  $x^{k+1} = $kx^k (x)$  assuming that the required  
inverse in the expression for  $$k$  (defined by (7.9c)) exists. In determining the  
expression for  $$k$ , as given by (7.9c), we have also utilized the matrix identity  
of Lemma 6.2 in simplifying some of the terms involved. Now, finally, by  
further making use of (x) in (vii), (v), (vi), (viii) and (ix), we arrive at (7.9a),  
(7.9b), (7.10a), (7.10b) and (7.10c), respectively.  $\square$$

**HIERARCHICAL (STACKELBERG) EQUILIBRIA 373** Remark 7.3 The  
coupled matrix difference equations (7.10a)-(7.10c) constitute a two-point  
boundary value problem, since the starting conditions are specified at both end  
points. If the problem is initially formulated as a (high dimensional) static  
quadratic game problem, however, such a two-point boundary value problem  
does not show up; but then one is faced with the task of inverting high  
dimensional matrices (i.e., the difficulty again shows up in a different form).  $\square$   
Remark 7.4 Corollary 7.1 extends naturally (without any conceptual  
difficulties) to affine-quadratic games where the state equation has an  
additional control-independent driving term and the cost functions have cross  
terms between the current values of the state vector and the players' control  
vectors. For this more general case, the follower's optimal response will again  
be an affine function of the leader's control, and the leader's open-loop  
Stackelberg control will be affine in the initial state  $x_i$ , obtained from the  
solution of a generalized quadratic optimization problem.  $\square$  7.3 Feedback  
Stackelberg Solution Under CLPS Information Pattern The feedback  
Stackelberg solution concept introduced in Def. 3.29 for finite games is  
equally valid for infinite dynamic games defined in discrete time and with  
additive cost functionals, provided that the underlying information structure  
allows the players to have access to the current value of the state (such as in



the case of CLPS or MPS information pattern). Furthermore, in view of Prop. 3.15 (which is also valid for infinite games denned in discrete time), it is possible to obtain the solution recursively (in retrograde time), by employing a dynamic programming argument and by solving static Stackelberg games at every stage. Toward this end, we consider, in this section, the class of two-person dynamic games described by the state equation (7.1), and the cost functionals  $J_i(u_1, u_2)$  given by (7.2). Under the delayed-commitment mode of play, we restrict attention to feedback strategies, i.e.,  $u_i^k = f_i(x^k)$ ,  $k \in K$ ,  $i = 1, 2$ , and seek to obtain the feedback Stackelberg solution that is valid for all possible initial states  $x^0 \in R^n$ . To complete the description of the dynamic Stackelberg game, let  $T_i$  denote the class of all permissible strategies of  $P_i$  at stage  $k$  (i.e., measurable mappings from  $R^n$  into  $U^i$ ). Then, the cost functional for  $P_i$ , when the game is in normal form, is (as opposed to (7.2))  $J_i(u_1, u_2) = \sum_{k \in K} \gamma^k g_i(x^k, u_1^k, u_2^k)$ ;  $j \in T_j$ ,  $k \in K$ ,  $j = 1, 2$ . Because of the additive nature of this cost functional, the following theorem now readily follows from Def. 3.29 and Prop. 3.15 (interpreted appropriately for the infinite dynamic game under consideration), and it provides a sufficient condition for a pair of strategies to constitute a feedback Stackelberg solution. Such

374 T. BAŞAR AND G. J. OLSDER a solution is clearly also a strongly time consistent one, using the terminology introduced in Def. 5.14 with the sol operator taken as the feedback Stackelberg equilibrium solution. Theorem 7.2 For the two-person discrete-time infinite dynamic game formulated in this section, a pair of strategies  $\{u_1^* \in T_1, u_2^* \in T_2\}$  constitutes a feedback Stackelberg solution with  $P_1$  as the leader if  $u_1^k \in R^1(x^k) = G_1(u_2^*, x^k)$ ,  $\forall x^k \in R^n$ , (7.11) where  $R^1(x^k)$  is a singleton set defined by  $R^1(x^k) = \{u_1 \in T_1 \mid G_1(u_1, u_2^*, x^k) = \min_{u_1 \in T_1} G_1(u_1, u_2^*, x^k)\}$ , (7.12a)  $G^k(u_1^k, u_2^k, x^k) = G^k(u_1^k, u_2^k, x^k)$  (7.12b)  $i = 1, 2$ ,  $k \in K$ ,  $v \in R^n$ ; and  $G^k$  is recursively defined by  $G^k(u_1^k, u_2^k, x^k) = G^k(u_1^k, u_2^k, x^k) + \gamma^{k+1} [G^{k+1}(u_1^{k+1}, u_2^{k+1}, x^{k+1}) - G^k(u_1^k, u_2^k, x^k)]$ ,  $G^{K+1} = 0$ ,  $i = 1, 2$ . Furthermore, every such solution is strongly time consistent. Remark 7.5 If the optimal response set  $R^2(x^k)$  of  $P_2$  (the follower) is not a singleton, then (in accordance with Def. 3.29) we have to take the maximum of the LHS of (7.11) over all these nonunique response functions. However, if  $G^k(u_1^k, u_2^k, x^k)$  is strictly convex on  $U^2$  ( $k \in K$ ), the underlying assumption of "unique follower response" is readily met, and the theorem provides an easily implementable algorithm for computation of

feedback Stackelberg strategies. □ Linear-quadratic games For the special case of linear-quadratic games with strictly convex cost functionals, the feedback Stackelberg solution of Thm. 7.2 can be determined in closed form. Toward this end, let  $f_k$  and  $g_k$  be as defined by (7.7a) and (7.7b), respectively, with  $Q_{i+1} > 0$ ,  $R_{k3} > 0 \forall k \in K$ ,  $i, j = 1, 2$ ,  $i \wedge j$ , and take  $U^k = R^m$ . Then, solving recursively for  $G_k$ , in view of (7.11) and (7.12a), we obtain the quadratic structure  $G^*(k+i) = i^k + i^{k+1} + \dots + u_i R^{i-1} u_i^T$ ,  $i = 1, 2$ ;  $j \wedge i$ ,

HIERARCHICAL (STACKELBERG) EQUILIBRIA 375 where  $L^k$  satisfies the backward recurrence relation  $L^k = (A_k - B_i S_i - B_i S_i^T) L^{k+1} + S_i^T \{ A_k - B_i S_i - B_i S_i^T \} + S_i^T R_i^{-1} S_i + Q_i$ ;  $L^{K+1} = Q^{k+1}$  ( $i, j = 1, 2$ ;  $j \neq i$ ),  $S_i = [B_i^T \{ I + L^{k+1} B_i B_i^T \}^{-1} \{ I + B_i^T X_{L^{k+1}} \}^{-1} B_i^T L^{k+1} X_{L^{k+1}} + B_i^T L^{k+1} B_i + T_i^* \{ I + B_i^T L^{k+1} B_i \}^{-1} R_i^{-1} \{ I + B_i^T L^{k+1} B_i \}^{-1} L^{k+1} X_{L^{k+1}} + L^{k+1} B_i \{ I + B_i^T L^{k+1} B_i \}^{-1} R_i^{-1} \{ I + B_i^T L^{k+1} B_i \}^{-1} L^{k+1} X_{L^{k+1}} \} A_k$ , (7.15a)  $S_i = (I + B_i^T L^{k+1} B_i)^{-1} B_i^T L^{k+1} (A_k - B_i S_i)$ . (7.15b) It should be noted that  $L^{k+1} > 0$ ,  $\forall k \in K$ ,  $i = 1, 2$  (since  $Q_{i+1} > 0$ ,  $R_{kj} > 0$ ), and therefore the required inverses in (7.15a) and (7.15b) exist. By the same token,  $G^k(x^k)$  is strictly convex on  $R^m$  and hence  $R^k$  is a singleton for every  $k \in K$ ; furthermore, the minimization problem (7.11) admits a unique solution, again because of the strict convexity (on  $R^m$ ). The implication, then, is the following corollary to Thm. 7.2, which summarizes the feedback Stackelberg solution of the linear-quadratic game.

Corollary 7.2 The two-person linear-quadratic discrete-time dynamic game, as defined by (7.7a) and (7.7b) and with strictly convex cost functionals and FB, CLPS or MPS information pattern, admits a unique feedback Stackelberg solution with  $P_i$  as the leader (under the delayed-commitment mode of play), which is strongly time consistent and linear in the current value of the state:  $Y_k^*(x_k) = -S_k x_k$ ,  $k \in K$ ;  $i = 1, 2$ . Here  $S_k$  and  $Q_k$  are defined by (7.15a) and (7.15b), respectively, via the recursive relations (7.14). Remark 7.6 The counterpart of Remark 7.4 holds here, in the sense that Corollary 7.2 extends naturally to anne-quadratic games where the state equation has an additional control-independent driving term and the cost functions have cross terms between the current values of the state vector and the players' control vectors. In this case, the unique feedback Stackelberg strategies of the players will be in the form:  $Y_k^*(x_k) = -S_k x_k + u_k$ ,  $k \in K$ ;  $i = 1, 2$ , where  $S_k$  and  $Q_k$  are some matrix sequences generated as in (7.15a) and (7.15b), and  $s_k$  and  $s_k^T$  are state-independent vector sequences, again generated in retrograde time. The precise expressions for these can readily be obtained by observing that in this more

general case the counterpart of  $G$  (cost-to-go for  $P_i$ ) will also contain a state-independent term, a term that is linear in  $x_k$ , and some (quadratic) cross terms between  $x_k$ ,  $u_j$ , and  $u_k$ .  $\square$

376 T. BAŞAR AND G. J. OLSDER 7.4 (Global) Stackelberg Solution Under CLPS Information Pattern We consider, in this section, the derivation of (global) Stackelberg solutions when the leader has access to dynamic information (e.g., CLPS information). Such decision problems cannot be solved by utilizing standard techniques of optimal control theory (as in Sections 7.2 and 7.3), because the reaction set of the follower cannot, in general, be determined in closed form, for all possible strategies of the leader, and hence the optimization problem faced by the leader on this reaction set becomes quite an implausible one (at least, using the standard available techniques of optimization). In finite games (cf. Section 3.6) this difficulty can definitely be circumvented by converting the original game in extensive form into normal form (which is basically a static (matrix) game) and then by performing only a finite number of comparisons. In infinite dynamic games, however, such an approach is apt to fail because of the "infinite" nature of the strategy space of the leader (i.e., he has uncountably many alternatives). In the sequel we first consider a specific example to elucidate the difficulties to be encountered in a brute-force derivation of the Stackelberg solution when the leader has access to dynamic information, and introduce (within the context of that example which involves two stages—each player acting only once, first the follower and then the leader) an indirect method to obtain the solution (see subsection 7.4.1). Then, we extend this indirect approach to a modified version of the example of subsection 7.4.1, in which the follower also acts at the second stage (see subsection 7.4.2) and we obtain in subsection 7.4.3 closed-form expressions for linear Stackelberg strategies in general linear-quadratic dynamic games, by utilizing the methods introduced in subsections 7.4.1 and 7.4.2. Finally, in subsection 7.4.4 we provide a brief discussion on the closely related topic of incentives.

7.4.1 An illustrative example (Example 7.1) Example 7.1 Consider the two-stage scalar dynamic game described by the state equations  $\dot{x}_2 = x_1 - u_1$  (7.16)  $x_3 = x_2 - u_2$  and cost functionals  $J_1 = (x_3)^2 + 2(u_1)^2 + (3u_2)^2$ ,  $3 > 0$ , (7.17a)  $J_2 = (x_3)^2 + (u_2)^2$ . (7.17b) (Note that we have suppressed the subindices on the controls, designating the stages, since every player acts only once in this game.) The information structure of the problem is CLPS, i.e.,  $P_1$  (the leader), acting at stage 2, has access to both

$x_1$  and  $X_2$ , while  $P_2$  (the follower) has only access to the value of  $x_1$  therefore, a permissible strategy  $\gamma_1$  for  $P_1$  is a measurable mapping from  $R \times R$

HIERARCHICAL (STACKELBERG) EQUILIBRIA 377 into  $R$ , and a permissible strategy  $\gamma_2$  for  $P_2$  is a measurable mapping from  $R$  into  $R$ . At this point, we impose no further restrictions (such as continuity, differentiability) on these mappings, and denote the strategy spaces associated with them by  $T_1$  and  $T_2$ , respectively. Now, to any announced strategy  $\gamma_1 \in T_1$  of the leader, the follower's optimal reaction (for fixed  $x_1$ ) will be the solution of the minimization problem  $\min_{u_2 \in R} \{ [x_1 - u_2 - \gamma_1(a_1 - u_2, :n)]^2 + [u_2]^2 \}$  (7.18) which determines  $R_2(\gamma_1)$ . For the sake of simplicity in the discussion to follow, let us now confine our analysis to only those strategies in  $T_1$  which lead to a singleton  $R_2(\gamma_1)$ . Denote the class of such strategies for the leader by  $T_1$ , and note that the follower's response is now a mapping  $\gamma_2 : R \times F_1 \rightarrow R$ , which we symbolically write as  $\gamma_2[a_1; \gamma_1]$ . There is, in general, no simple characterization of  $\gamma_2$  since the minimization problem (7.18) depends "structurally" on the choice of  $\gamma_1 \in T_1$ . Now, the optimization problem faced by the leader is  $\min_{\gamma_1 \in T_1} \{ \int_{t_1}^T [x_1 - \gamma_1(x_1)]^2 + 2 \int_{t_1}^T [x_1 - \gamma_1(x_1)]^2 + m_2 [x_1]^2 \}$ , where  $t_1 \in [0, T]$ . This is a constrained optimization problem which is, however, not of the standard type, since the constraint is in the form of the minimum of a functional. If we further restrict the permissible strategies of the leader to be twice continuously differentiable in the first argument (in which case the underlying  $T_1$  becomes a smaller set), the first- and second-order conditions associated with the minimization problem (7.18) can be obtained, which implicitly determine  $\gamma_1^*$ . The problem faced by the leader then becomes one of calculus of variations, with equality and inequality constraints, which is quite intractable even for the "simple" problem formulated in this subsection. Besides, even if the solution of this constrained calculus of variations problem is obtained (say, through numerical techniques), we cannot be sure that it constitutes a Stackelberg strategy for the leader in the game problem under consideration (i.e., it might not be the best choice for the leader) since the a priori assumption of twice differentiability might be overly restrictive. We shall, in fact, see in the sequel that, depending on the value of  $\lambda$ , the Stackelberg strategy of the leader could be nondifferentiable. These considerations then force us to abandon the direct approach toward the solution of the Stackelberg game under CLPS information pattern, and to seek indirect methods. One such indirect method involves determination of a (tight)

lower bound on the (Stackelberg) cost of the leader, and then finding an appropriate strategy for the leader that realizes that bound in view of possible rational responses of the follower.

378 T. BAŞAR AND G. J. OLSDER In an attempt to explore this indirect method somewhat further, let us first observe that the lowest possible cost the leader can expect to attain is  $J_1^*$  (7.19) which is the one obtained when the follower cooperates with him to (globally) minimize his cost functional. This quantity clearly provides a lower bound to the leader's Stackelberg cost; but it is not as yet known whether it can be realized,<sup>73</sup> and this is what we investigate in the sequel. First, solving the optimization problem (7.19) in view of (7.16) and (7.17a), we obtain (by utilizing dynamic programming or directly Prop. 5.1) the pair of feedback strategies  $u_1(x) = \dots$ , (7.20a)  $u_2(x) = [2/(3+2)]x$  (7.20b) as providing the unique globally minimizing solution within the class of feedback strategies. The corresponding optimum state trajectory, in this two-stage team problem, is described by  $x = [30/(30+2)]x_0$ ,  $\dot{x} = (2/3)x$ ,  $J^* = (7.20c)$  and the minimum cost is  $J_1^* = [2/(3+2)]x_0^2$ . (7.20d) The superscripts in (7.20a)-(7.20d) stand for "team", since (7.19) actually describes a two-member team problem with the common objective being the minimization of (7.17a), and (7.20a)-(7.20b) is the unique team-optimal solution in feedback strategies. We should note, however, that if PI is also allowed to know (and utilize) the value of the initial state  $x_0$ , (7.20a) is no longer a unique team-optimal strategy for PI, but any representation of it on the state trajectory (7.20c) also constitutes an optimal strategy (cf. Section 5.6). Denoting the class of all such strategies by  $T_1$ , we now have the more general result that any pair  $(u_1, u_2)$  with  $u_2(x) = [2/(3+2)]x$  constitutes a team-optimal solution to (7.19). For each such pair, the state trajectory and the corresponding cost value are still as given by (7.20c) and (7.20d), respectively. Now, if (7.19), or equivalently (7.20d), is going to be realized as the Stackelberg cost of the leader, there should be an element of  $T_1$ , say  $(u_1^*, u_2^*)$ , which forces  $J_1^*$ . Note that the follower in fact attempts to minimize his own cost functional which is different from  $J_1$ .

HIERARCHICAL (STACKELBERG) EQUILIBRIA 379 the follower to play  $u_2$  even though he is minimizing his own cost functional (which is  $J_2$ —the counterpart of  $L_2$  in the extensive form description). This, inevitably, asks for a relation of the form  $u_2^* = \arg \min J_2(u_1^*, u_2)$ , and this being so uniquely, i.e.,

the minimum of  $J_2(-y_u, -)$  on  $T_2$  should be attained uniquely at  $f_2$ . Intuitively, an appropriate strategy for the leader is the one that dictates (7.20a) if the follower plays (7.20b), and penalizes the follower rather heavily otherwise. Since the leader acts, in this game, after the follower does, and since the state equation is scalar and linear, he (the leader) is in a position to implement such a strategy (by utilizing solely the state information available to him). In particular, the strategy where  $k$  is any positive number, is one such candidate. Such a strategy is clearly in  $T_1$ . Moreover, one may readily check validity of the inequality  $\min J_2(71 = gX_2, 72) < \min J_2^1 = -f_{cci}, 72) 72ep2 v ' J \blacksquare'' / 72gr2 nil$  for  $k > 0, 75$  which implies that the strategy (7.21) does indeed force the follower to the team strategy (7.20b), since otherwise he incurs a cost which is higher than the one obtained under the strategy (7.20a). The conclusion, therefore, is that the strategy (7.21) constitutes, for  $k > 0$ , a Stackelberg strategy for the leader, with the unique optimal response of the follower being (7.20b).  $\square$  To recapitulate: (1) For the dynamic game described by (7.16), (7.17a) and (7.17b) and under the CLPS information pattern, a Stackelberg solution exists for all nonnegative values of  $k$ . (2) The Stackelberg cost of the leader ( $P_1$ ) is equal to the global minimum value of his cost function (obtained by cooperative action of the follower), even though the follower ( $P_2$ ) seeks to minimize his own (different) cost functional and may not know the cost functional of  $P_1$ . (3) The leader's Stackelberg strategy is a representation of his optimal feedback strategy in the related team problem described by (7.19), and it necessarily involves memory. (A feedback strategy for  $P_1$  cannot be a Stackelberg strategy.) <sup>74</sup>In case of nonunique solutions, we have to take the supremum of  $J_1$  over all those solutions, which should equal to  $J_1$ . <sup>75</sup>This inequality, in fact, holds for  $k > \frac{1}{1 + \sqrt{8/13}}$ , but we consider only positive values of  $k$  for the sake of simplicity in presentation.

380 T. BAŞAR AND G. J. OLSDER (4) The Stackelberg strategy of  $P_1$  is nonunique (parameterized, in this case, by a positive scalar  $k$ ), but the optimal response of  $P_2$  to all those strategies of  $P_1$  is unique (and independent of  $k$ ); it is the strategy that minimizes the leader's cost function, even though the follower's objective is quite different. (5) The Stackelberg strategy of the leader, as given by (7.21), is not even continuous; this, however, does not rule out the possibility for existence of continuous (and differentiable) Stackelberg strategies. (6) The strategy given by (7.21) may be viewed as a threat strategy on part of the leader, since he essentially threatens the follower to make his

cost worse if he does not play the optimal team strategy (7.20b). One natural question that arises at this point is whether there exists a Stackelberg strategy for the leader which is structurally more appealing (such as continuous, differentiable, etc.). In our investigation in this direction, we necessarily have to consider only the strategies in  $T_1$ , since the quantity (7.19) constitutes a tight lower bound on the cost of the leader. In particular, we now consider only those strategies in  $T_1$  which are linear in the available information, which can be written as  $j_1 \{x_2, x_1\} = -X_2 + p \{x_2 - 3g + 2X_1\}$  (7.22) where  $p$  is a scalar parameter. To determine the value of  $p$  for which (7.22) constitutes a Stackelberg strategy for the leader, we have to substitute (7.22) into  $J_2$ , minimize the resulting expression over  $j_2 \in T_2$  by also utilizing (7.16), and compare the argument of this minimization problem with the strategy (7.20b). Such an analysis readily leads to the unique value  $p = \frac{1}{3}$  provided, of course, that  $\beta > 0$ . Hence, we have the following. Proposition 7.1 Provided that  $\beta > 0$ , the linear strategy constitutes a Stackelberg strategy for the leader in the dynamic game described by (7.16), (7.17a) and (7.17b), and the unique optimal response of the follower is as given by (7.20b). The corresponding state trajectory is described by (7.20c) and the leader's Stackelberg cost is equal to the team cost (7.20d). Therefore, with the exception of the singular case  $\beta = 0$ , the leader can force the follower to the team strategy (7.20b) by announcing a linear (continuously differentiable) strategy, which is definitely more appealing than (7.21) which

HIERARCHICAL (STACKELBERG) EQUILIBRIA 381 was discontinuous on the equilibrium trajectory. For the case  $\beta = 0$ , however, a continuously differentiable Stackelberg strategy does not exist for the leader (see Başar (1979b), pp. 23-25) for a verification), and he has to employ a nondifferentiable strategy such as (7.21). Remark 7.7 The linear strategy presented in Prop. 7.1 and the discontinuous strategies given by (7.21) do not constitute the entire class of Stackelberg strategies of the leader in the dynamic game of this subsection; there exist other, nonlinear (continuous and discontinuous) strategies for the leader which force the follower to the optimal team strategy (7.20b). Each such strategy can be determined by basically following the analysis that led to Prop. 7.1 under a different nonlinear representation of (7.20a) (instead of the linear representation (7.22)). □ Remark 7.8 The results of this example can also be visualized graphically for each fixed  $x_1$  and  $\beta$ . Let us take  $X_1 = 1$  and leave  $\beta$  as a parameter, in which case

the iso-cost curves of P2 look as depicted in Fig. 7.1, and the (open-loop) team solution for PI can be written as (from (7.20a)-(7.20d))  $u_1' = (3/(2 + 3/3))$ ,  $u_2' = 2/(2 + 3/3)$ . Figure 7.1: Graphical illustration of Example 7.1. Since there is a one-to-one correspondence between  $x_i$  and  $u_2$ , any permissible strategy of PI can also be written as a function of  $u_2$ , i.e.,  $u_1 = \hat{u}_1(u_2)$ . Now, if we can find a strategy  $\gamma_1^*$  with the property that its graph has only the point  $(u_1, u_2)$  in common with the set  $D = \{(u_1, u_2) : \dots\}$ ,

382 T. BAŞAR AND G. J. OLSDER then the follower (P2) will definitely choose  $u_2$ . Thus  $\gamma_1^*$  constitutes a Stackelberg strategy for  $P_1$ , leading to the global minimum of his cost function. It easily follows from Fig. 7.1 that, for  $\beta > 0$ , there exists a unique linear  $\gamma_1^*$  which satisfies the above-mentioned requirement; it is the tangent at  $(u_1, u_2)$  to the corresponding iso-cost curve of P2. For  $\beta = 0$  this tangent line turns out to be vertical (parallel to the  $u_1$ -axis), and hence cannot be described by a linear strategy; however, a Stackelberg strategy exists in the class of continuous strategies for the leader. (In Fig. 7.1 we have drawn the tangent lines for  $\beta = 2$  and  $\beta = 0$ , and for the latter case the dashed curve provides a continuous nonlinear Stackelberg strategy for the leader.)  $\square$  Remark 7.9 A common property of the Stackelberg solutions obtained for the dynamic game described by (7.16), (7.17a) and (7.17b) and under CLPS information, is that they are also Nash equilibrium solutions. To see this, we first note the obvious inequality  $J_2(\gamma_1^*, \gamma_2) < J_2(\gamma_1^*, \gamma_2^*)$  for such a Stackelberg equilibrium solution pair. Furthermore, the inequality  $J_1(u_1, u_2^*) < J_1(u_1^*, u_2^*)$  also holds, since the Stackelberg solution is also team-optimal (under  $J_1$ ) in this case. These two inequalities then imply that every such pair also constitutes a Nash equilibrium solution. We already know that the dynamic game under consideration admits an uncountable number of Nash equilibrium solutions (cf. Section 6.3); the above discussion then shows that, since the leader can announce his strategy ahead of time in a Stackelberg game, he can choose those particular Nash equilibrium strategies (and only those) which lead to an equilibrium cost that is equal to the global minimum of his cost functional.  $\square$  7.4.2 A second example (Example 7.2): Follower acts twice in the game Example 7.2 The indirect method introduced in the foregoing subsection for the derivation of the Stackelberg solution of dynamic games under CLPS information basically involves two steps: (1) determination of a tight lower bound on the cost function of the leader, which coincides with the global minimum value of a particular team problem; (2)



adoption of a particular representation of the optimal team strategy of the leader in this team problem, which forces the follower to minimize (in effect) the cost function of the team problem while he is in fact minimizing his own cost function. Even though this general approach is applicable in a much more general context (i.e., applicable to dynamic games with several stages and more complex dynamics), the related team problem is not always the one determined solely by the cost function of the leader (as in the previous subsection), in particular if the follower also acts at the last stage

HIERARCHICAL (STACKELBERG) EQUILIBRIA 383 of the game. The following scalar dynamic game now exemplifies derivation of the related team problem in such a situation. As a modified version of the scalar dynamic game of subsection 7.4.1, consider the one with dynamics  $\dot{x}_2 = x_1 - u_1$ ,  $\dot{x}_1 = x_2 - u_1 - u_2$ . (7.23) and cost functionals  $J_1 = (x_1)^2 + 2(u_1)^2 + \theta(u_2)^2$ ,  $\theta > 0$ , (7.24a)  $J_2 = (x_1)^2 + (u_1)^2 + (u_2)^2$ . (7.24b) The information structure of the problem is CLPS, and P1 (the leader) acting at stage 2 has a single control variable  $u_1$ , while P2 (the follower) acting at both stages 1 and 2 has control variables  $u_1$  and  $u_2$ , respectively. Let us introduce the strategy spaces  $T_1$  (for P1),  $T_1$  (for P2 at stage 1) and  $T_2$ , (for P2 at stage 2) in a way compatible with the governing information structure. Then, to any announced strategy  $\gamma_1 \in T_1$  of the leader, the follower's optimal reaction at stage 2 will be  $u_2^*(x_2, \gamma_1) = -\frac{1}{\theta} [x_2 - \gamma_1(x_2, x_1)]$  which is obtained by minimization of  $J_2$  over  $\gamma_2 \in T_2$ . Hence, regardless of what strategy  $\gamma_1 \in T_1$  the leader announces, and regardless of what strategy the follower employs at stage 1, his (the follower's) optimal strategy at stage 2 is a unique linear function of  $x_2$  and  $\gamma_1$ , as given above. Substituting this structural form into  $J_1$  and  $J_2$ , derived from (7.24a) and (7.24b), respectively, we obtain the reduced cost functionals  $J_1 = \frac{1}{2} [x_2 - \gamma_1(x_2, x_1)]^2 + 2[\gamma_1(x_2, x_1)]^2 + 3[\gamma_2(x_1)]^2$ ,  $J_2 = \frac{1}{2} [x_2 - \gamma_1(x_2, x_1)]^2 + [\gamma_2(x_1)]^2$ , and therefore the lowest possible cost value the leader can achieve is  $\min_{\gamma_1} \min_{\gamma_2} J_1(\gamma_1, \gamma_2)$  (7.25) which is the quantity that replaces (7.19) for this game. The original Stackelberg game has thus been converted into one wherein the follower does not act at the last stage (his only strategy now being  $\gamma_2$ ), and the related team problem is described by (7.25), whose optimal team solution in feedback strategies is  $\gamma_1^*(x_2) = \frac{1}{2} x_2$ , (7.26a)  $\gamma_2^*(x_1) = \frac{2}{(9/\theta + 2)} x_1$  (7.26b)

384 T. BAŞAR AND G. J. OLSDER with the corresponding unique team trajectory being  $\frac{1}{2} x_2 + \frac{2}{(9/\theta + 2)} x_1$ , (7.26c)  $\frac{4}{9} \theta \cdot J^*$  The associated minimum

team cost is  $J_1' = [2/3/(9/3 + 2)](X_1)^2$  (7.26d) which forms a lower bound on the leader's cost in this game, and is definitely higher than the quantity  $\min_{x_1} J_1(x_1, 7_0, 7_1, 2)$ . Now, paralleling the arguments of subsection 7.4.1, we obtain the following representation of (7.26a), on the trajectory (7.26c), to be a Stackelberg strategy for the leader, forcing the follower at the first stage to the team strategy (7.26b):  $u_1 = -\frac{1}{k} \frac{1}{1 + \frac{9}{113}}$  where  $k > 1 + \frac{9}{113}$ . The optimal response of the follower to any such strategy is  $u_2^*(x_1) = [2/(9/3 + 2)]x_1$ , (7.28a)  $f(x_2, x_1) = -\frac{1}{2}x_2^2 - u_2^*(x_2, x_1)$  (7.28b) leading to (7.26d) as the Stackelberg cost of the leader.  $\square$  To obtain the counterpart of Prop. 7.1 for the dynamic game of this subsection, we again start with a general linear representation of (7.26a), and by following similar arguments we arrive at the following proposition. Proposition 7.2 Provided that  $\beta > 0$ , the linear strategy  $W = \frac{1}{2} \frac{1}{\beta} \frac{1}{1 + \frac{9}{113}} \frac{1}{\beta} \frac{1}{\beta} (x_2, X_1) = -\frac{1}{2} X_2^2 - u_2^*(x_2, x_1)$  constitutes a Stackelberg strategy for the leader in the dynamic game described by (7.23), (7.24a) and (7.24b), and the unique optimal response of the follower is given by (7.28a), (7.28b). The corresponding state trajectory is described by (7.26c), and the leader's Stackelberg cost is as given by (7.26d). Remark 7.10 It can again be shown that, for  $\beta = 0$ , no continuously differentiable Stackelberg solution exists for the leader; therefore, he has to announce a strategy which is discontinuous in the derivative, such as (7.27), or the one  $u_1 = \frac{1}{2} \frac{1}{\beta} \frac{1}{\beta} \frac{1}{\beta}$ . The reader is asked to fill in the details of this derivation.

**HIERARCHICAL (STACKELBERG) EQUILIBRIA** 385 analogous to that discussed in Remark 7.7. Furthermore, the reader should note that, for the dynamic game under consideration, there does not exist a Stackelberg solution which also constitutes a Nash equilibrium solution; i.e., a counterpart of Remark 7.8 does not apply here. The main reason for this is the fact that the follower also acts at the last stage of the game.  $\square$  7.4.3 Linear Stackelberg solution of linear-quadratic dynamic games We now utilize and extend the indirect methods introduced in subsections 7.4.1 and 7.4.2 to solve for the Stackelberg solution of two-person linear-quadratic dynamic games within the class of linear-memory strategies, so as to obtain the counterpart of Props 7.1 and 7.2 in this more general context. The class of dynamic games under consideration is described by the state dynamics  $x_{k+1} = A_k x_k + B_k u_k + B_{2k} u_2$  (7.29) and cost functional (for P1 and P2, respectively)  $J_1 = \frac{1}{2} \sum_{k=0}^{N-1} (x_k^T Q_k x_k + u_k^T R_k u_k + V_N(x_N))$  (7-30a)  $J_2 = \frac{1}{2} \sum_{k=0}^{N-1} (x_k^T W_k x_k + u_2^T M_k u_2)$  (7-30b)

$z_k = i$  with  $Q_{k+1} > 0$ ,  $R_{k+1} > 0$ ,  $Q_{k+1} > 0$ ,  $R_{k+1} > 0$ ,  $\forall A; e K = \{1, \dots, K\}$ ,  $\dim(x) = n$ ,  $\dim(u_k) = r_{ii}$ ,  $i = 1, 2$ . The information structure of the problem is CLPS, under which the closed-loop strategy spaces for  $P_1$  (the leader) and  $P_2$  (the follower) are denoted by  $T_k$  and  $r_{2k}$ , respectively, at stage  $k \in K$ . Furthermore, let  $J_1$  and  $J_2$  denote the cost functionals derived from (7.30a) and (7.30b), respectively, for the normal (strategic) form of the game, under these strategy spaces. Since  $B_{2K}$  does not necessarily vanish in this formulation, the follower also acts at the last stage of the game, and this fact has to be taken into account in the derivation of a tight lower bound on the leader's cost functional. Proceeding as in subsection 7.4.2, we first note that, to any announced strategy  $f$ -tuple  $\{f^k \in F_k; A; e K\}$  by the leader, there corresponds a "robust" optimal reaction by the follower at stage  $K$ , which is determined by minimizing  $J_2(x^k, u^k)$  over  $u^k \in F_{2k}$ , with the remaining strategies held fixed. This quadratic strictly convex minimization problem readily leads to the relation  $12k[x^k \setminus 1k] = \{I + B_{2k}Q_{2k+1}B_{2k}\}^{-1}B_{2k}Q_{2k+1}\{A_k x^k + B_{2k}u^k\}$ , (7.31) where  $r^k = \{x^k, u^k, \dots, x^k\}$ , and  $12k[x^k \setminus i] \leq 1$  stands for " $f_{2k}\{r^k\}$ —the strategy of  $P_2$  at stage  $K$ —and displays the explicit dependence of the choice of  $f^k$  on  $f^k$ . It should be noted, however, that the structure of this optimal response of

386 T. BAŞAR AND G. J. OLSDER the follower does not depend on the choice of  $f^k$ : regardless of what  $f^k$  and the previously applied strategies are,  $f^k$  depends linearly on  $x^k$  and  $u^k$ . The unique relation (7.31) can therefore be used in (7.30a) and (7.30b) without any loss of generality, in order to obtain an equivalent Stackelberg game wherein the follower does not act at the last stage. The cost functionals of this new dynamic game in extensive form are  $J_1 = \sum_{k=0}^{K-1} (x^k + u^k)^T Q_{k+1} (x^k + u^k) + x^k{}^T R_{k+1} x^k + u^k{}^T R_{k+1} u^k$ , (7.32a)  $J_2 = \sum_{k=0}^{K-1} (x^k + u^k)^T Q_{k+1} (x^k + u^k) + x^k{}^T R_{k+1} x^k + u^k{}^T R_{k+1} u^k$ , (7.32b) where  $x^k + u^k = A_k x^k + B_{2k} u^k$ ,  $Q_{k+1} = [I - B_{2k} K^T Q_{1k+1} [I - B_{2k} K^T + T^T R_{k+1} T, Q_{2k+1} = Q_{2k+1} [I - B_{2k} K^T], T \in \{B_{2k} Q_{2k+1}, B_{1k} + I\}^{-1} B_{2k} Q_{1k+1}$ . Let us further introduce the notation  $f_{2k} = \{f_{2k}^i, i = 1, \dots, K-1\}$ ,  $J_1(x^k, u^k) = J_1(x^k, u^k)$ ,  $k \in K$ ,  $u^k = 12k(u^k), k = 0, \dots, K-1$ . Then, a lower bound on the leader's Stackelberg cost is, clearly, the quantity  $\min_{f^k} \min_{u^k} J_1(x^k, u^k)$  (7.33) - which is the minimum cost of a team problem in which both  $P_1$  and  $P_2$  strive to minimize the single objective functional  $J_1$ . This team problem admits a unique optimal solution within the class of feedback strategies, which is given in the following lemma whose proof readily follows from Prop. 5.1 by

an appropriate decomposition. Lemma 7.2 In feedback strategies, the joint optimization (team) problem defined by (7.33) admits a unique solution given by  $l_k(x_k) = -L_k x_k$ ,  $k \in K$ ,  $y_k(x_k) = -L_k x_k$ ,  $k \in K$ .

HIERARCHICAL (STACKELBERG) EQUILIBRIA 387 and with minimum cost  $J_1 = -x^T M_1 x$ , where  $M_1 = [I + S_1 B_1^T S_1 A_1, k \in K, = [R_1^2 + S_2^2 B_2^2] - S_2^2 A_2, k < K-1, = B_2^2 Q_{k+1}, = B_1^2 M_{k+1} \{I - B_2^2 [R_k^2 + B_2^2 M_{k+1} B_2^2] - B_2^2 M_{k+1}\}, k < K-1, S_2^2 = B_2^2 M_{k+1} \{I - B_1^2 [I + B_1^2 M_{k+1} B_1^2] - B_1^2 M_{k+1}\}, k < K-1. M(\cdot)$  is defined recursively by  $M_K = Q_K + F^T K^T F_k + V K L^T, M_k = Q_k + F^T M_{k+1} F_k + L_k^T L_k + L_k^2 R_k^{-2} L_k^2, k < K-1$ , where  $F_k \pm A_k - B_1 L_1 K L_1 K, F_k \pm A_k - B_1 L_1 - B_2 L_2, k < K-1$ . The optimum team trajectory is then described by  $x_{k+1} = F_k x_k, k \in K, x_0 = x_0$ . (7.34) In view of the analysis of subsection 7.4.2, we now attempt to find a representation of the team strategy  $\{u^k; k \in K\}$  on the optimum team trajectory (7.34), which provides a Stackelberg strategy for the leader. The following proposition justifies this indirect approach for the problem under consideration. Proposition 7.3 Let  $\{u^k; k \in K\}$  be a representation of the team strategy  $\{y^k; k \in K\}$  on the optimum team trajectory (7.34), such that every solution of the minimization problem  $\min_{J_2} \{u^k; k \in K\}$  is a representation of  $\{u^k; k \in K\}$  on the same team trajectory. Then,  $\{u^k; k \in K\}$  provides a Stackelberg strategy for the leader.  $L_1 S_1 K S_1$

388 T. BAŞAR AND G. J. OLSDER Proof. We have already shown, prior to the statement of Lemma 7.2, that the quantity (7.33) provides a lower bound on the cost function of the leader, for the dynamic game under consideration. But, under the hypothesis of Prop. 7.3, this lower bound is tight, and is realized if the leader announces the strategy  $\{u^k; k \in K\}$  which, thereby, manifests itself as a Stackelberg strategy for the leader. To obtain some explicit results, we now confine our attention to linear one-step memory strategies for the leader (more precisely, to those representations of  $\{u^k; k \in K\}$  which are linear in the current and most recent past values of the state), viz.  $u^k(x_k, X_{k-1}) = -L^k x_k + P_k [x_k - F_{k-1} X_{k-1}], k \in K - \{1\}$ , (7.35) where  $\{P_k; k \in K - \{1\}\}$  is a matrix sequence which is yet to be determined. This matrix sequence is independent of  $x_k$  and  $X_{k-1}$ , but it may, in general, depend on the initial state  $x_0$  which is known a priori. Our objective for the remaining portion of this subsection may now be succinctly stated as (i) determination of conditions under which a Stackelberg strategy for the leader exists in the structural form (7.35) and also satisfies the hypothesis of Prop. 7.3, (ii) derivation of the corresponding

strategy of the leader (more precisely, the matrices  $P_k, k \in K - \{1\}$ ) recursively, whenever those conditions are fulfilled. Toward this end, we substitute (7.35) into  $J_2(7.1, 7.2)$  minimize the resulting functional over  $u_2 \in T_2$  and compare the minimizing solutions with  $\{u_2^*; k \in K\}$  and their representations on the optimum team trajectory. Such an analysis leads to Thm. 7.3 given in the sequel. Preliminary notation for Theorem 7.3  $\Phi(\cdot)$ : an  $(m_2 \times m_2)$ -matrix defined recursively by  $\Phi^k = Q_k + -\Phi^k S^k + \Phi^k F^k + L_k R_k L_k + L_k L_k$ ;  $S^{k+1} = Q^{k+1} A^k$ ; an  $(m_1 \times n)$ -matrix defined recursively by (as a function of  $\{P_{k+1}, \dots, P_k\}$ )  $A^k = B^k P^{k+1} F^k - R^k L^k + B^k X^{k+1} F^k$ ;  $A^{K+1} = 0$ . (7.36)

Condition 7.1 For a given  $x \in R^n$ , let there exist at least one matrix sequence  $\{P_k, P_{k-1}, \dots, P_2\}$  that satisfies recursively the vector equation  $B^k P^k x + X^k F^k + B^k L^k x - L^k x = 0, k + 1 \in K$  (7.37) where  $A^k$  is related to  $\{P_k, \dots, P_{k+1}\}$  through (7.36), and  $x_k$  is a known linear function of  $x$ , as determined through (7.34). Theorem 7.3 Let Condition 7.1 be satisfied and let  $\{P^k(x), \dots, P^2(a; x)\}$  denote one such sequence. Then, there exists a Stackelberg solution for the dynamic game described by (7.29)-(7.30b) and under the CLPS information, which

HIERARCHICAL (STACKELBERG) EQUILIBRIA 389 is given by  $u_1^* = \{x_k, X_{k-1}, x\} = [P_k(x) - L^k] x_k - P^k(x) F^k X_{k-1}, k \in K - \{1\}$  (7.38a)  $u_2^*(x) = -L^k x, k \in K - \{K\}$ , (7.38b) where  $L^k (i = 1, 2; k \in K)$  were defined in Lemma 7.2. The corresponding Stackelberg costs of the leader and the follower are given, respectively, by  $J_1^* = -x^T M^k X^k$ , (7.39a)  $J_2^* = -x^T E^k X^k$ . (7.39b) Proof. Substituting  $u_j = J_k^*(-), k \in K$ , as given by (7.38a), into  $L_2$  defined by (7.32b), we obtain a functional which is quadratic and strictly convex in  $\{u_2; t \in K - \{K\}\}$ . Let us denote this functional by  $L(u_2, u_1^k)$  which is, in fact, what the follower will minimize in order to determine his optimal response. Strict convexity of  $L$  readily implies that the optimal response of the follower to the announced strategy  $u_1^k$ -tuple (7.38a) of the leader is unique in open-loop policies, and that every other optimal response in  $T_2$  is a representation of this unique open-loop strategy on the associated state trajectory. Consequently, in view of Prop. 7.3, verification of the theorem amounts to showing that the minimum of  $L(u_2, \dots, u_2^{K-1})$  is attained by the open-loop representation of the feedback strategy  $(K - 1)$ -tuple (7.38b) on the state trajectory described by (7.34); in other words, we should verify that the set of controls  $u_1^* = -L^k x, k \in K - \{K\}$  minimizes  $L(u_2, \dots, u_1^{K-1})$ . Now, since  $L$  is quadratic and strictly convex, it is both necessary and sufficient that  $\{u_2^*\}$  be

a stage-by-stage optimal solution; i.e., the set of inequalities  $L(u^k) < L(u^{k-1})$  (i) should be satisfied for all  $u \in R^m$  and all  $k = 1, \dots, K-1$ , where  $\{u^k\}$ -k denotes the entire sequence  $u^1, \dots, u^{k-1}$  with only  $u^k$  missing. Because of the additive nature of the cost function  $L$ , (i) can equivalently be written as  $L_k(u^k) < L_{k-1}(u^{k-1}, u^k)$  (ii) for all  $u \in R^m$  and all  $k = 1, \dots, K-1$ , where, for fixed  $k$ ,  $L_k(u^k) = L(u^k) = -W_{k+1}Q_{k+1}V_{k+1} + \sum_{j=k+1}^K \lambda_j^k + 2 \sum_{j=k+1}^K (y_j^i + Q_j + M_j + \dots) + 2 \sum_{j=k+1}^K (U_j + V_j - 2u_j^k)$  (iii)

390 T. BAŞAR AND G. J. OLSDER with  $V_{k+1} = F_k V_k + B^T P_{k+1}^T (y_{k+1} - F_k x_k)$ ,  $V_{k+2} = F_{k+1} y_{k+1} + B_{k+1} P_{k+1}^T [y_{k+1} - F_{k+1} x_k]$ ,  $y_{k+1} = \{A_{k+1} - B_{k+1} L_{k+1}\} x_{k+1} + B_{k+1} u_{k+1}$ , and  $\lambda_{k+1}^k = -L_{k+1} y_{k+1} + P_{k+1}^T \{y_{k+1} - F_{k+1} y_{k+1}\}$ ,  $j > k+1$ ,  $H_{k+1} = I - L_{k+1} A_{k+1}$ ,  $j = k$ . It is important to note that, at the RHS of (ii),  $L_k$  is not only a function of  $u$  but also of  $x_k$ , which we take at its equilibrium value  $x^k$  because of the RHS of (i) (i.e., the sequence  $\{u^1, \dots, u^{k-1}\}$  that determines  $x^k$  is already at equilibrium while treating (ii)). It is for the same reason that  $y_k$ , on the last row of (iv), is taken to be equal to  $x_k$ . Now, if the relation in the last line of (iv) is used iteratively in the remaining relations of (iv), we can express  $y_j$ ,  $j > k+2$ , in terms of  $y_{k+1}$  and  $x_k$ , and if this is used in (v),  $\lambda_j^k$ ,  $j > k+1$ , can be expressed in terms of the same variables. The resulting expressions are  $V_j = F(j, k+1) y_{k+1} + K(j, k+1) B_{k+1} P_{k+1}^T [y_{k+1} - F_{k+1} x_k]$ ,  $j > k+2$ ,  $H_j = -L_j F(j, k+1) y_{k+1} + [Y(j, k+1) - L_j] H(j, k+1) B_{k+1} P_{k+1}^T [y_{k+1} - F_{k+1} x_k]$ ,  $j > k+1$ , where  $F(j, k) = F_j \dots F_k$ ,  $F(j, j) = I$ ,  $H(j, k) = H(j, k+1) B_{k+1} P_{k+1}^T + \dots + F_{j+1} U_{j+1}$ ,  $H(j, j-1) = I$ ,  $H(j, i) = 0$ ,  $j < i$ ,  $Y(j, k) = P_{k+1}^T B_{k+1} \dots P_{j+1}^T B_{j+1} Y(j, j) = I$ . Utilizing these expressions, we can next show that  $V_{k+1} L^*(u^k) = 0$ ,  $\forall k = 1, \dots, K-1$ . (vi) Toward this end, we first obtain  $V_{k+1} L^* [u^k] = [B_{k+1} P_{k+1}^T \{B_{k+1}^T H'(K+1, k+1) Q_{k+1} F(K+1, k) K + Y, [B_{k+1}^T H'(U, k+1) Z_j F(j, k) - Y'(j, k+1) R_{k+1}^T] F(j, k)\}] (vii) + B_{k+1}^T x_{k+1} F_k - L_{k+1} D_{k+1}$ , where  $Z_j = L_j R_{k+1}^T + L_j L_{k+1} + Q_{k+1}$ . (iv)

HIERARCHICAL (STACKELBERG) EQUILIBRIA 391 Some further manipulations prove that (vii) can equivalently be written as  $V_{k+1} L^* [u^k] = [B_{k+1} P_{k+1}^T \{A_{k+1} F_k + B_{k+1} X_{k+1} F_k - L_{k+1}\} x_k]$  where  $A_{k+1}$  satisfies the recursive equation (7.36) with  $P_{k+1}$  replaced by  $P_{k+1}$ . But, by Condition 7.1 [(7.37)], the foregoing expression identically vanishes; this proves (vi), and in turn establishes the desired result that  $u^k$  minimizes  $L$  at stage  $k$  and is in equilibrium with  $\{u^j\}_{j=k}$ . (Note that, since  $L$  is a strictly convex functional, (vi) is a sufficient condition for minimization.) Since  $k$  was an arbitrary

integer, the proof of the main part of the theorem follows. Expression (7.39a) is readily obtained from Lemma 7.2 under the stipulation that the optimum team trajectory is realized, since then  $J_1^*$  is identical with  $J_1$ . Expression (7.39b), on the other hand, follows from substitution of the team solution given in Lemma 7.2 into  $J_2$ . Remark 7.11 Condition 7.1, and thereby the Stackelberg solution presented in Thm. 7.3, depends, in general, on the value of the initial state  $a^k$ ; that is, the Stackelberg strategy of the leader at stage  $k$  is not only a function of  $x_k$  and  $x_{k-1}$ , but also of  $x^k$ . This latter dependence may be removed under a more stringent condition (than Condition 7.1) on the parameters of the game, in which case the vector equation (7.37) is replaced by the matrix equation  $B^{*k+1}A^{k+1}F_k + B^{2k}X^{k+1}F_k - L^{2k} = 0$  (HieK). (7.40) Such a restriction is, in fact, inevitable if we seek a solution to the infinite horizon problem (i.e., with  $K \rightarrow \infty$ ), in which case our attention is solely confined to stationary controls (see Başar and Seluz (1979b), for an illustrative example).

□ Remark 7.12 The analyses of subsections 7.4.1 and 7.4.2 have already displayed the validity of the possibility that a linear-quadratic dynamic game might not admit a Stackelberg solution in linear strategies. Therefore, depending on the parameters of the dynamic game, Condition 7.1 might not be fulfilled. In such a case, we have to adopt, instead of (7.35), a parameterized nonlinear representation of  $\{u^k; k \in K\}$  as a candidate Stackelberg strategy of the leader, and then determine those parameter values in view of Prop. 7.3. For a derivation of such a nonlinear Stackelberg solution, the reader is referred to Tolwin-ski (1981b). Another important point worth noting is that, for the linear-quadratic game of this section, and for general parameter values,  $J_1$  given in Lemma 7.2 is not always the Stackelberg cost of the leader, i.e., there may not exist a strategy  $u_1^*$  in  $\mathcal{F}_1$  (linear or otherwise) under which the follower is forced to choose a representation of  $u_2$  on the optimal team trajectory. This would be the case if, for example, there exist certain control variables of the follower, at intermediate stages, which do not affect the state variable, and thereby cannot be "detected" by the leader. In such a case, the leader cannot influence the cost function

392 T. BAŞAR AND G. J. OLSDER of the follower through these control variables and hence cannot enforce the team solution. This then necessitates derivation of a new team cost (which is realizable as the Stackelberg cost of the leader), by taking into account the robust optimal responses of the follower in terms of these controls (as we did at the final stage). The resulting

Stackelberg cost will be higher than  $J_1$ . Another such case occurs if the leader can detect (through his state observation) a linear combination of the control variables of the follower, but does not observe them separately. (This would be the case if, for example, the dimension of the state is lower than that of the control vector of the follower.) Here again, we have to determine a new minimum team cost (different from  $J_1$ ) to be realized as the Stackelberg cost of the leader, by taking into account the freedom allotted to the follower in the nondetectable (by the leader) region of his control space. We do not pursue this point any further here, and refer the reader to Tolwinski (1981b) for such a derivation, and also to Başar (1980a) and Başar (1982) for indirect derivations of the Stackelberg cost value of the leader in general dynamic games when the leader has access to closed-loop imperfect state information (cf. Def. 5.2). See also Problem 5 in Section 7.7. □

**7.4.4 Incentives (deterministic)** The idea of declaring a reward (or punishment) for a decision maker P1 according to his particular choice of action in order to induce a certain "desired" behavior on the part of another decision maker P2 is known as an incentive (or in case of the punishment, as a threat). Mathematical formulation and analysis of such decision problems bear strong connections with the theory of Stackelberg games presented in the previous subsections, which is what we will be discussing next, for deterministic scenarios.

Counterparts of these results in the stochastic case will be presented later in Section 7.5 (particularly, subsection 7.5.3). Following the convention of the previous subsections, we call, in the above scenario, P1 the leader and P2 the follower. Then, the action outcome desired by the leader is:  $(u^1, u^2) = \arg \min_{V(v^1, v^2)}$ . (7.41)  $u^1 \in S^1$  The incentive problem can now be stated as: Find a  $\gamma \in T_1$ , where  $T_1$  is an admissible subclass of all mappings from  $S^2$  into  $S^1$ , such that  $\arg \min_{L_2(\gamma(u^2), u^2)} = u^2$  (7.42a)  $u^2 \in \gamma(u^2) = u^1$  (7.42b) Note that (7.42a) and (7.42b) require choosing a set of  $m$  scalar functions which together map  $S^2$  into  $S^1$  so as to satisfy  $m_1 + m_2$  equations. If this set of  $m$  functions has  $m_1 + m_2$  or more parameters, then we might in general accomplish this by choosing the parameters appropriately. Incentive problems do arise in real life decision making. Think of P1 as a government and of P2 as a citizen. The income tax which P2 has to pay is a

**HIERARCHICAL (STACKELBERG) EQUILIBRIA** 393 fraction (say  $k$ ) of his income (before taxation)  $u^2$ . The amount of money that the government receives is  $u^1 = k u^2$ . It is up to P2 how hard to work and thus how much money



to earn. The incentive here is  $u_1 = \gamma_1(u_2) = ku_2$ . The government will choose  $k$  so as to achieve certain goals, but it cannot choose its own income,  $u_1$ , directly. In reality, the  $\gamma_1$ -functions will often be nonlinear, but that does not take away the incentive phenomenon. Example 7.3 Consider  $L_1 = (u_1)^2 + (u_2)^2$  and  $L_2 = (u_1 - 1)^2 + (u_2 - 1)^2$ , where the  $u_i$  are scalars. By inspection,  $u_1 = u_2 = 0$ . Consider the choice  $u_1 = ku_2$ , with  $k$  approaching  $\infty$  if necessary, as a possible incentive mechanism for  $P_1$ . The idea is that any choice of  $u_2 \neq 0$  will make  $L_2$  approach  $\infty$  if  $k \rightarrow \infty$  and thus force  $P_2$  to choose  $u_2$  arbitrarily close to  $u_2 = 0$  in his own interest. However, by substituting  $u_1 = ku_2$  into  $L_2$ , it is easily shown that the minimizing action is  $u_2 = (k + 1)/(k^2 + 1)$  and consequently  $u_1 = (k^2 + k)/(k^2 + 1)$ . Thus  $u_2$  approaches  $u_2 = 0$  and  $u_1$  approaches 1 (which is different from  $u_1 = 0$ ) as  $k$  approaches  $\infty$  and hence (7.42b) is violated. Consequently "infinite threat" as just described is not generally feasible. Besides, such a threat may not be credible in practice. Let us now consider an incentive  $\gamma_1$  of the form  $u_1 = \gamma_1(u_2) = u_1^t + g(u_2, u_2^t)$ , (7.43) where  $g$  is a function which satisfies  $g(u_2, u_2) = 0$ . With this restriction on  $g$ , equation (7.42b) is automatically satisfied. Let us try the linear function  $g = k(u_2 - u_2^t)$ . Equation (7.42a) reduces to  $(k + 1)/(k^2 + 1) = 0$  and hence  $k$  must be equal to  $-1$ . Graphically, the incentive  $u_1 = ku_2 = -u_2$  is a line through the team solution  $(u_1^t, u_2^t)$  and has only this point in common with the set of points  $(u_1, u_2)$  defined by  $L_2(u_1, u_2) < L_2(u_1^t, u_2^t)$ . By announcing the  $\gamma_1$ -function,  $P_1$  ensures that the solution  $(u_1, u_2)$  will lie on the line  $u_1 = -u_2$  in the  $(u_1, u_2)$  plane, independent of the action of  $P_2$ . Being rational,  $P_2$  will choose that point on this line which minimizes his cost function; such a choice is  $u_2 = 0$ . The relationship of this with the derivation of the Stackelberg solution discussed in subsection 7.4.1 and particularly with the depiction of Fig. 7.1 should be clear now. From this graphical interpretation, it should also be clear that any (nonlinear, continuous or discontinuous) incentive policy  $u_1 = \gamma_1(u_2)$  which passes through the point characterized by the team solution and has only this point in common with the set just described, will lead to the team solution for  $P_1$ . If we restrict ourselves to linear incentives, then the solution is unique in this example. This example exhibits yet another feature. Note that with  $k = -1$ ,  $L_2(ku_2, u_2) = 2(u_2)^2 + 2 = L_1(u_2) + 2$ . In other words, by this choice of incentive, the objectives of both players are identical (apart from a constant), thus fulfilling the old adage "if you wish other people to behave in your own interest, then make them see things your way".

394 T. BAŞAR AND G. J. OLSDER In general making the cost functions identical by the choice of an appropriate  $\gamma_1$ -function will be too strong a requirement. A weaker form, however, which also leads to the team solution, is  $\text{argmin}_{L^2}(\gamma_1(u_2), u_2) = \text{argmin}_{L^2}(u_1^2, u_2)$ , (7.44)  $u_1, u_2$  where it is assumed that  $\gamma_1$  is of the form as described by (7.43).  $\square$  Definition 7.1 The incentive problem, as defined in this subsection, is (linearly) incentive controllable, if there exists a (linear)  $\gamma$ -function such that (7.42a) and (7.42b) are satisfied. Of course, not all problems are incentive controllable. What can  $P_1$  achieve in problems that are not incentive controllable? The method to be employed to answer this question will be described in the context of an example, given below. It should then be clear how the method would apply to more elaborate problems.

Example 7.4 Consider  $L_1 = (u_1 - 4)^2 + (u_2 - 4)^2$  and  $L_2 = (u_1)^2 + (u_2 - 1)^2$ , where the  $u_i$  are scalars;  $u_1 \in S_1 = [0, 3]$  and  $u_2 \in S_2 = [0, 6]$ . The team solution in this case is:  $u_1 = 3, u_2 = 4$  and  $L^*(u_1, u_2) = 1$ . This is depicted in Fig. 7.2, where some contours of  $L_1$  have been drawn; here the horizontal axis stands for  $u_2$  (which is the independent variable) and the vertical one corresponds to  $u_1$ . The worst possible outcome for  $P_2$ , even if he minimizes his cost function with respect to his own decision variable, is  $\min [\max_{L^2}(u_1, u_2)]$ . This occurs for  $u_2 = 1, u_1 = 3$  (point A in the figure), and  $L_2(3, 1) = 9$ . Whatever choice  $P_1$  makes for  $\gamma_1$ , the cost for  $P_2$  will never be higher than 9. If  $P_1$  chooses  $u_1 = \gamma_1(u_2) = 3$  on the interval  $[0, 6]$ , then the outcome becomes  $u_2 = \text{argmin}_{L^2} u_1 = 3$ , (point A), and the costs for  $P_1$  and  $P_2$  become 10 and 9, respectively. This, however, is not optimal for  $P_1$ . He should instead consider  $\min_{L^1} L_1$  subject to  $L_2 < 9$ . This latter region has been shaded in the figure. The solution of this minimization problem is  $u_1 = 12/5, u_2 = 14/5$  (point B in the figure). Now, any  $\gamma_1$ -curve, in the rectangle  $0 < u_1 < 3, 0 < u_2 < 6$ , which has only the points A and B with the shaded region in common, would lead to a nonunique choice for  $P_2$ ; he might either choose  $u_2 = 1$  or  $u_2 = 14/5$ . Both choices lead to  $L_2 = 9$ . The costs for  $P_1$  are, respectively, 10 and 4 for these choices. Therefore,  $P_1$  will choose a  $\gamma_1$  function as just described, with one exception; it will have a little "dip" in the shaded area near point B, such that the choice of  $P_2$  will be unique again. (A possible choice is:  $\gamma_1(u_2) = 3$  for  $0 < u_2 < 14/5 - \epsilon$ , where  $\epsilon > 0$  and  $\gamma_1(u_2) = 12/5$  for  $14/5 - \epsilon < u_2 < 6$ .) The outcome will now be a point  $u_1, u_2^*$  near point B, just within the shaded area.  $P_1$  can keep his costs arbitrarily close to 4 (but not equal to 4).  $\square$

HIERARCHICAL (STACKELBERG) EQUILIBRIA 395 L=4 Figure 7.2:

Graphical illustration of Example 7.4. Extensions of the foregoing analysis are possible in different directions, such as the multi-stage problems (as in the case of the example of subsection 7.4.2) or problems with multiple hierarchies; see the Notes section 7.7 for selected references that cover these extensions. Another possibility is the many-followers case which we now briefly discuss. If there are two or more followers in the decision problem, the relationship (that is, the solution concept to be adopted) between the followers must be specified.<sup>77</sup> We shall illustrate a few of these, barring the formation of coalitions between followers. Let  $P_1$  be the leader,  $P_i, i = 2, \dots, M$ , be the  $M - 1$  followers. An incentive policy (equivalently, incentive mechanism) for the leader is a mapping  $\gamma_1 : S_2 \times \dots \times S_M \rightarrow S_1$ . Suppose that  $u^1, i = 1, \dots, M$ , is the  $M$ -tuple of actions desired by the leader. An incentive mechanism  $\gamma_1$  is said to induce a dominant strategy solution if  $\text{argmin}_{u^1} U\{-/X\{u^2, \dots, u^M\}, u^2, \dots, u^M\} = u^1 \setminus u^1$  with arbitrary  $u^j, \forall j \neq i, i = 2, \dots, M$ . For an illustration of a three decision maker problem, let  $u^1 = (u^2, u^3)$ , and  $L_1 = (u^2 + u^3)^2, V = \{u^i - I\}^2 + u^i, i = 2, 3$ . Then, the incentive mechanism  $u^2 = \gamma_1(u^3) = 2u^3$  will induce  $u^2 = 0$  regardless of the value of  $u^3$ , and similarly  $u^3 = \gamma_2(u^2) = 2u^2$  will induce  $u^3 = 0$  for all values of  $u^2$ , and hence the concatenation of  $\gamma_1$  and  $\gamma_2$  constitutes a dominant strategy. Such a policy is the most desirable one (for the leader), since it effectively decouples the followers from each other, and the leader can control each one's cost function separately. However, such a solution is generally difficult to realize, since the cost functions of the followers may not have the required structure. An alternative is the Nash equilibrium concept (among the followers), Recall the discussion in Section 3.6 in the context of finite games.

396 T. BAŞAR AND G. J. OLSDER where we only require  $\text{argmin}_{u^1} L_1(\gamma_1(\langle 2, \dots, \langle M \rangle, \dots, \langle M \rangle) = \hat{u}, \dots, \hat{u})$  (7-45) with  $u^i = u^i \forall j \neq i, i, j = 2, \dots, M$ , i.e., each follower will behave desirably conditioned on the fact (expectation) that the others will do the same. A particular subcase of this equilibrium occurs if the incentive mechanism can be chosen so that under it the cost functions of all followers become identical. The followers then face a team problem which has only one "reasonable" solution; see Problem 10 in Section 7.7 for an illustration of such a situation. The same problem also deals with the case of two followers facing a zero-sum game after implementation of the correct incentive mechanism. The saddle-point solution of this zerosum game leads to the team solution of the leader who thus practices the adage "divide and rule".

7.5 Stochastic Dynamic Games with Deterministic Information Patterns This section is devoted to a brief discussion on possible extensions of the results of the previous sections to stochastic dynamic games (cf. Def. 5.4) wherein the state evolves according to  $X_{k+1} = f_k(x_k, u_k) + g_k$ ,  $k \in K$ , (7.46) where  $\{g_1, \dots, g_k\}$  is a set of statistically independent Gaussian vectors with values in  $R^n$  and with  $\text{cov}(g_k) > 0$ ,  $k \in K$ . In a general context, the cost functional of  $P_1$ , for the game in extensive form, is again taken to be stage-additive, viz.  $K$   $J_1(u_1, u_2) = \sum_{i=1}^K s_i(x_{i+1}, u_i, u_i)$ ,  $i = 1, 2$ , (7.47)  $f_k = i$  and, abiding by our standard convention,  $P_1$  is taken to act as the leader and  $P_2$  as the follower. In the discussion to follow, we shall consider three different information structures; namely, (A) open-loop for both players, (B) open-loop for the leader and closed-loop perfect state (CLPS) for the follower, (C) CLPS for both players. In addition to derivation of the (global) Stackelberg solution (cf. subsection 7.5.1), we shall also discuss derivation of the feedback Stackelberg solution under the third information pattern listed above (cf. subsection 7.5.2), and stochastic incentive problems (as the counterpart of the material presented in subsection 7.4.4).

7.5.1 (Global) Stackelberg solution A. Open-loop information for both players If the underlying deterministic information structure is open-loop for both players (in which case the controls depend on the initial state  $x_0$  which is assumed to

HIERARCHICAL (STACKELBERG) EQUILIBRIA 397 be known a priori), then the Stackelberg solution can be obtained by basically converting the game into equivalent static normal form and utilizing the first method outlined in Section 7.2 for the deterministic open-loop Stackelberg solution. Toward this end, we first recursively substitute (7.46) into (7.47), and take expected values of  $J_1$  over the statistics of the random variables  $\{g_1, \dots, g_k\}$ , to obtain a cost functional  $J_1(u_1, u_2)$  which depends only on the control vectors,  $u_1, u_2$  (and also on the initial state vector  $x_0$  which is already known by the players). Since such static games have already been treated in Section 4.5, we do not discuss them here any further, with the exception of one special case (see the next paragraph and Prop. 7.4). It should be noted that the Stackelberg solution will, in general, depend on the statistical moments of the random disturbances in the state equation—unless the system equation is linear and the cost functionals are quadratic. For the special case of linear-quadratic dynamic games (i.e., when  $f_k$  in (7.46) is given as in (7.7a), and  $g_k$  in (7.47) is structured as in (7.7b), and under the open-loop information structure, the cost functional  $J_1(u_1, u_2)$  can be

written as  $K^{-1} r(u^2) = M^Q v V_{k+i} + 4*4 + u\{R\} + ?; j \pm i, j = 1,2, k=i$  where  $l_{k+i} = A_k y_k + B_k u + B_k u + E[9_k]$ ;  $V_i = *i$ , and  $\xi_a$  is independent of the controls and  $X$ , and depends only on the covariances of  $\{9_1, \dots, 9_K\}$ . This result is, of course, valid as long as the set  $\{9_1, \dots, 9_K\}$  is statistically independent of  $X$ , which was one of our underlying assumptions at the beginning. Hence, for the linear-quadratic stochastic game, the stochastic contribution completely separates out; and, furthermore, if  $E[9_k] = 0 \forall k \in K$ , the open-loop Stackelberg solution matches (completely) with the one given in Corollary 7.1 for the deterministic problem. This result is now summarized below in Prop. 7.4.

**Proposition 7.4** Consider the two-person linear-quadratic stochastic dynamic game described by (7.46)-(7.47) together with the structural assumptions (7.7a)-(7.7b), and under the parametric restrictions  $Q_{k+1} > 0, R_{k+1} > 0, E[9_k] = 0 (k \in K; i = 1,2)$ . It admits a unique open-loop Stackelberg solution with PI acting as the leader, which is given, under the condition of invertibility of the matrix appearing in the braces in (7.9c), by (7.8a)-(7.8b).

**B. OL information for PI and CLPS information for P2** For the deterministic Stackelberg game, we have already seen in Section 7.2 that an additional state information for the follower does not lead to any difference in the open-loop Stackelberg strategy of the leader, provided, of course, that he (the leader) still has only open-loop information (cf. Remark 7.1). The only

398 T. BAŞAR AND G. J. OLSDER difference between the OL-OL and OL-CLPS Stackelberg solutions then lies in the optimum response strategy of the follower, which, in the latter case, is any closed-loop representation of the open-loop response strategy on the equilibrium trajectory associated with the OL Stackelberg solution; hence, the optimum strategy of the follower under the OL-CLPS information pattern is definitely not unique, but it can, nevertheless, be obtained from the OL Stackelberg solution. For the stochastic Stackelberg game, however, the issue is more subtle, especially if the state dynamics is not linear and the cost functional is not quadratic. In general, the OL-OL Stackelberg solution does not coincide with the OL-CLPS Stackelberg solution, and the latter has to be obtained independently of the former. The steps involved in this derivation are as follows: (1) For each fixed  $\{u \in U; k \in K\}$ , minimize  $J_2(u) = E \{ \int_0^T g(x_{k+i}, u, X_k) dt \mid u = f(x_t, e^k), k \in K \}$  subject to (7.46) and over the permissible class of strategies,  $F_2$ . Any solution  $u^* \in F_2$  of this stochastic control problem satisfies the dynamic programming equation (see Section 5.6, equation (5.61b)) (7.48)  $V(k,x) = \min$

$E_k[g_1(f_k(x_k, u_1, u_2) + 9_k, u_1, u_2, x_k) + V(k+1, f_k(x_k, u_1, u_2) + 6_k)] + v(k+1, f_k(x_k, u_1, u_2) + e_k)$ , where  $r_k = \{x_i, e_k\}$ , and  $E_k$  denotes the expectation operation with respect to the statistics of  $9_k$ . (2) Now, minimize the cost functional  $J^*(u) = E[L(u) | u] = J^*(W = j^*(j^2), * \in K)$  (7.49) over the permissible class of strategies,  $T_1$ , and subject to the constraints imposed by the dynamic programming equation (7.48) and the state equation (7.46) with  $u$  replaced by  $-y^*(77^*)$ . The solution of this optimization problem constitutes the Stackelberg strategy of the leader in the stochastic dynamic game under consideration. Note that  $J^*(u)$  is, in fact, dependent on  $u$ , but this dependence cannot be written explicitly (in closed-form), unless the cost functional of  $P_2$  and the state equation have a specific simple structure (such as quadratic and linear, respectively). The underlying assumption here is that step (i) provides a unique solution  $\{u^*; k \in K\}$  for every  $\{u \in U; k \in K\}$ .

**HIERARCHICAL (STACKELBERG) EQUILIBRIA 399** For the class of linear-quadratic games, these two steps can readily be carried out, to lead to a unique Stackelberg solution in closed-form. In this case, the leader's equilibrium strategy, in fact, coincides with its counterpart in the OL-OL game (presented in Prop. 7.4)—a property that can be verified without actually going explicitly through the two steps of the foregoing procedure. This result is given below in Prop. 7.5, whose proof, as provided here, is an indirect one.

**Proposition 7.5** Consider the two-person linear-quadratic stochastic Stackelberg game of Prop. 7.4, but under the OL-CLPS information pattern. Provided that the condition of matrix invertibility of Prop. 7.4 is satisfied, this stochastic game admits a unique Stackelberg solution with  $P_1$  as the leader, which is given for  $k \in K$  by  $J^*(u) = -B_1' K_1 x$  (7.50a)  $u^*(x_k, *i) = -P_1 s_{k+1} A_k x_k - P_2 (s_{k+1} - s_{k+i} B_1 B_1' K_1 x)$ , (7.50b) where  $P^i = I + B_1 s^i B_1' - B_1'$ , (7.51a)  $s_k = Q^{-1} + A' k s_{k+1} [I - B_2 P_2 s_{k+1}] A_k$ ;  $s_{k+i} = Q_2^{k+1}$ , (7.51b)  $s_k = A' k [I - B_2 P_1 s_{k+1}] [s_{k+1} - s_{k+i} B_1 B_1' K_1 x^* k]$ ;  $s_{k+1} = 0$ , (7.51c) and  $K_k$  is defined by (7.9a), and  $x$  by (7.10e). Proof. For each fixed control strategy  $\{u_k = x^*(x^* - j_k \in K)\}$  of the leader, the dynamic programming equation (7.48) can explicitly be solved to lead to the unique solution  $J^*(u) = -P^i s_{k+1} A_k x_k - P_1 (s_{k+1} + s_{k+i} B_1 k u)$ ,  $k \in K$ , (i) where  $P_k$  and  $s_{k+i}$  are defined by (7.51a) and (7.51b), respectively, and  $s_k$  is defined by  $s_k = A' k [I - B_2 P_k s_{k+1}] [s_{k+1} + s_{k+i} B_1 k u]$ ;  $s_{k+1} = 0$ . (ii) (This result follows from Prop. 5.4 by taking  $E[6_k] = B_k u$ .) Now, at step (2) of the derivation, (i) and (ii) above will have

to be used, together with the state equation, as constraints in the minimization of (7.49) over  $T_1$ . By recursive substitution of (ii) and the state equation into (i), we obtain the following equivalent expression for (i), to be utilized as a constraint in the optimization problem faced by the leader:  $K k^{-1} 7f = 53T^*(0 \ll i + MkXl + Y, Nk(1)9e$ , (iii) where  $Tk(i)$ ,  $Mk$ ,  $Nk(i)$  are the coefficient matrices associated with this representation, whose exact expressions will not be needed in the sequel. Now, (iii)

400 T. BAŞAR AND G. J. OLSDER being a linear relation, and  $L_1$  being a quadratic expression in  $\{x^{+i}, u, u; k \in K\}$ , it follows that, with  $u_1 = (u, \dots, u)^T$ ,  $J_1(71, 720)$  can be written as  $J'_{ij1, -/20} = E\{v\}'txux + v\}'T2Xl + x\}Tixl I u = IUxi) \in K] + J_m$ , where  $T_1, T_2, T_3$  are deterministic matrices, and  $J_m$  is independent of  $u_1$  and is determined solely by the covariance matrices of  $\{9k', k \in K\}$ —this decomposition being due to the fact that the leader's information is open-loop and hence his controls are independent of the random variables  $\{9k, k \in K\}$  which were taken to have zero mean. Since the Stackelberg strategy of the leader is determined as the global minimum of  $J_1(71, 720)$ , it readily follows that it should be independent of the covariance of  $\{9k', k \in K\}$ . Now, if there were no noise in the state equation, the follower's response (though not unique) would still be expressible in the form (i), and therefore the leader's Stackelberg strategy would still be determined as the control that minimizes  $[u_1' Txux + u_1' T2X \backslash + x'1T3x \backslash]$ . Hence the leader's Stackelberg strategy in this stochastic game is the same as the one in its deterministic counterpart—thus leading to (7.50a) (see Corollary 7.1 and Remark 7.1). The follower's unique optimal response (7.50b) then follows from (i) by substituting (7.50a) into (ii). Note that this strategy is, in fact, a particular closed-loop representation of (7.8b) on the open-loop equilibrium trajectory of the deterministic problem (which is described by (7.10e)). D C. CLPS information for both players When the leader has access to dynamic information, derivation of the Stackelberg solution in stochastic dynamic games meets with insurmountable difficulties.<sup>80</sup> First, any direct approach readily fails at the outset (as in its deterministic counterpart, discussed in Section 7.4) since the optimal response of the follower to any announced strategy of the leader (from a general strategy space) cannot be expressed analytically (in terms of that strategy), even for linear-quadratic games. Second, the indirect method of Section 7.4, which is developed for deterministic dynamic games, cannot be extended to stochastic games, since, in

the latter case, every strategy has a unique representation (see Section 5.6) as opposed to the existence of infinitely many closed-loop representations of a given strategy in a deterministic system. Consequently, derivation of the closed-loop Stackelberg solution of stochastic dynamic games remains, today, as a challenge for the researchers. If, however, we make some structural assumptions on the possible strategies of the leader—which is tantamount to seeking sub-optimal solutions—then the problem may become tractable. In particular, if, under the stipulated structural assumptions, the class of permissible strategies of the leader can be described by a finite number of parameters, and if the follower's optimal response can be determined analytically as a function of these parameters, then the original game

80 Consistent with the CLPS information pattern, we are assuming here that the leader does not have direct access to the follower's control. If this is not the case, and the leader's information exhibits redundancy, then the problem might be more tractable. This case will be discussed later in subsection 7.5.3 in the context of stochastic incentive problems.

HIERARCHICAL (STACKELBERG) EQUILIBRIA 401 may be viewed as a static one in which the leader selects his strategy from a Euclidean space of appropriate dimension; such a static Stackelberg game is, in general, solvable—but more often numerically than analytically. The following example now illustrates this approach and demonstrates that even for simple stochastic dynamic games, and under the crudest type of structural assumptions, the corresponding (sub-optimal) Stackelberg solution cannot be obtained analytically, but only through some numerical minimization techniques.

Example 7.5 Consider the two-stage scalar stochastic dynamic game described by the state equations  $x_2 = \xi_1 - u_2 + 9X$ ,  $\dots$   $x_3 = x_2 - u_1 + 6z$  (7'52) and cost functional  $L_1 = (x_3)^2 + 2(u_1)^2 + (u_2)^2$ ,  $L_2 = (x_3)^2 + (u_2)^2$ . Here,  $\xi_1$  and  $z$  are taken as independent random variables with mean zero and variances  $a_1$  and  $a_2$ , respectively. The leader ( $P_1$ ) acts at stage 2 and has access to both  $x_1$  and  $x_2$ , while the follower ( $P_2$ ) acts at stage 1 and has only access to  $X$ . If  $Y \in r_i$  denotes a general strategy of  $P_i$  ( $i = 1, 2$ ), the expected (average) cost functional of  $P_2$  can be written as  $J_2(l_2) = E\{[x_2 - y_1(x_2, x_1) + z]^2 + [y_2(x_1)]^2\} = E\{[x_2 - l_1(x_2, x_1)]^2 + [l_2(x_1)]^2\} + a_2$  which has to be minimized over  $Y \in T_2$ , to determine the optimal response of  $P_2$  to  $Y \in r_1$ . Barring the stochastic aspects, this is similar to the problem treated in subsection 7.4.1, where the difficulties involved in working with a general  $Y \in r_1$  have been



delineated. We therefore now restrict our investigation (also in line with the discussion preceding the example) to a subclass of strategies in  $T_1$  which are affine in  $x_2$ , that is, to strategies of the form  $J_1(z_2, \xi_i) = ax_2 + (3x_2)$ , (7.53) where  $a$  and  $\beta$  are free parameters which are yet to be determined. They will, in general, be dependent on  $x_1$  which is, though, known a priori by both players. Under the structural restriction (7.53),  $J_2$  admits a unique minimum, thus leading to the optimal response strategy (for the follower)  $J_2^*(x_2, \xi_i)$  (7.54)  $J_2^*(x_2, \xi_i) = \frac{1}{2} \left[ \frac{1-a-\beta}{2} x_2 + \frac{a+2\beta - a\beta}{2} \right]$ . The apparent linear structure of the second term below is adopted for the sake of convenience in the analysis to follow; it could also have been taken as a single function  $\beta/3$ .

402 T. BAŞAR AND G. J. OLSDER which explicitly depends on the parameters  $a$  and  $\beta$  that characterize the leader's strategy. To determine their optimal values, we now substitute (7.53)-(7.54) into  $J_1(x_1, x_2)$ , together with the corresponding values of  $X_3$  and  $X_2$  from (7.52), to obtain the function  $F$  given below, which has to be minimized over  $a \in \mathbb{R}$ ,  $\beta \in \mathbb{R}$  for fixed  $x_1$ :  $F(a, \beta) = \frac{1}{2} \left[ \frac{1-a-\beta}{2} x_1 + \frac{a+2\beta - a\beta}{2} x_1 \right]^2$ . Let us now note that (i)  $F$  is jointly continuous in  $(a, \beta)$ ,  $F(a, \beta) > 0 \forall (a, \beta) \in \mathbb{R} \times \mathbb{R}$ , and  $F(a, \beta) \rightarrow \infty$  as  $|a|, |\beta| \rightarrow \infty$ . Therefore, we can restrict our search for a minimum on  $\mathbb{R}^2$  to a closed and bounded subset of  $\mathbb{R}^2$ , and consequently there exists (by the Weierstrass theorem) at least one pair  $(a^*, \beta^*)$  that minimizes  $F$  for any given pair  $(x_1, a)$ . (ii) The optimum pair  $(a^*, \beta^*)$  depends on  $(x_1, c_i)$ , but not on  $C_2$ , and it cannot be expressed analytically as a function of  $(x_1, a_i)$ . Hence, for each fixed  $(x_1, (T_i))$ ,  $F(a, \beta)$  has to be minimized numerically, by utilizing one of the available optimization algorithms (Luenberger, 1973). (iii) With  $(a^*, \beta^*)$  determined as above, the linear-in- $a$  strategy  $r(x_2, X) = a^*(x_1)x_2 + \beta^*(x_1)x_1$  is only a sub-optimal Stackelberg strategy for the leader, since he may possibly achieve a better performance by announcing a strategy outside the "linear-in- $a$ " class.  $\square$

### 7.5.2 Feedback Stackelberg solution

In this subsection, we extend the results of Section 7.3 to stochastic dynamic games described by (7.46)-(7.47), and, in particular, we obtain the counterpart of Thm. 7.2 in the present context (see Thm. 7.4 below). In this analysis, we do not restrict our attention at the outset to feedback strategies (as it was done in Section 7.3), but rather start with general closed-loop strategies  $u_k = J_k(x_t, \xi < k)$ ,  $k \in K$ ,  $i = 1, 2$ . The conclusion, however, is that the feedback Stackelberg solution can be realized only in feedback strategies. To see this, we start (in view of Def. 3.29) at the last stage  $k = K$  and solve basically a static Stackelberg game

between  $P_1$  and  $P_2$ , with their strategies denoted by  $J_k$  and  $j_K$ , respectively. Because of the additive nature of the cost functional of each player, and since  $\text{cov}(9_k) > 0$ ,  $9_k$  is Gaussian and statistically independent of the remaining random vectors, every Stackelberg solution at this stage will depend only on  $x_k$  and not on the past values of the state vector. Proceeding to the next stage  $k = K - 1$ , after substitution of the Stackelberg solution at stage  $k = K$  into the state equation, we conclude, by the same reasoning, that the Stackelberg strategies at stage  $k = K - 1$  are only functions of  $x_{k-1}$ . An inductive argument (as in the proof of Thm. 6.10) then verifies See Section 5 of Appendix A.

**HIERARCHICAL (STACKELBERG) EQUILIBRIA** 403 the "feedback" property of the feedback Stackelberg strategies, and it also simultaneously leads to the equations that yield the corresponding solution for the stochastic dynamic game. Theorem 7.4, below, now provides the complete solution under the assumption of singleton reaction sets for the follower. Theorem 7.4 Every feedback Stackelberg solution of a two-person  $K$ -stage stochastic infinite dynamic game described by (7.46)-(7.47) and with the CLPS information pattern (for both players) comprises only feedback strategies, and is strongly time consistent. A pair of strategies  $\{-y_1^* \in T_1^2, -y_2^* \in T_2\}$  constitutes a feedback Stackelberg solution with  $P_1$  as the leader if  $\min_{j_1} G_1(j_1, j_2, x_k) = G_1(j_1^*, j_2^*, x_k), \forall x_k \in R^n, (k \in K), \forall j_2 \in T_2, \forall j_1 \in T_1$  where  $R \setminus (-)$  is a singleton set defined by  $R \setminus (-) = \cup_{k \in K} G_2(j_1, j_2, x_k) = \min_{j_2} G_2(j_1, j_2, x_k) \forall x_k \in R^n, G_1(j_1, j_2, x_k) = E_k[G_i\{f_k(x_k, i_1(x_k), i_2(x_k)) + g_k(x_k, i_1(x_k), i_2(x_k), x_k)\}], i = 1, 2, k \in K$  and  $G_k$  is recursively defined by  $G_k(x_{k+1}, J_k(x_k), l_k(x_k), x_k) = E_k[G_{k+1}\{f_{k+1}(x_{k+1}, l_{k+1}(x_{k+1}), -y_{k+1}^*(x_{k+1})) + g_{k+1}(x_{k+1}, j_1^*(x_{k+1}), f_1^*(x_{k+1}), x_{k+1})\} + g_k(x_{k+1}, l_k(x_k), J_k(x_k), x_k)]; G_{i, K+1} = 0, i = 1, 2$ , Proof. This result follows from a direct application of Def. 3.29 and Prop. 3.15 (interpreted appropriately for the infinite stochastic dynamic game under consideration), by employing the techniques and lines of thought used in the proof of Thm. 6.10 for the feedback Nash solution. Remark 7.3 has a natural counterpart here for the stochastic game; and for the special case of the linear-quadratic stochastic dynamic game in which  $E[8_k] = 0 \forall k \in K$ , it may readily be verified that the feedback Stackelberg solution of Thm. 7.4 is precisely the one given in Corollary 7.2.

**7.5.3 Stochastic incentive problems**  
 We have seen in subsection 7.5.1 that for stochastic dynamic games, and under CLPS information, it is generally very difficult, if not impossible, to obtain the

global Stackelberg solution—with the indirect method developed in the deterministic case (cf. Section 7.4) not applicable here because of uniqueness of strategy representations. What if, however, the leader has also access to the follower's past control actions, in addition to the state information based on

404 T. BAŞAR AND G. J. OLSDER which these actions were obtained? In this enlarged information structure, a given strategy of the leader will have multiple representations, thus opening the possibility of enforcement of a team solution (to the leader's advantage) by selecting an appropriate representation of the team-optimal strategy (of the leader). To illustrate this line of thought now, let us revisit Example 7.5, but with the following enlarged information structure. Example 7.6 Consider the two-stage scalar stochastic dynamic game of Example 7.5, but with the enlarged information structure that allows  $P_i$  to have access to  $u_2$ , in addition to  $X$  and  $x_2$ . Note that since  $u_2$  depends only on  $x_i$ , this enlarged information structure carries the same statistical information as the earlier one for each fixed (pure) policy of the follower; however, as we will see shortly, the informational redundancy that it generates will bring a substantial advantage to the leader. In the spirit of the analysis of Example 7.1 for the deterministic case, let us first determine the best performance the leader would achieve if the follower were cooperating with him (in the minimization of the leader's expected cost function). The associated team problem is  $J^* = \min_{T_1, F_2} J^*(T_1, F_2)$ , where  $J^*(T_1, F_2) = E\{[x_2 - 1/x_2, x_2] + 92\}^2 + 2[11(x_2, x_1)]^2 + [72(x_i)]^2\}$ . Here, the cost shows dependence on  $u_2 = \tau^2(x)$  not only directly, but also through  $x_2$  as given by (7.52), which has to be taken into account in the minimization. Furthermore, the strategy spaces  $T_1$  and  $F_2$  are taken as in Example 7.1, since the additional knowledge of  $u_2$  for  $P_i$  does not help in further reducing the minimum team cost  $J^*$ . Now, this team problem is in fact a standard LQ stochastic control problem of the type covered by Prop. 5.4, and its solution can readily be obtained as:  $t \in T_1(x_2, x_i) = -x_2$ ,  $72(z_i) = j^* x_i$  (7.55) which is the unique minimizing pair in  $T_1 \times F_2$ . It is not, however, unique in the enlarged strategy space for the leader, as (for example) the following parameterized strategy also constitutes an optimal solution, along with  $72$  given above, for every  $a \in \mathbb{R}$ :  $1 \leq 7^*(x_2, x_i, u_2) = -x_2 + a\{u_2 - a\}$ . (7.56) This in fact characterizes the complete class of linear (in  $x_2, x, u_2$ ) optimal strategies, but of course there are also nonlinear ones—all leading to the same (minimum) expected value for the leader. By a slight generalization of the terminology introduced in Def. 5.11, we will refer to all these "minimum

expected cost

HIERARCHICAL (STACKELBERG) EQUILIBRIA 405 achieving" strategies as representations of  $\Gamma_1$  under the team-optimal solution  $(\Gamma_1 * \Gamma_2)$ . This is a rich family of strategies, among which we seek one with the additional property that if the follower instead minimizes his own expected cost function, then the strategy in  $\Gamma_2$  that achieves this minimum is still  $j_2$ . The corresponding strategy (representation) for the leader would then clearly constitute a global Stackelberg solution, leading to the best possible performance for him. Let us now conduct the search in the family of linear representations (7.56), which leads to the quadratic optimization problem:  $\min E\{[x_2 - 1x_a\{x_2, x_1, u_2\} + 92]^2 + [u_2]^2\}$ , where  $x_2 = x_1 - u_2 + 9x$ . Since  $X$  is independent of  $8$  and  $92$ , which have zero mean, this problem is equivalent to the following deterministic optimization problem:  $\min_v \{v^2 + a^{-1}(v - x)^2\}$  where we have written  $v$  for  $u_2$ , to simplify the notation. Now, a simple optimization shows that for the value of  $a = 8/27$ , this optimization problem admits the unique solution  $v = (5/14)a^{-1}(x)$ , and hence the policy pair  $\Pi(x_2, x_1, v) = -x_2 + (v - x)$ ,  $\Pi(x) = -x$  provides a global Stackelberg solution. This is in fact the unique such solution in the linear class.  $\square$  Stochastic decision problems of the type above, where the leader is allowed to have access to past actions of the follower are known as stochastic incentive problems, which are the stochastic counterparts of those briefly discussed in subsection 7.4.4. In stochastic incentive problems, the information structure may not always be nested (for the leader), as in the example above, where the leader has access to all the information that the follower has access to (plus more). If, for instance, in Example 7.6 the leader has only access to  $x_2$  and  $u_2$ , then we have a problem with a nonnested information structure, to which the methodology presented above does not apply, since the dynamic information for the leader no longer exhibits redundancy. Discussion of such problems, where the follower possesses private information not known to the leader, is beyond the scope of our coverage here; the interested reader can consult with Ho, Luh and Olsder (1982) and Başar (1984, 1989a). For stochastic incentive problems with nested information, however, the methodology used in Example 7.6 can be applied. This equivalence holds as far as its optimum solution goes (which is what we seek), but not for the corresponding minimum values.

406 T. BAŞAR AND G. J. OLSDER developed into a general procedure as

briefly discussed below for a special class of such problems. Consider a two-person stochastic incentive problem with the cost functions  $L_1(u_1, u_2; \theta)$  and  $L_2(u_1, u_2; \theta)$ , for P1 (leader) and P2 (follower), respectively, where  $\theta$  is some random vector with a known distribution function. Let  $y_1 = h_1(\theta)$  be the measurement of P1 on  $\theta$ , and  $y_2 = h_2(\theta)$  be P2's measurement, with the property that what P2 knows is also known by P1 (but not necessarily vice versa).<sup>84</sup> Let  $T_1$  be the set of all measurable policies of the form  $u_1 = \gamma_1(j/1, y_2)$ , and  $T_2$  be the set of all measurable policies of the form  $u_2 = \gamma_2(j/2, y_1)$ . Introduce the pair of policies  $(\gamma_1, \gamma_2) = \arg \min_{\gamma_1, \gamma_2} E \{ L_1(h_1(\theta), \gamma_2(h_2(\theta)), \theta) + \int L_2(\gamma_1(j/1, y_2), h_2(\theta), \theta) dP \}$  assuming that the underlying team problem admits a minimizing solution. Then, a representation of the leader's strategy  $\gamma_1$  under the pair above is an element of  $T_1$ , say  $\gamma_1$ , with the property  $\int \gamma_1(h_1(\theta), \gamma_2(h_2(\theta))) dP = \int \gamma_1(h_1(\theta)) dP$ , a.s.<sup>85</sup> (7.57) The following result now readily follows: Proposition 7.6 For the stochastic incentive decision problem with nested information as formulated above, the pair  $(\gamma_1, \gamma_2)$  constitutes a global Stackelberg solution, leading to the best possible outcome for the leader. Equivalently, if a strategy  $\gamma_1 \in T_1$  exists satisfying (7.57), the stochastic decision problem is incentive controllable.<sup>86</sup> Remark 7.13 For the special class of linear-quadratic-Gaussian (LQG) problems, where the decision variables  $(u_1, u_2)$  belong to finite dimensional Euclidean spaces,  $\theta$  is a Gaussian random vector,  $h_1$  and  $h_2$  are linear, and  $L_1$  is jointly quadratic in the triple  $\{u_1, u_2, \theta\}$  and strictly convex in  $(u_1, u_2)$  for each  $\theta$ , the team-optimal policies  $\gamma_1$  and  $\gamma_2$  exist, are unique and linear in  $y_1$  and  $y_2$ , respectively (see any standard book on stochastic control, such as (Bertsekas, 1987)). If, furthermore,  $L_2$  is also a quadratic function, then except for some isolated cases one can restrict the search to linear representations of  $\gamma_1 : \int \gamma_1(v_1, v_2) dP = \int \gamma_1(v_1) dP + P[v_2^T v_2]$ , (7-58) where  $P$  is a matrix of appropriate dimensions. Now, invoking the condition (7.57) one can obtain an equation for  $P$ , whose solution (when used in (7.58)) leads to a linear incentive policy. This then makes the decision problem linear incentive controllable; for details see (Başar, 1979d).  $\square$  <sup>84</sup>In mathematical terms, this requirement can be stated as the sigma-field generated by  $y_1$  including the sigma-field generated by  $y_2$ . <sup>85</sup>The equality should hold for almost all values of  $\theta$ , under its assumed distribution function. <sup>86</sup>The terminology we have used here is the natural counterpart (in the stochastic case) of the one introduced in Def. 7.1 for deterministic incentive problems.

above does not cover (even in the LQG framework) the most general class of dynamic nested stochastic incentive problems, because the measurements of the decision makers have been taken to be static—not depending on the past actions. If the leader's measurement at stage  $k$  depends on the past actions of the follower ( $u_2, i < k$ ), then the approach discussed above can easily be adjusted to apply to such multi-stage problems too. If, however, the follower also has access to the leader's past control actions, then because of the nestedness of the information structure for the leader (which does not allow for the follower to have access to all measurements of the leader) the associated dynamic team problem becomes what is called a nonclassical stochastic control problem, for which no general theory exists. Issues such as learning, inference, and filtering become of relevance then, whose treatment requires background in stochastic processes, information theory and control, much beyond the level of our coverage here. For static stochastic incentive problems where there is one leader and several followers, with the followers playing according to the Nash concept, a natural counterpart of (7.45) exists, which is:  $J_i(u_1, u_2, \dots, u_M) = \inf_{u_i} \int_0^T \dots$  with  $y_j = u_j \wedge i, i, j = 2, \dots, M$ . Since the Nash equilibrium is defined here over strategy spaces, it admits natural (conceptual) extensions to the multi-stage case, where the informations of the players (leader as well as the followers) might be intertwined in more complicated ways.

### 7.6 Stackelberg Solution of Differential Games

In this section we discuss the continuous-time counterparts of some of the results presented in the previous sections, and in particular the open-loop and feedback Stackelberg solutions.

#### 7.6.1 The open-loop information structure

As counterparts of the results of Section 7.2 in the continuous time, we present here the open-loop Stackelberg solution of deterministic two-person differential games of prescribed fixed duration (cf. Def. 5.5), when  $P_i$  acts as a leader. The class of games under consideration are described by the state equation  $\dot{x}(t) = f(t, x(t), u_1(t), u_2(t)); x(0) = x_0$ , (7.59a) and cost functionals  $J_i(u_1, u_2) = \int_0^T g_i(t, x(t), u_1(t), u_2(t)) dt + q_i(x(T)); i = 1, 2$ , (7.59b) Jo

408 T. BAŞAR AND G. J. OLSDER where  $[0, T]$  denotes the fixed prescribed duration of the game,  $x_0$  is the initial state known by both players,  $x(t) \in \mathbb{R}^n$  and  $u_i(t) \in C^0([0, T], \mathbb{R}^m_i)$  ( $i = 1, 2$ ),  $\forall t \in [0, T]$ . The underlying information structure is open-loop for both players, so that the controls  $u_1(-)$  and  $u_2(-)$  depend only on the time variable  $t$  and the initial state  $x_0$ . We allow only for controls (or synonymously, in this case, strategies) that are continuous in the

time variable; and this determines the strategy sets  $(r_1, r_2)$  which are in this case equivalent to the control function sets  $\{U_1, U_2\}$ . We further assume that  $f$  satisfies the conditions of Thm. 5.1 so that a unique continuously differentiable trajectory exists as a solution to (7.59a) for every permissible control pair  $(u^1, u^2)$ . To determine the set of relations to be satisfied by an open-loop Stackelberg solution, we first obtain the optimal reaction of the follower (P2) to every announced control  $u^1$  of the leader by minimizing  $L_2\{u^1, u^2\}$  over  $u^2 \in U_2$ . Lemma 7.3 In addition to the conditions of Thm. 5.1, let i)  $f\{t, x, u^1, u^2\}$  be continuously differentiable on  $R^n \forall t \in [0, T]$ , ii)  $g_2(t, x, u^1, u^2)$  and  $q_2\{x\}$  be continuously differentiable on  $R^n \forall t \in [0, T]$ . Then, if  $u^1 \in U_1$  is a fixed control of P1, and  $u^2 \in U_2$  denotes a corresponding optimal response of P2, there exists a function  $p(-) : [0, T] \rightarrow R^n$  such that the following relations are satisfied:  $\dot{x} = f(t, x(t), u^1(t), u^2(t))$ ,  $x(0) = x_0$ ,  $u^2(t) = u^2(t, p(t), x(t), u^1(t))$ ,  $H_2\{t, p(t), x(t), u^1(t), u^2(t)\}$ ,  $p'(T) = -q_2'(x(T))$ , where  $H_2(t, p, u^1, u^2) = g_2(t, x, u^1, u^2) + p'f(t, x, u^1, u^2)$ ,  $t \in [0, T]$ . If, furthermore,  $H_2\{t, p, x, u^1, u^2\}$  is continuously differentiable and strictly convex on  $S_2$  which is taken as an open set, the second relation above is replaced by  $\forall u^2 \in U_2 H_2\{t, p, x, u^1, u^2\} = 0$ . Proof. This result follows directly from Thm. 5.4. To proceed further, we now assume that there exists a unique  $u^2 \in U_2$  under which the set of relations of Lemma 7.3 is satisfied for a given  $u^1 \in U_1$ , a sufficient condition for which is strict convexity of  $L_2(u^1, u^2)$  on  $U_2$ , with  $x(-)$  substituted from (7.59a). Then, to determine his Stackelberg strategy, the leader will be faced with the optimal control problem:  $\min_{u^1} L(u^1, u^2)$  (7.61a)

HIERARCHICAL (STACKELBERG) EQUILIBRIA 409 such that  $x = f(t, x, v, u^2)$ ;  $x(0) = x_0$ , (7.61b)  $p' = -H_2(t, p, x, u^1, u^2)$ ;  $p'(T) = -q_2'(x(T))$ , (7.61c)  $\forall u^2 \in U_2 H_2(t, p, x, u^1, u^2) = 0$ . (7.61d) This optimization problem is not of the type covered by Thm. 5.4 since its differential constraints involve specified boundary conditions at both ends. The problem, however, is still tractable, and can be solved by utilizing some other standard results of optimal control theory, particularly from (Bryson and Ho, 1975, p. 65). Toward this end, we introduce the "Hamiltonian"  $H = g + p'f + \lambda'(X - \phi)$  (7.62) where  $A_1(-) : [0, T] \rightarrow R^n$  is the costate function corresponding to (7.61b),  $A_2(-) : [0, T] \rightarrow R^n$  is the costate function corresponding to (7.61c) and  $A_3(-) : [0, T] \rightarrow R^m$  is the Lagrange multiplier function associated with the equality constraint (7.61d). Under suitable differentiability conditions (to be made precise in Thm. 7.5 to follow),  $A_1(\bullet)$  and  $A_2(-)$  satisfy the set of

differential equations  $A_j = -\text{ifl}(\langle p, A_i, A_2, A_3, u_1, u_2 \rangle)$ ;  $A_i(T) = J^*(T) - \text{ifl}(\langle p, A_1, A_2, A_3, u_1, u_2 \rangle)$ ;  $A_2(0) = 0$ ; and with  $S_1$  taken as an open set, the Stackelberg open-loop control of  $P_i$  satisfies the relations  $V_{i+1} = 0$ . These results are now summarized in Thm. 7.5 below. Theorem 7.5 For the class of two-person differential games under consideration in this subsection, assume, in addition to i), ii) and  $H_i$  of Lemma 7.3, that iv)  $f(t, x, u_1, u_2)$  is twice continuously differentiable on  $R^n \times V \times \epsilon [0, T]$ , v)  $q_2(t, x)$  and  $q_2(x)$  are twice continuously differentiable on  $R^n \times V \times \epsilon [0, T]$  and  $q_1(t, x)$  are twice continuously differentiable on  $R^n \times V \times \epsilon [0, T]$ ,

410 T. BAŞAR AND G. J. OLSDER vii)  $f(t, x, u_1, u_2)$ ,  $g_i(t, x, u_1, u_2)$ ,  $i = 1, 2$ , are continuously differentiable on  $R^n \times V \times \epsilon [0, T]$ , viii)  $S_1$  is an open set. Then, if  $\{u_i^*(t, x_0) = u_i^*(t); i = 1, 2\}$  provides an open-loop Stackelberg solution with  $P_i$  as the leader, and  $\{x^*(t), 0 < t < T\}$  denotes the corresponding state trajectory, there exist continuously differentiable functions  $p(-)$ ,  $X(-)$ ,  $A_2(-)$ :  $[0, T] \rightarrow R^n$ , and a continuous function  $A_3(-)$ :  $[0, T] \rightarrow R^m$ , such that the following relations are satisfied:  $i^*(t) = f(t, x^*(t), u_1^*(t), u_2^*(t))$ ;  $x^*(0) = x$ ,  $p'(t) = -\text{ifl}(\langle p, x^*(t), u_1^*(t), u_2^*(t) \rangle)$ ;  $p'(T) = -q_2(x^*(T))$ ,  $X'(t) = -H_1(t, p, u_1^*(t), u_2^*(t), x^*(t))$ ,  $X'(T) = -q_2(x^*(T)) - X_2(T)$ , (7-63)  $A_2(t) = -\text{ifl}(\langle p, x^*(t), u_1^*(t), u_2^*(t) \rangle)$ ,  $A_2(0) = 0$ ,  $V_{i+1} H_x(t, p, u_1^*(t), u_2^*(t), x^*(t)) = 0$ ,  $V_{u_2} H_1(t, p, u_1^*(t), u_2^*(t), x^*(t)) = V_{u_2} H_2(t, p, x^*(t), u_1^*(t), u_2^*(t)) = 0$ , where  $H_2$  is defined in Lemma 7.3, and  $H_1$  is defined by (7.62). Proof. Since the optimization problem faced by the leader is (7.61a)-(7.61d), the theorem follows, in view of Lemma 7.3, from a standard optimal control result that can, for instance, be found on p. 65 of the text by Bryson and Ho (1975). D As a specific application of this theorem, we now consider the special class of linear-quadratic differential games (cf. Def. 6.5), wherein we also assume  $R_1(-) > 0$ ,  $Q^i(-) > 0$ ,  $i = 1, 2$ . The Hamiltonians  $H_1$  and  $H_2$  can be written as  $H_2 = \pm (x'Q_2x + u_2'r_2u_2 + u_1'r_2u_1) + P'(Ax + B_1u_1 + B_2u_2)$ ,  $H_1 = -(x'Q_1x + u_1'r_1u_1 + u_2'r_1u_2) + X_1'(Ax + B_1u_1 + B_2u_2) - X_2'(Q_2x + A'p) + X_3'(R_2u_2 + B_2'p)$  and thus the set of relations (7.62) reduce to (with the arguments suppressed)  $x^* = Ax^* + BV + B_2u_2^*$ ;  $x^*(0) = x_0$ , (7.64a)  $p = -Q_2x^* - A'p$ ;  $p(T) = Q_2fx^*(T)$ , (7.64b)  $Ax = -QV - A'XX + Q_2X_2$ ; (7.64c)  $A_i(T) = Q)x^*(T) - Q)u_2^*(T)$ ,  $A_2 = .4A_2 - B_2X_3$ ;  $A_2(0) = 0$ , (7.64d)  $u_1^* = -R^{-1}B^X_i$ , (7.64e)

HIERARCHICAL (STACKELBERG) EQUILIBRIA 411  $u_2 = -R^{-1}B_2'p$ ,



(7.64f)  $A_3 = -R^{-1}x^*R^{-1}$ . (7.64g) Hence, Thm. 7.5 says, for this special case, that every open-loop Stackelberg solution of the linear-quadratic differential game is determined by the set (7.64a)-(7.64g). Because of the specific structure of the problem, we can actually obtain a stronger result, which is that the linear quadratic differential game admits a unique solution, and this, in turn, implies that (7.64a)-(7.64g) admits a unique solution set. To establish this result, let us first observe that there exist, for the linear-quadratic problem, bounded linear operators  $L_1 : U_1 \rightarrow U_1$ ,  $L_2 : U_2 \rightarrow U_2$ ,  $L_{12} : U_1 \rightarrow U_2$ ,  $L_{21} : U_2 \rightarrow U_1$  with  $L_1 > 0$ ,  $L_2 > 0$  and elements  $r_i \in U_i$ ,  $r_i \in \mathbb{R}$ ,  $i = 1, 2$ , so that  $J(u_1, u_2) = \int_0^T \{ u_1^T L_1 u_1 + u_2^T L_2 u_2 + u_1^T L_{12} u_2 + u_2^T L_{21} u_1 + r_1 u_1 + r_2 u_2 \} dt$ , where  $(u, v) = \int_0^T u^T(t)v(t)dt$  denotes the inner product of two vectors in  $U_i$  which is assumed to be structured as a complete vector space. Now, by hypothesis,  $J(u_1, u_2)$  is quadratic and strictly convex on  $U_1 \times U_2$ , and therefore it follows from Prop. 4.7, together with Remark 4.8 (interpreted in infinite dimensional spaces), that the linear-quadratic Stackelberg game under consideration admits a unique solution given by  $u_1^* = -L_1^{-1}(L_{12}u_2^* + r_1)$ ,  $u_2^* = -L_2^{-1}(L_{21}u_1^* + r_2)$ , where  $L = L_1 + L_{12}L_2^{-1}L_{21} > 0$ . This solution, written in operator form, should clearly correspond to the one given by (7.64a)-(7.64g), and this can also be verified directly (see Simaan and Cruz (1973a)). Finally, let us note that, since the solutions of the coupled set of differential equations (7.64a)-(7.64d) depend linearly on  $x_0$ , and since  $x^*(t)$  can be recovered The analysis to follow is in the same spirit as the proof of Thm. 6.14, and requires some prior knowledge of functional analysis. Our treatment here is rather informal and without details; for a more complete version the reader is referred to Simaan and Cruz (1973a). Furthermore, for the terminology and notation used here, the reader is referred to the proof of Thm. 6.14.

412 T. BAŞAR AND G. J. OLSDER from  $x^*(s)$  by linear invertible transformation for all  $t, s > 0$ , the set (7.64a)-(7.64d) can be replaced by a set of matrix differential equations independent of  $x_0$ , which are obtained by letting  $\dot{x}(t) = A(t)x^*(t)$ ,  $\dot{A}_2(t) = A_2(t)x^*(t)$ ,  $\dot{p}(t) = P(t)x^*(t)$ , where  $A$ ,  $A_2$  and  $P$  are the corresponding matrices. This set of coupled matrix differential equations is  $\dot{P} + P(A - B^{-1}R^{-1}B^T A^T - B_2^{-1}R_2^{-1}B_2^T P) + A^T P + Q = 0$ ;  $P(T) = Q$ ,  $Ax + A_1(A - B_1^{-1}R_1^{-1}B_1^T - B_2^{-1}R_2^{-1}B_2^T P) + A^T A + Q_1 - Q_2 A = 0$ ;

$A_1(T) = Q_1 - Q_2 A_2(T)$ ,  $U^b \& j A_2 + A_2(A - B_1^{11} - 1^{1^1} - B_2 R_{22} - 1 B_2' P) - A A_2 + B_2 R_{22} - 1 R_{12} B_2' P - B_2 R_{22} - 1 B_2' A_1 = 0$ ;  $A_2(0) = 0$ . The following theorem now summarizes the result.

**Theorem 7.6** The two-person linear-quadratic differential game (cf. Def. 6.5) characterized by the additional parametric restrictions  $R_{12} \{-\} > 0$ ,  $Q_1 > 0$ ,  $Q_2 > 0$ ,  $i = 1, 2$ , admits a unique open-loop Stackelberg solution with  $P_i$  acting as the leader, which is given by  $y_u\{t, x_0\} = -R^{-1} \{t\} m^x \{t\}$ ,  $j_2^*(t, x_0) = -R_{22}^{-1}(t) B_2(t) P \{t\} x^* \{t\}$ , where  $P$  and  $A$  are uniquely determined from (7.65), and  $x^*(-)$  satisfies the differential equation  $\dot{x}^*(t) = (A - B_1^{11} - 1^{1^1} - B_2 R_{22} - 1 B_2' p) x^*(t)$ ;  $x^*(0) = x_0$ . Proof. It follows from the preceding discussion.

**Remark 7.14** All results of this subsection (with the exception of uniqueness of follower response strategy) are valid if the follower is allowed to have access to CLPS information, and the discussion of Remark 7.1 has a natural counterpart here.

**7.6.2 The CLPS information pattern** In discrete-time dynamic games (see Sections 7.3, 7.4), the CLPS information (for the leader) has led to two types of equilibria—the global Stackelberg and feedback Stackelberg solutions. In the continuous time, we have direct counterparts of these, which we briefly discuss below.

**Global Stackelberg solution** For the global Stackelberg solution, a direct approach is again quite unwieldy, since the optimum response of the follower to an arbitrary CLPS strategy of

**HIERARCHICAL (STACKELBERG) EQUILIBRIA** 413 the leader cannot be obtained using standard methods of optimal control—the major difficulty stemming from the fact that a general control strategy for the leader will also incorporate memory, and may even be discontinuous. To circumvent this difficulty two alternatives exist. One alternative is to make some structural assumption on the permissible strategies of the leader (such as, taking  $\gamma_1$  to be a general CLNM policy, or one that is parameterized in terms of its dependence on the current and past values of the state), which then would make it possible to write down a set of necessary (and/or sufficient) conditions for the follower's reaction function to satisfy. Then, the leader will have to solve a dynamic optimization problem over an infinite dimensional strategy space subject to these infinite dimensional dynamic constraints. A theory for such nonclassical control problems can be found in Papavassilopoulos and Cruz (1979a), which is also applied there to Stackelberg problems. If, on the other hand, the strategy set of the leader is parameterized (say, by a finite dimensional vector), then the leader will have to solve an optimization

problem on the corresponding parameter space, which might be more feasible<sup>88</sup> but definitely sub-optimal. The second alternative approach to the global Stackelberg solution is the counterpart of the methodology discussed in Section 7.4, which is to find an equivalent team problem the global solution of which forms a tight lower bound on the leader's cost, and then to obtain a particular representation of the team solution (for the leader) on the optimum team trajectory that achieves this bound. This methodology has been applied in (Başar and Olsder, 1980b) to linear-quadratic differential games, and conditions have been obtained for a finite dimensional linear representation (i.e., a linear dynamic compensator) to provide a global Stackelberg solution. Counterparts of these results when the leader has access to sampled state information has been presented in (Başar, 1981b); here representation of the leader's team strategy will have to meet the additional side condition that it can depend only on the current and past sampled values of the state. Details of these derivations can be found in the references cited. Feedback Stackelberg solution For the feedback Stackelberg solution, we have to extend the definition from discrete to continuous time. Let us recall that in discrete time a feedback Stackelberg solution is one that retains the Stackelberg property at every stage— with the leader having only stagewise advantage over the follower. The continuous-time problem can be viewed as the limit of the discrete-time game as the number of stages becomes unbounded in a finite interval, which means that two consecutive decision points get arbitrarily close to each other. Hence, in a continuous-time dynamic game stagewise advantage of the leader (on the follower) turns into instantaneous advantage. Formally, the feedback <sup>88</sup>This has been demonstrated by Medanic (1977) for linear-quadratic games by taking the initial state uniformly distributed on the unit sphere.

414 T. BAŞAR AND G. J. OLSDER Stackelberg solution in a differential game can be obtained as the limit of the feedback Stackelberg solutions of a sequence of discrete-time dynamic games, each one obtained by time-discretization of the original differential game, with the  $k + 1$ st game in the sequence corresponding to a finer discretization (sampling) than the  $k$ th one. One possible construction would be to divide the finite interval  $[0, T]$  into  $k$  uniform subintervals, and to assume that over the  $l$ th subinterval  $[j(l-1), jt)$ , where  $0 < l < k$ , the players' policies depend only on the values of state prior to and including  $t - \Delta t$ , and that the leader can enforce his policy on the

follower over each such interval—thus denning the  $k$ th game in the sequence. If necessary, one can take a subsequence of these games, corresponding to  $k = 1, 2, 4, 8, \dots$ , so that the information set of each time-discretized game is strictly included in the information set of the next such game. An appropriate terminology to use for these games is "sampled-state", as the only time-discretization is in the information set, which comprises the state vector. The feedback Stackelberg solution of each state-sampled game can be obtained by again using a dynamic programming-type argument, by solving a sequence of open-loop Stackelberg games of the type covered in Section 7.2, with the one on the time interval  $[j\Delta, j(t+1))$  having the initial state  $x(j\Delta)$ . The Stackelberg costs associated with this open-loop game (which will be expressed in terms of  $x(j\Delta)$ ), will constitute the terminal state cost part of the cost functions of the respective players in the open-loop game denned on the next subinterval (in retrograde time)  $[(j-1)\Delta, j\Delta)$ . A complete analysis of the convergence of these solutions as  $k \rightarrow \infty$  is beyond the scope of our coverage here; but, if a limit exists, then the limiting solution should involve solutions of a sequence of open-loop Stackelberg games, each one denned on an infinitesimally small subinterval, which means that we now have to obtain Stackelberg solutions based on incremental costs at each time  $t$ . If  $V^i(t, x)$  denotes the feedback Stackelberg cost-to-go of  $P_i$  at time  $t$ , at state  $x$ , then the counterpart of the pair of Nash equilibrium PDEs (6.52) are in this case (for the game described by (7.59a)-(7.59b))

$$\frac{dV^i(t, x)}{dt} + f(t, x, u) + g(t, x, u) + \frac{1}{2} \frac{d^2 V^i(t, x)}{dx^2} = 0 \quad (7.66)$$

where "sol" stands for the static Stackelberg solution with  $P_i$  as the leader. By analogy with the discrete-time case, we will call any set of policies obtained from (7.66) the continuous-time feedback Stackelberg solution, which is clearly strongly time consistent (by definition). To bring (7.66) to a more explicit form, let us introduce the instantaneous

**HIERARCHICAL (STACKELBERG) EQUILIBRIA** 415 reaction function of  $P_2$  by

$$r_2(t, x; u_1) = \arg \min_{u_2} \{ f(t, x, u_1, u_2) + g_2(t, x, u_1, u_2) \} \quad (7.67)$$

Substituting this into the RHS of the first line of (7.66) and minimizing the resulting expression over  $u_1$  yields the instantaneous Stackelberg solution for the leader:

$$r_1(t, x) = \arg \min_{u_1} \{ f(t, x, u_1, r_2(t, x; u_1)) + g_1(t, x, u_1, r_2(t, x; u_1)) \} \quad (7.68a)$$

with the corresponding one for the follower being

$$r_2^*(t, x) = \arg \min_{u_2} \{ f(t, x, r_1(t, x), u_2) + g_2(t, x, r_1(t, x), u_2) \} \quad (7.68b)$$

Using these, (7.66) can equivalently be written as

... , ... NN gt ' = gx 'ffax,-,1 {t,xW (t,x)) +gi(t,xn1'(t,x)n2'(t,x)) (7.69) V1{T,x) = q(x), t = 1,2, which is a pair of coupled PDEs, whose solution yields the feedback Stackelberg cost-to-go pair. Hence, we have the following.

**Proposition 7.7** For the differential game described by (7.59a)-(7.59b), and under the CLPS information pattern, the pair of policies (7.68a)-(7.68b) constitutes a feedback Stackelberg solution, where  $V_i(t,x)$ , satisfying (7.69), is the corresponding cost-to-go function for  $P_i$  ( $i = 1,2$ ). Since the asymmetry in the roles of the players in a continuous-time feedback Stackelberg solution is only incremental, one may be led to the conclusion that the feedback Stackelberg solution should coincide with (or be very close to) the feedback Nash solution (cf. Thm. 6.16). This, however, is not necessarily the case as the feedback Nash solution corresponds to choosing the static sol operator in (7.66) as Nash equilibrium, whereas in the present case it is the Stackelberg equilibrium, and the two are not generally the same. To illustrate this point, as well as the possibility that they may sometimes be equivalent, we consider in the sequel a class of linear-quadratic games with a coupling term between the controls of the players:  $\dot{x} = Ax + B_1u_1 + B_2u_2$ , (7.70a)  $g^* = \int_0^T (x'Qix + u_1'u_1 + \sum_{i,j=1,2} R_{ij}u_j) dt$ , (7.70b)  $q_i = \int_0^T (x'Q)x dt$ ,  $S_i = R_m \setminus \{i\}$ ,  $i = 1,2$ . (7.70c)

416 T. BAŞAR AND G. J. OLSDER Then, the instantaneous optimum response function of the follower is  $u_2^*(t,x) = -R_2^{-1}(B_2' + K_2(dV_2/dx))'$  leading to the following structure for (7.68a)-(7.68b):  $\dot{V}_1(t,x) = -(A - J_1^* K_1^{-1} J_1^* - \sum_{i=1}^2 K_i^{-1} J_i^* K_i^{-1} J_i^*) V_1 + (B_1' + \sum_{i=1}^2 K_i^{-1} J_i^* K_i^{-1} J_i^* B_i)' V_1 + (B_1' + \sum_{i=1}^2 K_i^{-1} J_i^* K_i^{-1} J_i^* B_i)' (dV_2/dx)'$ , (7.71a)  $\dot{V}_2(t,x) = -R_2^{-1}(B_2' + K_2(dV_2/dx))' - L_1(dV_1/dx)' - L_2(dV_2/dx)'$ . (7.71b) The PDEs (7.69) then become  $dV_1/dx = Ax + B_1 u_1 + B_2 u_2 + W_1 V_1 + \sum_{i=1}^2 K_i^{-1} J_i^* K_i^{-1} J_i^* V_1 + K_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + J_1^* Ax + B_1^* T$ ,  $dV_2/dx = (dV_2/dx) L_1 + L_2(dV_2/dx) + R_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} K_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} K_2^{-1} J_2^* K_2^{-1} J_2^* V_2$ , (7.72a)  $\dot{x} = Ax + B_1 u_1 + B_2 u_2 + W_1 V_1 + \sum_{i=1}^2 K_i^{-1} J_i^* K_i^{-1} J_i^* V_1 + K_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + J_1^* Ax + B_1^* T$ ,  $dV_2/dx = (dV_2/dx) L_1 + L_2(dV_2/dx) + R_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} K_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} K_2^{-1} J_2^* K_2^{-1} J_2^* V_2$ , (7.72b)  $\dot{x} = Ax + B_1 u_1 + B_2 u_2 + W_1 V_1 + \sum_{i=1}^2 K_i^{-1} J_i^* K_i^{-1} J_i^* V_1 + K_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + J_1^* Ax + B_1^* T$ ,  $dV_2/dx = (dV_2/dx) L_1 + L_2(dV_2/dx) + R_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} K_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} J_2^* K_2^{-1} J_2^* V_2 + R_2^{-1} K_2^{-1} J_2^* K_2^{-1} J_2^* V_2$ , (7.72c)  $B_1^* = \{B_1 K_1 + B_2 L_1\}$ ,  $B_2^* = -(B_1 K_2 + B_2 L_2)$ . (7.73) For the feedback Nash equilibrium solution, however, the relevant set of PDEs is in the same form as (7.72a)-(7.72c) but with  $K_i$ ,  $K_2$ ,  $L_1$ ,  $L_2$ ,  $B_1$  and  $B_2$ , respectively, replaced by the "hat'ted" quantities  $\hat{K}_i = (I - R_u R^y B_1')$ ,  $\hat{K}_2 = (I - R_1 R_2 R_2^{-1})^{-1} R_1 R_2 B_2'$ ,  $\hat{L}_1 = (I - R_2 R_1 R_1^{-1})^{-1} R_2 B_1'$ ,  $\hat{L}_2 = (I - R_2 R_1 R_2^{-1})^{-1} R_2 B_2'$ ,  $\hat{S}_1 = B_1 K_1 + B_2 L_1$ ,  $\hat{B}_2 = B_1 K_2 + B_2 L_2$ .

HIERARCHICAL (STACKELBERG) EQUILIBRIA 417 Note that if the cross

terms in the cost functions are absent (i.e.,  $R_{12} = 0, R_{21} = 0$ ), we have the simple relations  $K_1 = K_1 = B_1'$ ,  $K_2 = K_2 = 0$ ,  $L_1 = -L_2 = 0$ ,  $L_2 = L_2 = B$  which imply that the two sets of PDEs become identical, thus admitting the same set of solutions. If the cross terms are not absent, however, the two sets of PDEs are intrinsically different and admit different sets of solutions. Hence, even in linear-quadratic games with generalized quadratic cost functionals, the feedback Stackelberg and Nash solutions may be different. We now conclude this subsection by reporting a result on the existence and structure of the feedback Stackelberg solution of the linear-quadratic differential game described by (7.70a)-(7.70c). Proposition 7.8 If  $T$  is sufficiently small and the matrix inverse in (7.11a) exists, the linear-quadratic differential game described by (7.70a)-(7.70c), and with  $P_1$  as the leader, admits a feedback Stackelberg solution given by  $\tilde{v}(t,x) = -(K_1 P_1 + K_2 P_2)x$ , (7.74a)  $\tilde{J}_2^*(t,x) = -(L_1 P_1 + L_2 P_2)x$ , (7.74b) where  $\{P_1(t), P_2(t)\}$  are symmetric solutions of the set of coupled matrix Riccati differential equations  $A = -PXL - L'P_1 - Q_1 - (K_1 P_1 + K_2 P_2)'(K_1 P_1 + K_2 P_2) + (L_1 P_1 + L_2 P_2)'R^{-1}2\{KXPX + K_2 P_2\}$  (7.75a)  $+(L_1 P_1 + K_1 P_1 + K_2 P_2)'R^{-1}2\{L_1 P_1 + L_2 P_2\}; P_2(T) = Q_2 - (L_1 P_1 + L_2 P_2)'(L_1 P_1 + L_2 P_2) + (L_1 P_1 + L_2 P_2)'R^{-1}21(K_1 P_1 + K_2 P_2)$  (7.75b)  $+(K_1 P_1 + K_2 P_2)'R^{-1}21\{LXPX + L_2 P_2\}; P_2(T) = Q_2, L = A + BXPX + B_2 P_2$ . (7.75c) Proof This result follows by substituting  $V = (\sqrt{2})x'P_1 x$  into the pair of PDEs (7.72a)-(7.72b) and observing that (7.75a)-(7.75b) imply satisfaction of (7.72a)-(7.72b) by such quadratic cost-to-go functions. Existence of a (unique) solution to (7.75a)-(7.75b) when  $T$  is sufficiently small follows from a standard property of ordinary differential equations with continuous right-hand sides. D Remark 7.15 Proposition 5.1 has a natural counterpart in the context of feedback Nash equilibria, simply with  $K^* L_1$  and  $B_t$  replaced by their corresponding "hatted" versions introduced earlier. The solution to the resulting set of Riccati equations will not be the same as the solution to (7.75a)-(7.75b), unless  $R_{12} = 0, R_{21} = 0$ . For the latter case, these equations are equivalent to the two-player version of (6.17a) with  $R_{12} = R_{21} = 0$ .  $\square$

418 T. BAŞAR AND G. J. OLSDER 7.7 Problems 1. Show that, for the special case of two-person zero-sum dynamic games (i.e., with  $L_1 = -L_2$ ), Thm. 7.1 coincides with Thm. 6.3, and Corollary 7.1 coincides with Thm. 6.4. 2. Obtain the open-loop Stackelberg solution of the two-stage dynamic game described by (7.23), (7.24a) and (7.24b) when (i)  $P_1$  is the leader, (ii)  $P_2$  is the leader. For what values of the parameter  $\theta$  is the open-loop Stackelberg solution

concurrent? (See Section 4.5 for terminology.) 3. Obtain the feedback Stackelberg solution of the dynamic game of the previous problem under the CLPS information pattern when (i)  $P_1$  is the leader, (ii)  $P_2$  is the leader. 4. Consider the scalar  $K$ -stage dynamic game described by the state equation  $x_{k+1} = 0.5x_k + u_k - 2u_k$ ,  $x_1 = 1$ , and cost functional  $J_1 = \sum_{k=1}^K (u_k^2 + 2(t^*)^2)$ ,  $J_2 = \sum_{k=1}^K (u_k^2 + 0.5(u_k^2) + 1.5(u_k^2))$ . Obtain the global and feedback Stackelberg solutions of this dynamic game under CLPS information pattern when (i)  $P_1$  is the leader, (ii)  $P_2$  is the leader. Compare the realized values of the cost functions of the players under these different Stackelberg solutions for the cases when  $K = 5$  and  $J \rightarrow \infty$ . 5. Consider the two-stage dynamic game characterized by the two-dimensional state equation  $x_2 = x_1 - u_2$ ,  $x_1 = (1, 1)'$ ,  $x_3 = -(M)''1$  and cost functional  $J = x_1 x_3 + (u_1^2) + (u_2^2)$ . Here,  $u_1$  is the scalar control variable of  $P_1$  (the leader) who has access to  $V = (1, 1)^*2$ ,

HIERARCHICAL (STACKELBERG) EQUILIBRIA 419 and  $u_2$  is the two-dimensional control vector of  $P_2$  (the follower) who has open-loop information. Show that the global minimum of  $J_1$  cannot constitute a tight lower bound on the leader's Stackelberg cost function and a tight lower bound can be obtained by first minimizing  $J_2$  over  $u_2 \in \mathbb{R}$  subject to the constraint  $2 - \{ \setminus \setminus \} u_2 = y$  and then minimizing  $J_1$  over  $u_1 \in \mathbb{R}$  and  $y \in \mathbb{R}$ , with  $u_2$  determined as above. (For more details on this problem and the method of solution, see Başar (1980a, 1982).) 6. Consider the scalar dynamic game of subsection 7.4.1 with  $X \setminus = 1$ ,  $0 = 2$ , and under the additional restriction  $-1 < u^* < 0$ ,  $i = 1, 2$ . Show (by graphical display) that the Stackelberg cost of the leader is higher than the global minimum value of his cost function, and a strategy that leads to this cost value (which is  $J^* = 8/9$ ) is  $y \setminus = 1/3$ . 7. Prove Prop. 7.4 by carrying out the two steps preceding it in subsection 7.5.1. 8. Consider the scalar stochastic two-stage dynamic game described by the state equation  $x_{k+1} = 0.5x_k + u_k - 2u_k + 9k$ ,  $x_1 = 1$  and cost functional  $J_1 = (u_1^2) + \sum_{k=1}^K (u_k^2 + 2(u_k^2))$ ,  $J_2 = (u_1^2) + 5 \sum_{k=1}^K (u_k^2) + 0.5(u_k^2)$ , where  $\epsilon_k$  and  $\delta_k$  are statistically independent Gaussian variables with mean zero and equal variance 1. The underlying information structure is CLPS for both players. Determine the feedback Stackelberg solution and the linear (global) Stackelberg solution with  $P_1$  as the leader. Can the leader improve his performance in the latter case by employing nonlinear strategies? 9. Obtain the linear (global) Stackelberg solution of the previous problem when the stochastic terms are correlated so that  $E[\epsilon_k \delta_k] = 1$  (in other words, there exists a single Gaussian variable  $\eta$  with

mean zero and variance 1, so that  $\mathbb{E}X = \mathbb{E}2 = 9$ ). 10. Consider the three-player incentive problem where  $P_1$  is the leader and  $P_i, i = 2,3$ , are the followers. The controls are  $u_1 = (u, ul)$  for the leader

420 T. BAŞAR AND G. J. OLSDER and  $u_i$  (scalar quantities) for  $P_i, i = 2,3$ . The cost functions are  $L_1 = (u)^2 + (ul)^2 + (u_2)^2 + (u_3)^2$ ;  $L_2 = (u - 3v)^2 + \{u_2 - 1\}^2 + (u_3 - 1)^2$ ;  $L_3 = (u + ul + (u_2 + 1)^2 + (u_3 + 1)^2$ . Determine an incentive mechanism  $u_j = \gamma, 1(u_2, u_3), i = 1,2$ , such that the two followers face a team problem; i.e., after substitution of the incentive policy into the cost functions of the followers, these cost functions become identical. Moreover, the solution of this resulting "team" problem of the two followers must lead to the team solution of the leader, i.e.,  $u = ul = u_2 = u_3 = 0$ . Next determine another incentive mechanism such that the followers face a zero-sum game and such that the saddle-point solution of this game leads to the team solution of the leader. 11. Consider the two-person incentive problem with  $P_1$  the leader and  $P_2$  the follower. The decision variables are scalars and the cost functions are:  $L_1 = (u_1)^2 + (u_2)^2$ ;  $L_2 = (u_1 - \theta)^2 + (u_2 + 1)^2$ , where  $\theta$  is a zero-mean Gaussian random variable with unit covariance. The value of  $\theta$  is known to  $P_2$  only, i.e., the information is nonnested. Show that any continuous and deterministic incentive mechanism  $\gamma_1$  will not lead to  $L_1 = 0$  (depict the contours for the possible  $L_2$ -functions). If, however, the leader applies the following incentive mechanism in mixed strategies:  $\{0$  w.p. 1 for  $u_2 > 0$ ;  $-\theta u_2$  w.p. 0.5 for  $u_2 < 0$ ;  $+\theta u_2$  w.p. 0.5 for  $u_2 < 0$ , where  $A$  is a constant larger than 1, show that it indeed leads to achievement of his optimum team cost. 12. Extend the result of Thm. 7.5 to generalized linear-quadratic differential games, where  $\dot{x} = Ax + BV + B^2u + c, i = 1,2$ ;  $Q_i = -(\dot{x}^T Q_i x + \dot{v}^T [L_i u_i + Y^T x]), i = 1,2$ ;  $j = i$ ;  $Q_i = \dot{x}^T Q_i x, S_i = R_i m_i, i = 1,2$ . Here  $c(-)$  is a continuous vector-valued function of dimension  $n$ , and the entries of all matrices are continuous in  $t$ . What conditions on the weighting matrices in the cost functions will ensure the existence and uniqueness of open-loop Stackelberg solution?

HIERARCHICAL (STACKELBERG) EQUILIBRIA 421 13. Obtain the counterpart of Prop. 7.7 for a three-player differential game with CLPS information pattern (with general state dynamics, and general cost functions), where  $P_1$  is the leader, and  $P_2$  and  $P_3$  are followers who play according to the feedback Nash equilibrium for every announced strategy of the leader. Furthermore, the equilibrium solution concept adopted between the two levels



of the hierarchy is the feedback Stackelberg. Under what conditions on the state dynamics and/or the cost functions would the equilibrium strategies of the players (obtained under this two-level hierarchy) also constitute a Nash equilibrium solution for the three-player differential game (without any decision hierarchy)?

7.8 Notes Sections 7.2 and 7.3. The open-loop Stackelberg solution concept in infinite dynamic games was first treated in the continuous time in the works of Chen and Cruz (1972) and Simaan and Cruz (1973a,b), where the latter has also introduced the feedback Stackelberg concept in the context of finite and discrete-time infinite dynamic games. Some other references that discuss the open-loop and feedback Stackelberg solutions in discrete-time infinite dynamic games are Kydland (1975), Hamalainen (1976), Başar (1979b) and Cruz (1978), where the latter also discusses derivation of feedback Stackelberg solution of continuous-time dynamic games under sampled-data state information. For applications in microeconomics, see Okuguchi (1976).

Section 7.4. Derivation of the global Stackelberg solution in infinite dynamic games with CLPS information has remained a challenge for a long time, because of the difficulties explained in this section prior to subsection 7.4.1 and illustrated in subsection 7.4.1. The indirect approach presented in this section was first introduced in Başar and Selbuz (1979a,b) and later in Tolwinski (1981b) and Başar (1980a); the specific examples included in subsections 7.4.1 and 7.4.2 are taken from Başar (1980a, 1982). This approach was extended later to more general (infinite dimensional Hilbert and Banach) spaces in (Zheng and Başar, 1982). The analysis of subsection 7.4.3 follows closely the one of Başar and Selbuz (1979a) which also includes an extension to  $N(> 2)$ -person games with one leader and  $N - 1$  followers, and with the followers playing according to the Nash equilibrium solution concept among themselves. Extensions of this approach to other types of  $N(> 2)$ -person dynamic games (with different types of hierarchy in decision making) can be found in Başar (1981a). Other representative articles devoted to this topic are Başar (1981b), Olsder (1977a), Ho, Luh and Olsder (1982) and Tolwinski (1980, 1981a). Subsection 7.4.4 on incentives contains material from Ho, Luh and Olsder (1982); see this reference, as well as Ho and Olsder (1981) and Zheng, Başar and Cruz (1984) for results on multi-stage decision problems, the nonnested case, and problems with multiple hierarchies. The method presented in Example 7.4 as to what the leader can achieve if his team solution is beyond reach is due to Tolwinski. The fact that the optimum incentive mechanism is generally nonunique has opened the possibilities for

further refinement, by choosing the one that is least sensitive to changes in the values of some parameters defining the game. Some results along this direction have been presented in (Cansever and Başar. 1983).

422 T. BAŞAR AND G. J. OLSDER Section 7.5. This section follows the lines of Başar (1979b) and extends some of the results given there. The difficulties illustrated by Example 7.5 were first pointed out in Castanon (1976). For a discussion of the derivation of the Stackelberg solution when players have access to noisy (but redundant) state information, see Başar (1979a); this is related to the topic of stochastic incentives as discussed in subsection 7.5.3. Furthermore, Başar (1979c) discusses the feedback Stackelberg solution in linear-quadratic stochastic dynamic games with noisy observation. A general existence theory for stochastic incentive problems, with nested information, can be found in Başar (1984), which was later extended to problems with multiple levels of hierarchy in Başar (1989a). Cansever and Başar (1985a,b) discuss a further refinement of the notion of optimum schemes in stochastic incentive problems based on insensitivity to system parameters. Section 7.6. The open-loop Stackelberg solution to two-player differential games was first studied by Chen and Cruz (1972), and subsequently refined by Simaan and Cruz (1973a). Wishart and Olsder (1979) have shown (based on necessary conditions) that the open-loop Stackelberg solution of some simple investment problems could be discontinuous. Details of the indirect approach discussed in subsection 7.6.2 for the global Stackelberg solution of differential games with CLPS information pattern can be found in Başar and Olsder (1980b) and Papavassilopoulos and Cruz (1980). The material on the continuous-time feedback Stackelberg solution in this subsection is from Başar and Haurie (1984). This reference, besides developing a theory for the feedback Stackelberg solution in differential games, has also introduced and analyzed a general framework for differential games where the role of leadership, and the nature of the solution concept to be adopted by the players change as the game evolves, with the change being prompted by the past policy choices of the players, as well as the outcome of a chance mechanism (modeled as a jump process). An application of the continuous-time feedback Stackelberg solution in economics can be found in (Başar, Haurie and Ricci, 1985).

Chapter 8 Pursuit-Evasion Games 8.1 Introduction This chapter deals with two-person deterministic zero-sum differential games with variable terminal

time. We have already discussed zero-sum differential games in Chapter 6, but as a special case of nonzero-sum games and with fixed terminal time. The class of problems to be treated in this chapter, however, are of the pursuit-evasion type for which the duration of the game is not fixed. Actually, it was through this type of problems (i.e., through the study of pursuit and evasion between two objects moving according to simple kinematic laws) that the theory of differential games was started in the early 1950s. Extensions to nonzero-sum dynamic games, as treated in Chapters 6 and 7, were subsequently considered in the late 1960s. Section 8.2 discusses the necessary and sufficient conditions for existence of saddle-point equilibrium strategies. Sufficiency conditions are provided by a natural two-person extension of the Hamilton-Jacobi-Bellman equation, which is called the "Isaacs equation", after Isaacs—the acknowledged father of pursuit-evasion games. A geometric derivation of the Isaacs equation is given, which utilizes the principle of dynamic programming, and by means of which the concept of semipermeable surfaces is introduced; such surfaces play an important role in the remaining sections of the chapter. Subsequently, necessary conditions are derived, which form the two-person extension of the Pontryagin minimum principle (cf. Section 5.5). These conditions are valid not for feedback strategies but for their open-loop representations, and in order to obtain the feedback strategies these open-loop solutions have to be synthesised. We also briefly discuss upper and lower value functions (which are important if the so-called Isaacs condition does not hold), and their relation with viscosity solutions. In Section 8.3, we treat "capturability", which addresses the question of whether the pursuer can "catch" the evader or not. The answer to this is completely determined by the kinematics, initial conditions and the target set; the cost function does not play any role here. It is in this section that the 423

424 T. BAŞAR AND G. J. OLSDER homicidal chauffeur game and the two-cars model are introduced. In Section 8.4, we return to the derivation of saddle-point strategies for quantitative games. What makes the application of necessary and sufficient conditions intellectually interesting and challenging is the presence of a large number of singular surfaces. These surfaces are manifolds in the state space across which the backward solution process of dynamic programming fails because of discontinuities in the value function or in its derivatives. The complexity that arises in the derivation of the solution of a differential game due to the presence of singular surfaces is best illustrated

by the game of the "lady in the lake," which is the main topic of Section 8.5. In Section 8.6, a problem in maritime collision avoidance is treated: under what circumstances can one ship avoid collision with another? The approach taken is one of worst- case analysis, which quite naturally leads to a differential game formulation. The solution method, however, is also applicable to other collision avoidance problems which do not necessarily feature noncooperative decision making. In Section 8.7 we address the problem of role determination in a differential game wherein the roles of the players (viz. pursuing or evading) are not specified at the outset, but rather are determined as functions of the initial conditions. Role determination is applied to an aeronautical problem, which involves a dogfight between two airplanes.

8.2 Necessary and Sufficient Conditions for Saddle-Point Equilibria

The systems considered in this chapter can all be described by  $\dot{x}(t) = f(t, x(t), u_1(t), u_2(t))$ ,  $x(0) = x_0$ , (8.1) where  $x(t) \in \mathbb{R}^n$ ,  $u_i(t) \in \mathbb{R}^{m_i}$ ,  $i = 1, 2$ . The function  $f$  is continuous in  $t$ ,  $u_1$  and  $u_2$ , and is continuously differentiable in  $x$ . The first player, P1, who chooses  $u_1$ , is called the pursuer, which we shall abbreviate as P. The second player is called the evader and we shall refer to him as E instead of P2. The final time  $T$  of the game is defined by  $T = \inf\{t \in \mathbb{R}^+ : (x(t), t) \in A\}$ , (8.2) where  $A$  is a closed subset, called the target set, in the product space  $\mathbb{R}^+ \times \mathbb{R}^n$ . The boundary  $\partial A$  of  $A$  is assumed to be an  $n$ -dimensional manifold, i.e., a hypersurface, in the product space  $\mathbb{R}^+ \times \mathbb{R}^n$ , characterized by a scalar function  $\phi(t, x) = 0$ . This function is assumed to be continuous in  $t$  and continuously differentiable in  $x$ , unless stated differently. Since all differential games in this chapter will be of the zero-sum type, we have a single objective function  $J(u_1, u_2) = \int_0^T g(t, x(t), u_1(t), u_2(t)) dt + q(T, x(T))$ , (8.3)  $J_0$

PURSUIT-EVASION GAMES 425 where  $g$  is continuous in  $t$ ,  $u_1$  and  $u_2$ , and continuously differentiable in  $x$ ;  $q$  is continuous in  $T$  and continuously differentiable in  $x(T)$ . We define  $J$  as the cost function of the game in normal form, which is determined from  $L$  in terms of the strategies  $j_i(t, x(t)) = u_i(t)$ ,  $i = 1, 2$ . Only feedback strategies will be considered, and we assume  $j_i(\phi, a_i)$  to be piecewise continuous in  $t$  and  $x$ . With the permissible strategy spaces denoted by  $T_1$  and  $T_2$ , let us recall that a strategy pair  $(j_1^*, j_2^*)$  is in (saddle-point) equilibrium if  $J(t, x(t), j_1^*, j_2^*) \leq J(t, x(t), j_1, j_2^*) \leq J(t, x(t), j_1^*, j_2)$  (8.4) for all  $j_1^* \in T_1$ . The reason for considering feedback strategies, and not strategies of a more general class (such as strategies with memory), lies in the continuous- time counterpart of Thm. 6.9: as long as a saddle point in feedback strategies exists, it is not

necessary to consider saddle points with respect to other classes of strategies. Actually, as it will be shown in the sections to follow, "solutions" to various problems are, in general, first obtained in open-loop strategies, which may then be synthesized to feedback strategies, provided that they both exist.

8.2.1 The Isaacs equation In Corollary 6.6 a sufficiency condition (in the form of the solution of a partial differential equation) was obtained for the feedback saddle-point equilibrium of zero-sum differential games with a fixed final time. It will now be shown that this result essentially holds true also if the final time, instead of being fixed, is determined by the terminal constraint (8.2). The underlying idea here, too, is the principle of optimality; wherever the system is at any given arbitrary time, from then onwards the pursuer (respectively, evader), minimizes (respectively, maximizes) the remaining portion of the cost function under the feedback information structure. The function describing the minimax (upper) value of the cost function, when started from the position  $(t, x)$ , is  $V(t, x) = \min_{u_1} \max_{u_2} \int_t^T [g(s, x(s), u_1(s), u_2(s)) ds + q(T, x(T))]$ , (8.5) which is called the value function. Under the assumption that such a function exists and is continuously differentiable in  $x$  and  $t$ , it satisfies the partial differential equation  $dV/dt + \min_{u_1} \max_{u_2} [f(t, x, u_1, u_2) \cdot \nabla_x V + g(t, x, u_1, u_2)] = 0$  (8.6) which is known as the Isaacs equation. The lower (maximin) value function is defined analogously, with (8.5) and (8.6) replaced, respectively, by  $Y_-(t, x) = \max_{u_1} \min_{u_2} \int_t^T [g(s, x(s), u_1(s), u_2(s)) ds + q(T, x(T))]$  (8.7)

426 T. BAŞAR AND G. J. OLSDER  $dV/dt + \min_{u_1} \max_{u_2} [f(t, x, u_1, u_2) \cdot \nabla_x V + g(t, x, u_1, u_2)] = 0$  (8.8) If the differential game has equal upper and lower values, then we denote this common value by  $V$ , which satisfies (8.3) or equivalently (8.8), with the ordering of the min and max operations being irrelevant here. For future reference, let us rewrite this PDE explicitly:  $dV/dt + \min_{u_1} \max_{u_2} [f(t, x, u_1, u_2) \cdot \nabla_x V + g(t, x, u_1, u_2)] = 0$  (8.9) —  $f(t, x, u_1, u_2) + g(t, x, u_1, u_2)$  We now give a geometrical derivation of (8.9) in the case of a variable terminal time (as determined by (8.2)). It is also possible to extend the derivation given in Thm. 6.16 by allowing the terminal time to be variable, but we prefer the geometric derivation here, since it provides more insight into the problem, in particular if there is only a terminal pay-off. Accordingly, we assume now that  $g = 0$ , which in fact leads to no loss of generality because of the argument of Remark 5.1 (suitably modified to apply to continuous-time systems).  $q(U) = c_j$  Figure 8.1: Geometrical derivation of

the Isaacs equation, where  $c \leq c < C_2$ . The value function  $V$ , if it exists, is then defined by  $V(t,x) = \min \max q(T,x(T)) = \max \min q(T,x(T))$ , (8.10) when the initial position is  $(t,x)$ . One particular saddle-point trajectory is sketched in Fig. 8.1; it is indicated by curve  $m$  and the value corresponding to any point of  $m$  is  $c$ . We now make the assumption that for initial points in a certain neighborhood below  $m$ ,  $V(t,x) < c$ , and similarly for initial points 89 If not explicitly stated otherwise, it will be assumed henceforth that the upper and lower values are equal, that is, the max and min operations in (8.5) (or (8.7)) commute. Only in the next subsection, where we deal with viscosity solutions, will these operations not necessarily commute.

PURSUIT-EVASION GAMES 427 above  $m$ ,  $V(t,x) > c$ .<sup>90</sup> At point  $A$ , on curve  $m$ ,  $P$  will choose  $u_1 \in S_1$  so as to make the inner product between  $(dV/dx, dV/dt)'$  and  $(u_1, 1)'$  as small as possible, i.e., he tries to make the vector  $(u_1, 1)'$  point downward towards points corresponding to lowest possible costs. On the other hand,  $E$  wants to choose  $u_2 \in S_2$  in such a way so as to maximize this inner product, because then the system will move in  $E$ 's favorite direction, i.e., towards points corresponding to the highest possible costs. Hence  $u_1^*$  and  $u_2^*$  are chosen as the arguments of  $(dV, dV \setminus (dV \cdot dV \setminus \dots \min \max \dots f + \dots = \max \min \dots / + \dots \cdot (8'H) u_1 \ll 2 \setminus dx dt) \ll 2 u_1 V dx dt) K'$  Since  $m$  is an equilibrium trajectory, the system will move along  $m$  if  $u_1^*$  and  $u_2^*$  are applied, and that results in  $f dV dV \setminus \min \max \dots / + \dots = 0$ , (8.12)  $\ll iesi \ll 2gs2 \setminus dx J dt J K'$  which is equivalent to (8.6) if  $g = 0$ . If  $E$  plays optimally, i.e., he chooses  $u_2^*$ , then  $\wedge f(t,x,u_2^*) + \wedge > 0$  (8.13) for all  $u_1 \in S_1$ , which implies that the outcome of the game will be  $> c$ , since the system will move into the area where  $V > c$ . However,  $P$  can keep the outcome at  $c$  by playing  $u_1^*$ , since then the equality sign will apply in (8.13). Thus, without  $E$ 's cooperation (i.e., without  $E$  playing non-optimally),  $P$  cannot make the system move from  $A$  towards the area where  $V < c$ . Analogously, by interchanging the roles of the players in this discussion,  $E$  cannot make the system move from  $A$  in the "upward" direction if  $P$  plays optimally. Only if  $P$  plays non-optimally,  $E$  will be able to obtain an outcome higher than  $c$ . Because of these features, the  $n$ -dimensional manifold in  $R^n \times R^+$ , comprising all initial positions  $(t, x)$  with the same value, is called semipermeable, provided that  $\delta = 0$ . In conclusion, if the value function  $V(t,x)$  is continuously differentiate, then the saddle-point strategies are determined from (8.9). This equation also provides a sufficiency condition for saddle-point strategies as stated in the following theorem, the

proof of which is a straightforward two-player extension of Thm. 5.3; compare it also with Corollary 6.6. Theorem 8.1 // (i) a continuously differentiable function  $V(t,x)$  exists that satisfies the Isaacs equation (8.9), (ii)  $V(T, x) = q(T,x)$  on the boundary of the target set, defined by  $\mathcal{L}(t,x) = 0$ , and (Hi) either  $wu(t) = 71^*(t,x)$ , or  $u2^*(t) = 72^*(\mathcal{L},a:)$ , as derived from (8.9), generates trajectories that terminate infinite time (whatever  $\wedge^2$ , respectively  $71$ , is), then  $V(t,x)$  is the value function and the pair  $(71^*,72^*)$  constitutes a saddle point. 90If this assumption does not hold (which is the case of  $V$  being constant in an  $(n + 1)$ - dimensional subset of  $R^n \times R^+$ ) the derivation to follow can easily be adjusted.

428 T. BAŞAR AND G. J. OLSDER Remark 8.1 The underlying assumption of interchangeability of the min and max operations in the Isaacs equation is often referred to as the Isaacs condition. The slightly more general condition  $\min \max [p'f + g] = \max \min [p'f + g]$ ,  $u_1 \ u_2 \ u_2 \ u_1$  for all  $n$ -vectors  $p$ , is sometimes also referred to as the Isaacs condition. Regardless of which definition is adopted, the Isaacs condition will hold if both  $f$  and  $g$  are separable in  $u_1$  and  $u_2$ , i.e., they can be written as  $f(t,x,u_1,u_2) = f_1(t,x,u_1) + f_2(t,x,u_2)$ ,  $g\{t,x,v\},u_2) = g_1(t,x,u_1) + g_2(t,x,u_2)$ . If the Isaacs condition does not hold, and if only feedback strategies are allowed, then one has to seek for equilibria in the general class of mixed strategies defined on feedback strategy spaces. In this chapter, we will not extend our investigation to mixed strategies, and deal only with the class of differential games for which the Isaacs condition holds (almost) everywhere. Note, however, that, even if separability holds, this does not necessarily mean that the order of the actions of the players is irrelevant, since the underlying assumption of  $V$  being continuously differentiable may not always be satisfied. In regions where  $V$  is not continuously differentiable, the order in which the players act may be crucial.  $\square$  In the derivation of saddle-point strategies, equation (8.9) cannot directly be used, since  $V$  is not known at the outset. An alternative is to use Thm. 8.2, given below, which provides a set of necessary conditions for an open-loop representation (or realization) of the feedback saddle-point solution. In spite of the fact that it does not deal with feedback solutions directly, the theorem is extremely useful in the computation (synthesis) of feedback saddle-point solutions. It can be viewed as the variable terminal-time extension of Thm. 6.13 and the two-player extension of Thm. 5.4. The proof follows by direct application of Thm. 5.4, which is also valid under the weaker assumption of  $f$  and  $g$  being measurable in  $t$  (instead of being

continuous in  $t$ ) (see Berkovitz (1974), p. 52). Theorem 8.2 Given a two-person zero-sum differential game, described by (8.1)- (8.4), suppose that the pair  $\{u_1^*, u_2^*\}$  provides a saddle-point solution in feedback strategies, with  $x^*(t)$  denoting the corresponding state trajectory. Furthermore, let its open-loop representation  $\{u_i(t) = j^0(t, x^*(t)), i = 1, 2\}$  also provide a saddle-point solution (in open-loop policies). Then there exists a costate func-

PURSUIT-EVASION GAMES 429 tion  $p(-) : [0, T] \rightarrow \mathbb{R}^n$  such that the following relations are satisfied:  $x^*(t) = f(t, x^*(t), u_1^*(t), u_2^*(t))$ ,  $x^*(0) = x_0$ ,  $H(t, p, x, u_1, u_2) < H(t, p, x^*, u_1^*, u_2^*) < H(t, p, x^*, w_1, u_2^*)$ ,  $\forall u_1 \in S_1, \forall u_2 \in S_2$ ,  $p'(t) = -\frac{\partial}{\partial x} H(t, p(t), x^*(t), u_1^*(t), u_2^*(t))$ ,  $p'(T) = \frac{\partial}{\partial x} q(T, x^*(T))$  along  $i(T, x) = 0$ ,  $>$  (8.14) where  $H(t, p, x, u_1, u_2) = g(t, x, u_1, u_2) + p^T f(t, x, u_1, u_2)$  Remark 8.2 Provided that both open-loop and feedback equilibria exist, conditions of Thm. 8.2 will lead to the desired feedback saddle-point solution. Even if an open-loop saddle-point solution does not exist, this method can still be utilized to obtain the feedback saddle-point solution, and, by and large, this has been the method of approach adopted in the literature for solving pursuit- evasion games. For an example of a differential game in which this method is utilized, but an open-loop (pure) saddle point does not exist, see the "lady in the lake" game treated in Section 8.5. Also, for certain classes of problems, the value function  $V(t, x)$  is not continuously differentiate, and therefore the sufficiency conditions of Thm. 8.1 cannot be employed. Sufficiency conditions for certain types of nonsmooth  $V(t, x)$  have been discussed in Bernhard (1977), but they are beyond the scope of our treatment here.  $\square$  In the literature on pursuit-evasion differential games, it is common practice to write  $V_x$  instead of  $p'$  (for the costate vector), and this will also be adopted here. The reader should note, however, that in this context,  $V_x$  is only a function of time (but not of  $x$ ), just the way  $p'$  is. After synthesis, the partial derivative with respect to  $x$ , of the value function corresponding to the feedback equilibrium solution, coincides with this function of time, thus justifying the notation. We now illustrate the procedure outlined above by means of a simple example. As with most other problems to be considered in this chapter, this example is time-invariant, and consequently the value function and the saddle- point strategies do not explicitly depend on  $t$  (cf. Remark 5.5). Therefore the  $t$ -dependence has been dropped from the notation. Example 8.1 Consider the two-dimensional system described by  $\dot{x} = 2 + \sqrt{2}u_2 - \sin u_1$ ,  $\dot{u}_1 = -2 \cos u_1$  (0 where the initial value of the state lies in the half plane  $x^1 > 0$ ,  $u_2$  satisfies the constraint  $|u_2| < 1$  and no



restrictions are imposed on  $u_1$ . The target set is

430 T. BAŞAR AND G. J. OLSDER  $X_2 < 0$  and the objective functional is  $L = x_1(T)$ . Note that the game terminates when the  $x_1$ -axis is reached, and since  $\dot{x}_2 < -1$ , it will always terminate. In order to apply Thm. 8.2, we first write the Hamiltonian  $H$  for this pursuit-evasion game as  $H = VX_1 \{ \dot{x}_1^2 + V^2 \} u_2 - \sin u_1 \} + VX_2 \{ -2 - \cos u_1 \}$ , where  $VX_1 (= p_1)$  and  $VX_2 (= p_2)$  satisfy, according to (8.14),  $\dot{VX}_1 = 0$ ,  $\dot{VX}_2 = 0$ . (ii) Since the value function does not explicitly depend on time, we have  $V(x_1, x_2 = 0) = x_1$ , and hence  $VX_1 = 1$  along the  $x_1$ -axis. The final condition for  $VX_2$  along the  $x_1$ -axis cannot be obtained directly; it will instead be obtained through the relation  $\min_u \max_v H = 0$  at  $t = T$ .

Carrying out these min-max operations, we get  $u_2 = \text{sgn}(VX_1) = 1$  and hence  $\dot{y}_2^*(x) = 1$ , and the vector  $(\sin u_1, \cos u_1)$  is parallel to  $(VX_1, VX_2)$ . Substitution of these equilibrium strategies into  $H$  leads to  $2 + v/2 - \sqrt{(r + \dot{x}_1^2) - 2y_1^2} = 0$ , which yields  $VX_2 = 1$  at  $t = T$ . From (ii) it now follows that  $VX_1 = 1$  and  $VX_2 = 1$  for all  $t$ . In this example it is even possible to obtain  $V$  explicitly:  $y(x_1, a; 2) = x_1 + X_2$ . Hence, the feedback saddle-point strategies are constants:  $\gamma_1^*(a) = 7r/4$ ,  $\gamma_2^*(x) = 1$ . The corresponding state trajectories are straight lines, making an angle of  $7r/4$  radians with the negative  $X_2$ -axis (see Fig. 8.2). Figure 8.2: Vectogram for Example 8.1. Figure 8.2 also helps to visualize the geometric derivation of the saddle-point strategies. Suppose that the system is at point A corresponding to the value 6, say. Emanating from point A, the velocity vector  $\{x_1, \pm 2\}'$  has been drawn in terms of the vectors  $(0, -2)'$ ,  $((2 + V^2)\gamma_2^*, 0)'$  and  $(-\sin \alpha^*, -\cos \gamma_1^*)'$ , which add up to the RHS of (i). E, the maximizer, would like to make the angle  $\alpha$  (see Fig. 8.2) as large as possible. The best he can do, irrespective of P's decision, is to choose  $\gamma_2^* = +1$ . If E would play  $\gamma_2^* < 1$ , i.e.,  $\gamma_2 < 1$ , the angle  $\alpha$  will become smaller, and consequently the outcome will become smaller than 6.

PURSUIT-EVASION GAMES 431 Similarly, P, the minimizer, would like to have a vector  $(\pm 1, \pm 2)'$  with angle  $\alpha$  as small as possible, which leads to his optimal strategy  $\gamma_1^* = 7r/4$ ; for any other  $\gamma_1$ , the vector  $(x_1, \pm 2)'$  will point more to the right, and this is exactly what P wants to avoid. All conditions of Thm. 8.1 are satisfied and therefore the strategies obtained indeed constitute a saddle-point solution in feedback strategies.  $\square$  In addition to the necessary conditions of Thm. 8.2, we now exemplify some of the intricacies which might occur with respect to termination. Example 8.2 Consider the one-dimensional

system  $x = u_1 + u_2$ ,  $x(0) = x_0 > 0$ , where the controls  $u_1$  and  $u_2$  are constrained to  $-1 < u_i < 0$ ,  $i = 1, 2$ . The target set is the half line  $x < 0$ . The cost functional is  $L = \int_0^T (1-x) dt$ . The costate equation for this problem is and the optimal  $\{u_1^*, u_2^*\}$  satisfy  $(-1 \text{ if } p > 0, 0 \text{ if } p < 0, u_1^* = \text{undetermined if } p = 0, u_2^* = \text{undetermined if } p = 0)$ , (i)  $0 \text{ if } p < 0, 0 \text{ if } p > 0$ . How do we determine  $p(T)$ ? The appropriate condition in (8.14) cannot be used since that condition, which gives  $p(T)$  along the boundary of the target set, requires variations in  $x$  within  $\mathcal{L}(T, x) = 0$ ; but this latter relation only yields  $x = 0$ . In its stead, we make use of the definition of the derivative for  $p(T)$  and also make use of the functional form  $V(x) = Ax - \frac{1}{2}(Ax)^2$  for sufficiently small (positive)  $Ax$  (which is obtained through some analysis) to arrive at the value  $p(T) = 1$ . (See Problem 1 in Section 8.8 for another derivation.) Integration of the costate equation and substitution into (i) leads to  $u_1^* = -1, u_2^* = 0$ , for  $0 < x < 1$ . (ii) For  $x = 1$  we have  $p = 0$  and (i) does not determine the equilibrium strategies. For  $u_i = 0$  ( $i = 1, 2$ ), the system would remain at  $x = 1$  and the game would not terminate in finite time. To exclude the occurrence of such singular cases, we have to include an additional restriction in the problem statement, which forces  $P$  to terminate the game in finite time; this can be achieved by defining  $L = 0$  at  $T = 0$ . Then, at state  $x = 1$ , the preceding restriction forces  $P$  to choose  $u_1^*(x) = -1$ . Now we can integrate further in retrograde time. For

432 T. BAŞAR AND G. J. OLSDER  $t < T - 1$ ,  $p(t) = V_x(x^*(t))$  changes sign, and therefore, for  $x > 1$ , we obtain from (i):  $u_1^* = 0, u_2^* = -1$ ,  $x > 1$ . (iii) Now the question is whether the strategy pair  $\{u_1^*, u_2^*\}$ , defined by (ii) and (iii), constitutes a saddle point. To this end, Thm. 8.1 will be consulted. The value function is easily determined to be  $V(x) = x(2 - x)/2$ , and all conditions of Thm. 8.1 hold, except condition (iii). The reason why condition (iii) does not hold is because, for initial points  $x > 1$ , the strategy pair  $(u_1^*, u_2^*) = (0, -1)$  does not lead to termination in finite time, and hence, by playing  $u_2 = 0$ ,  $E$  would be much better off. However, in such a situation,  $P$  can do much better if he is allowed to have an informational advantage in the form  $u_1 = j_1(x(t), u_2(t))$ , (iv) i.e., at each instant of time,  $P$  is informed of  $E$ 's current decision (open-loop value).<sup>91</sup> If  $P$  chooses his new strategy as  $-1$  if  $x < 1$  or  $(x > 1 \text{ and } u_2 > -1)$ ,  $0$  otherwise, the saddle-point equilibrium will be restored again by means of the pair  $(u_1^*, u_2^*)$ , which follows from direct comparison. Another way of getting out of the "termination dilemma" is to restrict the class of permissible strategies to "playable pairs", i.e., to pairs for

which termination occurs in finite time (cf. Example 5.1 and Def. 5.7). Such an assumption excludes for instance the pair  $\{y_1 = 0, -y_2 = 0\}$  from consideration.  $\square$

8.2.2 Upper and lower values, and viscosity solutions In the previous subsection it was tacitly assumed (with the exception of (8.6) and (8.8)) that the Isaacs condition holds. In situations where it does not, one common approach is to endow one player with an instantaneous informational advantage over the other player, which we now briefly discuss.<sup>92</sup> An assumption throughout this subsection (only) will be that the final time  $T$  is fixed. With  $r^*$  taken as the set of all open-loop controls for  $P_i$ , let us introduce a mapping  $\gamma_i$  from  $T_1$  to  $T_2$  with the property that for each  $t \in [0, T]$  and  $u_1, v_1 \in F_1$  the following holds: if  $u_1(t) = v_1(t)$  for almost all  $r \in [0, t]$ , then  $\gamma_i[u_1](r) = \gamma_i[v_1](r)$  almost everywhere. In more popular terms this says that  $u_2(t)$  will depend on the past and current values of  $u_1$ , but not on its future values. It should be emphasized that the current value of  $u_1$  is part of the information for  $P_2$ . The set of all such strategies  $\gamma_i$  for  $P_2$  will be denoted by  $A_2$ . Similarly,  $\gamma_1$ <sup>91</sup> This is, of course, possible if  $E$  has access to only open-loop information, and can therefore only pick open-loop policies. If  $E$  has access to feedback information, however, (iv) is not physically realizable and a time delay has to be incorporated. <sup>92</sup>At the outset, this is different from the case of minimax ((8.5)) or maximin ((8.7)) values where one of the players is assumed to know not only the current but also the future values of the other player's controls.

PURSUIT-EVASION GAMES 433 and  $A_1$  are defined with the roles of the players reversed. Formally, a mapping  $\gamma_1$  from  $T_2$  to  $T_1$  will be called a strategy for  $P_i$  provided that for each  $t \in [0, T]$  and  $u_2, v_2 \in T_2$  the following holds: If  $u_2(t) = v_2(t)$  for almost all  $r \in [0, t]$ , then  $\gamma_1[u_2](r) = \gamma_1[v_2](r)$  almost everywhere. Definition 8.1 The value  $V^+$  of the zero-sum differential game described by (8.1) and (8.3), if the maximizer has instantaneous informational advantage over the minimizer, is defined as  $V^+(t, x) = \inf \sup \{ \int_t^T g(t, x(t), u_1(t), v_2(t)) dt + g(t, x(T)) \}$ , and the value  $V^-$  of the same game, but now with the minimizer having instantaneous informational advantage over the maximizer, is defined as  $V^-(t, x) = \sup \inf \{ \int_t^T g(t, x(t), u_1(t), v_2(t)) dt + q(t, x(T)) \}$ . Remark 8.3 Clearly,  $V^+(t, x) > V^-(t, x)$ .  $\square$  Remark 8.4 In the study of the existence of strategies that achieve the values  $V^+$  and  $V^-$ , the starting point is to partition the time interval of interest,  $[0, T]$ , into  $n$  subintervals each of length  $\delta = T/n$ . The informational advantage of one player with respect to the

other then concerns the entire subinterval to which the current time belongs. Thus one can define the so-called upper  $\epsilon$ -value and lower  $\delta$ -value. Under some regularity conditions, it can be shown that in the limit as  $n \rightarrow \infty$  and hence  $\epsilon \rightarrow 0$ , these values become equal to  $V^+$  and  $V^-$ , respectively.  $\square$

**Definition 8.2** The upper Hamiltonian and the lower Hamiltonian are, respectively,  $H^+(x,t,p) = \min [\max (p'f(t,x,u_1) + g(t,x,v), u_2)]$  and  $H^-(x,t,p) = \max [\min (p'f(t,x,u_1) + g(t,x,v), u_2)]$ .  $u_1 \in S_1$   $u_2 \in S_2$  **Remark 8.5** To be very explicit, the maximizing  $u_2$ -value in the definition of  $H^+$  will in general depend on the minimizing  $u_1$ -value, i.e.,  $u_2(u_1)$ , whereas in the definition of  $H^-$  it will be exactly the other way around. Note also that  $H^+ > H^-$  for all values of their arguments.  $\square$  The following theorem now makes the connection between Defs. 8.1 and 8.2, and also relates them to (8.5) and (8.7). The precise regularity conditions under which it is valid, as well as its proof can be found in Barron, Evans and Jensen (1984).

434 T. BAŞAR AND G. J. OLSDER **Theorem 8.3** Subject to appropriate regularity conditions on the functions  $f, g$  and  $q$ , the functions  $V^+$  and  $V^-$  are uniformly Lipschitz continuous, and hence differentiable almost everywhere, and they are solutions, respectively, of  $\Delta^+ + H^+(x,t, \Delta^+ p) = 0$ ,  $V^+(T,x) = g(T,x(T))$ ,  $dV^- (dV^- \Delta^- + H^-(x,t, \Delta^-) = 0$ ,  $V^-(T,x) = q(T,x(T))$ , and hence are equal to  $V$  and  $V_-$ , respectively. One can also define viscosity solutions for zero-sum differential games. In Chapter 5 the starting point to define viscosity solutions for optimal control problems was (5.65) (or (5.66)). In zero-sum differential games the equivalent starting point is (8.6) and (8.8), or their time-invariant equivalents if  $V$  does not directly depend on  $t$ . Since (5.65) and (8.6) are similar PDEs, the theory of viscosity solutions as presented in Section 5.8 remains equally valid for differential games, simply with  $H$  replaced by  $H^+$  or  $H^-$ . The following theorem, proved in Barron, Evans and Jensen (1984), and in Evans and Souganidis (1984) for problems with fixed terminal time, relates upper and lower value functions to viscosity solutions. **Theorem 8.4** The upper value function  $V^+$  is the viscosity solution of (8.6) and the lower value function  $V^-$  is the viscosity solution of (8.8). **8.3 Capturability** Consider the class of pursuit evasion games formulated in Section 8.2, under the additional assumption of time-invariance, i.e.,  $f, g$  and  $I$  are independent of  $t$ , and  $q$  does not depend on  $T$ . The target set  $A$  is therefore a tube in  $R^n \times R^+$ , with rays parallel to the time axis. The projection of  $A$  onto the  $R^n$ -space is denoted by  $A$ . Before studying the derivation of a saddle-point solution, we must first

address the more fundamental question of whether the target set can be reached at all. If it cannot be reached, we simply say that the pursuit-evasion game is not well denned. Accordingly, we deal, in this section, with the following qualitative pursuit-evasion game. The evader tries to prevent the state from reaching  $A$ , whereas the pursuer seeks the opposite. Once it is known, for sure, that the target set can be reached, one can return to the original (quantitative) pursuit-evasion game and investigate existence and derivation of the saddle-point solution. For the differential game of kind introduced above, we define an auxiliary cost functional as  $T - J - 1$  if  $\{x(t), t\} \in A$  for some  $t, 0 < t < \infty$ ,  $\backslash +1$  otherwise,

PURSUIT-EVASION GAMES 435 Figure 8.3: The construction of the barrier. thus converting it into a differential game of degree. The value function for this differential game can take only two values,  $V(x) = -1$  and  $V(x) = +1$ , where the former corresponds to initial states that lead to capture, while the latter corresponds to initial states from where capture cannot be assured.

Consequently, the value function is discontinuous, and therefore the theory of subsection 8.2.1 is not directly applicable. It can, however, be modified suitably to fit the present framework, which we now discuss below. Let us first note that only those points  $x$  of  $dA$  for which  $\min \max v'f(x, u_1, u_2) = \max \min v'f(x, u_1, u_2) < 0$  (8.15)  $u_1 u_2 u_2 u_1$  are candidates for a terminal position of the game. Here, the  $n$ -dimensional vector  $v$  is the outward normal of  $A$  at  $a$ ; (see Fig. 8.3). If strict inequality holds in (8.15), then the state will penetrate  $A$ ; furthermore, the points  $x$  for which the equality sign holds may only be grazing or touching points. The set of all points  $x$  of  $dA$  that satisfy (8.15) is called the usable part (UP) of  $dA$ . Let us suppose that the initial points situated in the shaded region of Fig. 8.3, and only those, can be steered to the target set by  $P$ , regardless of how  $E$  acts, and investigate whether we can determine the surface  $S$  that separates these initial states from the others which do not necessarily lead to termination. This surface  $S$  is sometimes referred to as a barrier. The points that are common to  $S$  and  $A$  are called the boundary of the usable part (BUP), it consists of those points of the UP for which the equality sign holds in (8.15). Consider a point  $A$  on  $S$ , at which a tangent hyperplane relative to  $S$  exists. The outward normal at  $A$ , indicated by  $p$ , is unique, apart from its magnitude. The surface  $S$  is then determined by  $\min \max p'f(x, u_1, u_2) = 0$ , (8.16)  $u_1 u_2$  which is only a necessary condition and therefore does not characterize  $S$  completely. This semipermeable surface should in fact have the

property that without P's cooperation, E cannot make the state cross S (from the shaded area

436 T. BAŞAR AND G. J. OLSDER  $V = -1$  to the non-shaded area  $V = +1$ ) and conversely, without E's cooperation, P cannot make the system cross S (from  $V = +1$  to  $V = -1$ ). Hence if a "tangential" penetration of S is impossible, (8.16) is also sufficient. Points of S at which no tangent hyperplane exists (such as  $B_i$  and  $E_2$  in Fig. 8.3) must be considered separately. Such points will be considered later, in Sections 8.6 and 8.7, while discussing intersections of several smooth semipermeable surfaces. To construct S, we now substitute the saddle-point solution of (8.16), to be denoted by  $u_i^* = j^*(x)$ ,  $i = 1, 2$ , into (8.16) to obtain the identity  $p'(x)f(x, u_1^*(x), u_2^*(x)) = 0$ . Differentiation with respect to  $x$  leads to (also in view of the relation  $(d/dx)(p'f) - (d^2f/dx^2) = 0$ , which follows from the discussion included in subsection 5.5.3 right after (5.42)) which is evaluated along S. The boundary condition for (8.17) is  $p\{T\} = v$ , (8.18) where  $v$  is the outward normal of A at S. We now introduce various features of the barrier S by means of the following example, which is known as the "homicidal chauffeur game".

Example 8.3 Consider a pursuer and an evader, both moving in a two-dimensional plane (with coordinates  $x, x-i$ ) according to  $\dot{X} = v \sin \theta$ ,  $\dot{\theta} = v \cos \theta u$ ,  $\dot{X}_e = v_e \cos \theta_e$ ,  $\dot{\theta}_e = v_e \sin \theta_e u_e$ , (8.19)  $\dot{X}_p = v_p \sin \theta_p$ ,  $\dot{\theta}_p = v_p \cos \theta_p u_p$ , where the subscripts p and e stand for pursuer and evader, respectively. The controls  $u_1$  and  $u_2$  satisfy the constraints  $|u_i| < 1$ ,  $i = 1, 2$ . The scalars  $v$  and  $v_2$  are the constant speeds, and  $u_p$  and  $W_2$  are the maximum angular velocities of P and E, respectively. Since the angular velocities are bounded, three coordinates ( $x_1, x_2$  and  $\theta$ ) are needed for each player to describe his position. For  $u_1 = 0$  (respectively,  $u_2 = 0$ ), P (respectively, E) will move in a straight line in the  $(x_1, X_2)$  plane. The action  $u = +1$  stands for the sharpest possible turn to the right;  $u = -1$  analogously stands for a left turn. The minimum turn radius for P is given by  $R_p = v_p / u_p$  and for E by  $R_2 = v_e / u_e$ . For this differential game, yet to be formulated, it is the relative positions of P and E that are important (and relevant), and not the absolute positions by means of which the model (8.19) has been described. Model (8.19) is six-dimensional; however, by means of the relative coordinates, defined as  $x = (x_e - x_p) \cos \theta_p - (x_2e - x_2p) \sin \theta_p$ ,  $X_2 = (x_1e - x_1p) \sin \theta_p + (x_2e - x_2p) \cos \theta_p$ , (8.20)  $\theta = \theta_e - \theta_p$ ,  $\theta_e = \theta_p + \theta$ ,

PURSUIT-EVASION GAMES 437 it can be converted into a three-dimensional model, as to be shown in the sequel. In this relative coordinate system,  $(x, X_2)$  denotes E's position with respect to P and the origin coincides with P's position. The  $x_2$ -axis is aligned with P's velocity vector, the  $x_1$ -axis is perpendicular to the  $x_2$ -axis (see Fig. 8.4), and the angle  $\theta$  is the relative direction (heading) in which the players move.  $v_P > v_E$  if  $v_P / v_E > 1$ . Figure 8.4: The relative coordinate system. With respect to the relative coordinates (8.20), model (8.19) can be described as  $\dot{x} = -U_1 v \{x_2 + V_2 \sin \theta, \dot{x}_2 = -v \{x_1 + U_1 \cos \theta + V_2 \cos \theta\}$ ,  $\dot{\theta} = U_2 / U_1 - U_1 u_2$ , which is sometimes referred to as the two-cars model. We shall consider a special version of (8.21), namely the one in which E is allowed to change his direction instantaneously (i.e.,  $\dot{\theta} = 0$ ). Then E will have complete control over  $\theta$ , which we therefore take as E's new control  $u_2$ . Also, normalizing both P's speed and minimum turn radius to one, we obtain  $\dot{X} = -U_1 X_2 + V_2 \sin u_2$ ,  $\dot{x}_2 = -1 + U_1 X_1 + V_2 \cos u_2$ , with no bounds imposed on  $u_2$ , while  $u_1$  still satisfies the constraint  $|u_1| < 1$ . The two normalizations introduced above determine the distance scale and time scale, respectively. Figure 8.4 depicts model (8.22); note that the angle  $u_2$  is measured clockwise with respect to the positive  $x_2$ -axis. This convention of measuring  $u_2$  has been initiated by Isaacs (1975), and since then it has been adopted in the literature on pursuit-evasion games. Note that the coordinate system is affixed to P and moves along with P through the "real" space. The target set A is defined through the inequality  $x_1 + x_2 < 1/2$ , that is, P tries to get within a distance  $1/2$  of the evader, whereas the evader tries to avoid this. The game thus formulated is referred to as the homicidal chauffeur game, in which a chauffeur (P) tries to overrun a pedestrian (E). The (8.22)

438 T. BAŞAR AND G. J. OLSDER homicidal chauffeur game is trivial for  $v_2 > 1 = v_1$ , since capture will then never be possible, provided that E plays the appropriate strategy. Therefore we shall henceforth assume  $0 < v_2 < 1$ . The UP (part of the circle described by  $x_1^2 + x_2^2 = 1/4$ ) is determined by those  $(x_1, x_2)$  for which the following relation holds (note that at the point  $(1/2, 0)$ , the vector  $(1/2, 0) = (0, 1, 0, 2)'$  is an outward normal to A):  $\min_{u_1} \max_{u_2} \{x_1(-u_1 x_2 + v_2 \sin u_2) + 1/2(-1 + u_1 x_1 + v_2 \cos u_2)\} < 0$ . Hence only those  $(x_1, x_2)$  for which  $1/2 > v_2/3$  on the circle  $x_1^2 + x_2^2 = 1/4$  constitute the UP. Intuitively this is clear: directly in front of P, E cannot escape; if E is close to P's side, he can side-step and avoid capture, at least momentarily; if E is

behind P, i.e.,  $0:2 < 0$ , there is no possibility for immediate capture. In order to construct the surface S, we shall start from the end point  $(0:1 = (3y/(1 - v^2), x_2 = v^2p)$ , on the capture set. Another branch of S can be obtained if we start from the other end of the UP, which is  $(0:1 = -3y/(1 - v^2), x_2 = v^20)$ , but that one is the mirror image of the first one; i.e., if the first branch is given by  $(x(t), x_2(t))$ , then the second one will be given by  $(-x(t), x_2(t))$ . Therefore we shall carry out the analysis only for the "first" branch. Equation (8.16) determines S and reads in this case  $\min \max \{p(-u_1x_2 + v^2\sin u_2) + p^2(-1 + u_1x + v^2\cos u_2)\} = 0, u_1 u_2$  whereby  $\sin u_2 = \frac{P_1}{\sqrt{P_1^2 + P_2^2}}$ ,  $\cos u_2 = \frac{P_2}{\sqrt{P_1^2 + P_2^2}}$ . (8.23) The differential equations for  $p$  and  $p_2$  are, from (8.17),  $\dot{p} = -u_1 p^2, \dot{p}_2 = u_1 p x$ . (8.24) The final condition for  $p$ , as expressed by (8.18), will be normalized to unit magnitude, since only the direction of the vector  $p$  plays a role, i.e.,  $P_1(T) = \sqrt{1 - w^2} \cos \alpha; p_2(T) = \sqrt{1 - w^2} \sin \alpha$  (8.25) for some appropriate angle  $\alpha$ . In order to solve for  $p(t)$ , we need to know  $u_1^*$ . At  $t = T, s = 0$  and therefore  $w x(T)$  cannot be determined directly. For an indirect derivation we first perform  $\frac{ds}{dt} = \frac{1}{2} \frac{d}{dt} (P_1^2 + P_2^2) = -P_1$  (8-26)

PURSUIT-EVASION GAMES 439 Since  $p t > 0$  at  $t = T$ , we have  $ds/dt < 0$  at  $t = T$ , which leads to  $s(t) > 0$  for  $t$  sufficiently close to  $T$ , and therefore  $u_1 = +1$  just before termination. The solution to (8.24), together with (8.25) and  $u_1 = +1$ , now becomes  $p_1(t) = \cos(t - T + \alpha), p_2(t) = \sqrt{1 - w^2} \sin(t - T + \alpha)$ , (8.27) and the equations determining the branch of S emanating from  $x = \frac{3y}{1 - w^2} > \frac{3\sqrt{1 - v^2}}{2} x_2 = \frac{1 + x + \sqrt{1 - w^2} \sin(t - T + \alpha)}{2}, x_2(T) = \frac{1 - \cos(t - T)}{2} \sqrt{1 - w^2}$ , which are valid as long as  $s(t) > 0$ . The solution is  $x_1(t) = (0 + \sqrt{1 - w^2}(t - T)) \cos(t - T + \alpha) + 1 - \cos(t - T), x_2(t) = (\sqrt{1 - w^2}(t - T) \sin(t - T + \alpha) - \sin(t - T)) \sqrt{1 - w^2}$ , which leaves the circle  $x_1^2 + x_2^2 = \frac{3}{2}$  tangentially. What we have to ensure is that the trajectory (8.28) lies outside A. This is what we tacitly assumed in Fig. 8.3, but which unfortunately does not always hold true as we shall shortly see. If we define  $r = y/(x_1 + x_2)$ , we require  $r > 0$  at  $t = T$  (note that  $r = 0$  at  $t = T$ ), and some analysis then leads to the conclusion that this holds if, and only if,  $P_2 + v < 1$ . (8.29) For parameter values  $(f, \sqrt{1 - w^2})$  for which  $0 + v > 1$ , the semipermeable surface S moves inside A and thereby loses its significance.<sup>93</sup> In the latter case the proposed construction (which was to split up the state space into two parts,  $V = +1$  and  $V = -1$ ) obviously fails. In the present example we henceforth assume that (8.29) holds. Let us now explore the trajectories described by (8.28). The following two



possibilities may be distinguished: (i)  $X_i(t_2) = 0$  for some  $t_2 < T$  and (ii)  $x^i(t) > 0$  for all  $t < T$ , i.e., as long as  $s(t) > 0$ . These two cases have been sketched in Fig. 8.5. From (8.28) it follows that the two branches of  $S$  intersect (if they do, the point of intersection will be on the positive  $X_2$ -axis) if  $f_3 < v^2 M c \sin(v^2) + y^j \{1 - v^2\}^{-1}$ , (8.30) whose verification is left as an exercise to the reader. Case 1 (Fig. 8.5a) For the situation depicted in Fig. 8.5a, let us first assume that  $E$  is initially somewhere within the shaded area. Then he cannot escape from the area, except 93 In fact, it is still a semipermeable surface, but for a different ("dual") game in which  $E$  is originally within the set  $x^i + x^j < 0.1$  and tries to remain there, whereas  $P$  tries to make the state cross the circle  $X^j + x^i = 0.2$ . (8.28)

440 T. BAŞAR AND G. J. OLSDER (a) (b) Figure 8.5: Possible configurations for the barriers. for moving through the UP (= capture), provided that  $P$  plays optimally, which is that  $P$  will never let  $E$  pass through the two semipermeable surfaces indicated by  $S_1$  and  $S_2$ . If  $E$  is initially outside the shaded area, however, he will never be captured (assuming that he plays optimally when in the neighborhood of  $S_1$  or  $S_2$ ). Q  $S_1$  \*1 Figure 8.6: A leaking corner. If  $E$  is initially in the shaded area, the intersection point  $C$  requires closer attention. Toward this end let us assume that the state is initially on  $S_1$ ; then  $P$  should try to force it to move in the direction towards the UP (which fortunately can be done in this example), and not towards  $C$  and pass  $C$  eventually. This latter possibility has been depicted in Fig. 8.6, where  $P$  tries to keep  $E$  in the shaded region. Suppose that, in Fig. 8.6, the state is initially at point  $Q$ . In order for  $P$  to keep the state within or on the boundary of the shaded area, and for  $E$  to try to escape from it, both players should employ their strategies according to (8.16). If these strategies cause the state to move in the direction of  $C$ , and eventually pass  $C$ , then  $P$  cannot keep  $E$  within the shaded area, though he can prevent  $E$  from passing either  $S_1$  or  $S_2$ . The intersection point of the two semipermeable surfaces "leaks" and it is therefore called a leaking corner. As already noted, point  $C$  in Fig. 8.5a does not leak. The next question is: once  $E$  is inside the shaded area, can  $P$  terminate the

PURSUIT-EVASION GAMES 441 game in finite time? In other words, would it be possible for  $E$  to maneuver within the shaded area in such a way so that he can never be forced to pass the UP? Even though such a stalemate situation may in general be possible, in the present example  $E$  can always be captured in

finite time, which, for the time being, is left to the reader to verify; Example 8.5 in Section 8.4 will in fact provide the solution. Case 2 (Fig. 8.5b) We now turn our attention to the situation depicted in Fig. 8.5b in which case  $0 > v_2 \arcsin(w_2) + v/(1 - \ll) - 1$ . (8.31) Figure 8.7: The failure of continuation of the semipermeable surface. In the figure, the semipermeable surfaces have been drawn up to  $t_i = T - t_v - 2a$  which corresponds to the first switching time of  $u_1$  in retrograde time from  $t = T$ . (From (8.26) and (8.27) it follows that  $s(t) = -\sin(f - T + a) + \sin a$ .) For  $K_f$  we have to continue with  $u_1^* = -1$ . Carrying out the corresponding analysis, it follows that the resulting continuation of  $S_i$ , designated by  $S_i$ , would be as sketched in Fig. 8.7. The trajectory leaves point C in exactly the opposite direction as it arrived, and the semipermeable surface has a sharp corner (which is also known as a cusp). For the semipermeable surface  $S_i$ , P can keep the state on the shaded side of  $S_i$ , and the same is true for  $S_i$  (the reader should note the directions of  $p(t)$  in Fig. 8.7). Clearly,  $S_i$  and  $S_i$  together do not constitute a semipermeable surface. Therefore,  $S_i$  will be disregarded altogether, and we shall consider only the semipermeable surface of Fig. 8.5b. It is now claimed that termination is possible from any initial point. If the state is originally at I, say, (see Fig. 8.5b), E can prevent the state from passing  $S_i$ ; P, on the other hand, can force the system to move along the dotted line, and therefore termination will occur. We have to ensure that no little "island" exists anywhere in the  $(\mathbb{R}^2)$  plane, surrounded by semipermeable surfaces, and not connected to A, and wherein E is safe. Such a possibility is easily ruled out in this example, on account of the initial hypothesis that P's velocity

442 T. BAŞAR AND G. J. OLSDER is greater than E's, since then P can always force the state to leave such an island, and therefore such an island cannot exist to begin with.  $\square$  In conclusion, the foregoing example has displayed various features of the barriers which separate states from where capture is possible from those from which it is not possible. If in particular two semipermeable surfaces intersect, then we have to ensure that the composite surface is also semipermeable, i.e., there should not be a leaking corner. Furthermore, in the construction discussed above, within the context of Example 8.3, we have treated only the case where the barrier does not initially move inside the target set. 8.4 Singular Surfaces An assumption almost always made at the outset of every pursuit-evasion game is that the state space can be split up into a number of mutually disjoint regions, the value function being

continuously differentiable in each of them. The behavior and methods of construction of  $V\{t,x\}$  are well understood in such regions. The boundaries of these regions are called singular surfaces, or singular lines if they involve one-dimensional manifolds, and the value function is not continuously differentiable across them. In this book we in fact adopt a more general definition of a singular surface. A singular surface is a manifold on which (i) the equilibrium strategies are not uniquely determined by the necessary conditions (8.14), or (ii) the value function is not continuously differentiable, or (iii) the value function is discontinuous. In general, singular surfaces cannot be obtained by routinely integrating the state and costate equations backward in time. A special singular surface that manifests itself by backward integration is the switching surface, or equivalently, a transition surface. In Example 5.2, we have already encountered a switching line within the context of an optimal control problem, and its appearance in pursuit-evasion games is quite analogous. The procedure for locating a switching or transition line is fairly straightforward, since it readily follows from the state and costate equations. A singular surface which does not manifest itself by backward integration of state and costate equations is called a dispersal surface. The following example now serves to demonstrate what such a surface is and how it can be detected.

Example 8.4 The system is the same as that of Example 8.1:  $\dot{x}_1 = u_2 + V_2$ ,  $\dot{x}_2 = -\sin u_1$ ,  $\dot{u}_2 = -2 - \cos u_1$ , with  $|u_2| < 1$  and with no bounds on  $u_1$ . The initial state lies in the positive  $x_1$  half plane and the game terminates when  $x_2(t) = 0$ . Instead of taking  $L = x_1(T)$  as in Example 8.1, we now take  $L = x_2(T)$ . In order to obtain the solution, we

PURSUIT-EVASION GAMES 443 initially proceed in the standard manner. The Hamiltonian is  $H = V_{x_1} u_2 + V_{x_2} (-\sin u_1) + V_{u_2} \{-2 - \cos u_1\}$ , (i) where  $V_{x_1} = 0$ ,  $V_{x_2} = 0$ . Since  $V$  is the value function, we obviously have  $V_{x_1} = 2x_1$  along the  $x_1$ -axis. Then, as candidate equilibrium strategies, we obtain  $u_2 = \text{sgn}(V_{x_1}) = \text{sgn}(x_1(T))$ ,  $(\sin u_1 \parallel \cos u_1)$ , where  $\parallel$  stands for "parallel to". Note that the  $u_1$  and  $u_2$  in (ii) are candidates for open-loop strategies and therefore the expression  $u_2(t) = \text{sgn}(x_1(T))$  does not contradict with causality. Substituting (ii) and  $V_{x_1} = 2x_1$  into (i) with  $t = T$  and setting the resulting expression equal to zero, we obtain  $V_{x_2}(x(T)) = 2|x_1(T)|$ . The costate variables  $V_{x_1}$  and  $V_{x_2}$  are completely determined now. For terminal state conditions of  $(x_1 > 0, x_2 = 0)$  we find, as candidate equilibrium strategies,  $u_1 = \pi/2$  if  $x_1 < 0$ ,  $u_2 = 1$ , and in the case of terminal state conditions  $(x_1 < 0, x_2 = 0)$  their

counterparts are  $U_1 = -$ ,  $U_2 = -1$ . The corresponding state trajectories are sketched in Fig. 8.8, wherefrom we observe that they intersect. Starting from an initial position  $\mathbb{I}$ , there are two possibilities, but only one of them—the trajectory going in the "south-east" direction—is optimal. Starting from  $7_2$ , the trajectory in the "south-west" direction would be optimal. The separation line between the south-east (SE) and south-west (SW) directions is obviously the  $a^{\wedge}$ -axis, since it is on this line where the two values (one corresponding to a trajectory going SW, the other one to SE) of the cost function match. While integrating the optimal paths backward in time, we should stop once the  $X_2$ -axis is reached, which is called a dispersal line. More precisely, it is a dispersal line for the evader, since, if the initial condition is on this line, it is E who chooses one of the two optimal directions along which to proceed; along both of these directions, the system trajectory leaves the dispersal line and both lead to the same outcome.  $\square$  (ii) Figure 8.8: Intersection of candidate optimal trajectories.

444 T. BAŞAR AND G. J. OLSDER In the foregoing example we have discovered a dispersal line for the evader. Of course, dispersal lines for the pursuer also exist. Toward that end, consider the homicidal chauffeur game introduced in Example 8.3. Suppose that the parameters of the game ( $l$ ,  $v_2$ ) are such that termination is always possible and that the cost functional stands for the time elapsed before capture. If E is far behind P (i.e., if, in Fig. 8.4, the state is  $X \setminus = 0, x_2 < -C$ ), then P will turn either to the right or to the left as quickly as possible, so as to have E directly in front of him. Therefore, the half line ( $x \setminus = 0, x_2 < 0$ ) for some negative  $a$ , is a dispersal line for the pursuer. By means of the same game we now introduce another type of a singular surface, viz. the so-called universal line or universal surface. Assume that (8.29) and (8.31) hold, which means that E can be captured by P, whatever the initial position is. Suppose that E is situated far in front of P, i.e.,  $x_2 > l$ , relatively close to the  $X_2$ -axis. Then, the best P can do is to turn towards E, such that, in the relative ( $x \setminus, X_2$ ) space, E's position moves towards the  $X_2$ -axis. Once E is on the  $X_2$ -axis, the remaining phase of the game is trivial, since it is basically a chase along a straight line. A line, or more generally a surface, to which optimal trajectories enter from both sides and then stay on, is called a universal line (surface). It turns out that, in the case when both (8.29) and (8.30) hold, the  $X_2$ -axis is also a universal line, as shown in the following example. Example 8.5 Consider the homicidal chauffeur game as introduced in

Example 8.3, in which the cost for P is the "time to capture", and where the parameters  $V_1$  and  $V_2$  are such that (8.29) and (8.30) hold. In order to construct candidates for optimal paths by backward integration from the UP, we now follow the analysis of Example 8.3 rather closely, along with the formulas (8.23) and (8.24). For the present example, equation (8.6) becomes  $\min \max \{ \sqrt{-u_1^2 + v_2^2 \sin^2 \theta} + V_1 \sqrt{1 - u_1^2 + v_2^2 \cos^2 \theta} + 1 \} = 0$ , whereby  $u_1 = \text{sgn}(V_1 x_2 - V_2 x_1)$ ;  $(\sin \theta, \cos \theta) \parallel (V_2 T_1, V_1)$ , (8.32) where  $V_1$  and  $V_2$  satisfy along optimal trajectories. The final conditions for the costate equations are  $\sin \theta_1(r) \cos \theta_1(r) = 1 - \cos \theta_0(T) - t; 2' 2 \cos \theta_0(T) - t; 2'$  the derivation of which is left as an exercise to the reader. (Rewrite the problem in polar coordinates  $(r, \theta)$ , and utilize the boundary condition  $V_2 = 0$  at termination.)

PURSUIT-EVASION GAMES 445 Exactly as in Example 8.3, it can be shown that  $u_1^* = +1$  near the end. For the right half plane, the terminal conditions are  $x_1(T) = r \sin \theta_0$ ,  $x_2(T) = r \cos \theta_0$ , where  $\theta_0$  is the parameter determining the final state, and the solution of the state equation close to termination can be written as  $x_1(t) = 1 - \cos T + (\sqrt{3} - w_2 - r) \sin(\theta_0 + T)$ ,  $x_2(t) = \sin T + (\sqrt{3} - w_2 T) \cos(\theta_0 + T)$ ,  $J(t)$  where  $r$  is the retrograde time  $T - t$ . Equation (8.33) displays two facts: (i)  $x_2(t) > 0$  for  $t > 0$ , which means that part of the shaded region in Fig. 8.5a is not yet filled with (candidate) optimal trajectories, (ii) for  $r = \sqrt{3}/v_2$ , the state  $(x_1(r), x_2(T))$  is independent of the arrival angle  $\theta_0$ . Figure 8.9: The decision points in the homicidal chauffeur game. Figure 8.9 depicts these trajectories. Point  $A$  in this figure has the coordinates  $(\sqrt{3}/v_2, x_2(P/v_2))$ ; it is a so-called decision point, E has various options, all leading to the same pay-off. Point  $A_2$  is the mirror image of  $A$  with respect to the  $x_2$ -axis. In order to fill the gap with optimal paths, we try a singular arc for which the switching function is identically zero, i.e.,  $V_1 x_2 - V_2 x_1 = 0$ . (8.34) For this to be true, its time derivative must also be identically zero, which yields  $V_1(-1 + v_2 \cos \theta) - V_2 v_2 \sin \theta = 0$ . (8.35) For (8.34) and (8.35) to be true when not both  $V_1$  and  $V_2$  are zero, the determinant of the coefficients must be zero, which results in  $x_1 - \sqrt{3} V_2 (x_1 \cos \theta - a \sin \theta) = 0$ .

446 T. BAŞAR AND G. J. OLSDER Figure 8.10: The  $a$ -axis is a universal line. Because of (8.32) and (8.34),  $x_1 \cos \theta - a \sin \theta \geq 0$ , which ultimately leads to  $\theta = 0$  as the only candidate for a singular surface. Along this line, the switching function is  $V_1 x_2$  and therefore we must have  $V_1 = 0$ , leading to  $\{u_1^* = 0, u_2^* = 0\}$ . Along the line  $x_1 = 0$ , both P and E follow the same straight

path in real space. From the singular arc, paths can leave in retrograde time at any instant, leading to Fig. 8.10, thus filling the whole shaded area in Fig. 8.5b with optimal trajectories. The line  $x \setminus = 0$  is a universal line. It should be noted that both  $V$  and  $dV/dx$  are continuous across the universal line (see Problem 2, Section 8.8). □ dispersal line equivocal line transition line XX universal line focal line switching envelope Figure 8.11: Classification of singular surfaces. There are also other types of singular surfaces. In Fig. 8.11 we have sketched all hitherto known singular surfaces of co-dimension one (i.e., these surfaces are manifolds of dimension one less than that of the state space) and across which  $V(x)$  is continuous. It is believed that no other such singular surface exists, but

PURSUIT-EVASION GAMES 447 a verification of this claim is not yet available. The singular surfaces not treated heretofore in this chapter are the equivocal line, focal line and switching envelope. They will not be treated in this book; for examples of games where equivocal lines arise, the reader is referred to Isaacs (1975), Lewin and Olsder (1979) and Lewin (1994); the focal line appears for instance in the "obstacle tag" example (Breakwell, 1971) and a game with a switching envelope is treated in Breakwell (1973). Problem 7 in Section 8.8 is also believed to feature a switching envelope in its solution. In the cases of the equivocal line and the switching envelope, one of the players has the option of either staying on the singular surface or leaving it; both possibilities lead to the same outcome. The difference between these two surfaces is, that in the switching envelope case the trajectories enter (in forward time) tangentially, whereas in the equivocal line case they do not. An analogous difference exists between the universal line and the focal line. Some of the singular surfaces can only be approximated by strategies of the form  $u_1 = \gamma_1(x), u_2 = \gamma_2(x)$ . In the case of the focal line, for instance, one player, say P, tries to get the state vector off the focal line, but E brings it back to the focal line. Strategies of the form  $u_1 = \gamma_1(x), u_2 = \gamma_2(x, u_1)$ , with an instantaneous informational advantage to the evader, such as those described in Section 8.2 with respect to upper and lower values, would be able to keep the state precisely on the focal line. Singular surfaces across which  $V(x)$  is discontinuous are the barrier and the safe contact. We have already encountered barriers in Example 5.2 in the context of optimal control, and in Section 8.3 where games of kind were treated. Their appearance in games of degree will now be delineated within the context of the homicidal chauffeur

game. An example of a safe contact can be found in Chigir (1976). Example 8.6 Consider the homicidal chauffeur game with the cost function defined as the time-to-go. In Example 8.3 we have described how semipermeable surfaces, starting from the end points of the UP of the target, could be constructed. We now treat the case depicted in Fig. 8.5b, i.e., the case when the parameters  $V_2$  and  $\beta$  are such that (8.31) holds. If  $E$  is initially situated at point  $C_i$  (see Fig. 8.5b),  $P$  cannot force him across  $S_i$ , but instead he can force him to go around it. Therefore it is plausible—and this can be proven rigorously (see Merz, 1971), that capture from  $C^+$  will take considerably more time than capture from  $C^-$ . Hence, the value function is discontinuous across  $S_i$  which, therefore, is a barrier.  $\square$  This concludes the present section on singular surfaces. In conclusion, the crucial problem in the construction of the value function is to locate the singular surfaces, but hitherto this problem has not been solved in a systematic way. On the other hand, once a particular  $V$  has been constructed, which is continuously differentiable in each of a finite number of mutually disjoint regions of the state space, some conditions (known as "junction conditions") exist to check whether the  $V(x)$  obtained is really the value function or not. Since these conditions are

448 T. BAŞAR AND G. J. OLSDER not yet very well understood, we do not treat them here; but for some discussion on this topic the reader is referred to Bernhard (1977). 8.5 Solution of a Pursuit-Evasion Game: The Lady in the Lake In this section we obtain the complete solution of a pursuit-evasion game called "the lady in the lake" by utilizing the techniques developed in the previous sections. This game features a dispersal surface for the evader; there is also a decision point. A lady ( $E$ ) is swimming in a circular pond with a maximum constant speed  $V_2$ . She can change the direction in which she swims instantaneously. A man ( $P$ ), who has not mastered swimming, and who wishes to intercept the lady when she reaches the shore, is on the side of the pond and can run along the perimeter with maximum speed 1. He, also, can change his direction instantaneously. Furthermore, it is assumed that both  $E$  and  $P$  never get tired.  $E$  does not want to stay in the lake forever, though; she wishes eventually to come out without being caught by the man. (On land,  $E$  can run faster than  $P$ .)  $E$ 's goal is to maximize the pay-off, which is the angular distance  $PE$  viewed from the center of the pond, at the time  $E$  reaches the shore (see Fig. 8.12).  $P$  obviously wants to minimize this pay-off. To make the game nontrivial, it is assumed that  $v_2 < 1$ . Figure 8.12: The "lady in the lake". Even

though the problem could be cast in a rectangular coordinate system fixed in space, a description in the relative coordinates  $\theta$  and  $r$  turns out to be simpler, where  $\theta$  is the angle between  $P$  and  $E$  (see Fig. 8.12 for the direction) and  $r$  is  $E$ 's distance from the center of the lake. This coordinate system is called relative since it is attached to  $P$ ; it is not fixed in space, and yet it describes the game completely. The kinematics of the game are  $\dot{\theta} = -\frac{v_2 \sin \theta}{r}$ ,  $\dot{r} = v_1 \cos \theta - v_2$ ,

PURSUIT-EVASION GAMES 449 where  $R$  is the radius of the pond.  $P$ 's control  $u_1$  is restricted by  $|u_1| \leq 1$  while  $E$ 's control  $u_2$ , which stands for the angle of  $E$ 's velocity vector with respect to the radius vector  $CE$  (see Fig. 8.12), is not restricted in any way. We now seek for the equilibrium strategies in feedback form. The cost function is  $J(\theta, r)$ , where  $T$  is defined as  $T = \min\{t: r(t) = R\}$ , and  $-T \leq t \leq +T$ . The Isaacs equation for this problem is  $\max_{u_1} \min_{u_2} \left( \frac{dV(\theta, r)}{dt} + \frac{1}{2} \frac{d^2V(\theta, r)}{dt^2} \right) = 0$  (8-36) whence the differential equations for the costate variables along the optimal trajectories are  $\frac{d\lambda_\theta}{dt} = -\frac{\partial H}{\partial \theta}$ ,  $\frac{d\lambda_r}{dt} = -\frac{\partial H}{\partial r}$  and the value function at  $t = T$  is given by  $V(\theta(T), r(T)) = J(\theta(T), r(T))$ . The optimal control  $u_1^*$ , taken as the open-loop representation of the feedback equilibrium strategy,<sup>94</sup> can now be written as  $u_1(t) = \text{sgn}(\frac{dV}{d\theta}) = \text{sgn}(\lambda_\theta)$ . Since we assumed  $-T \leq t \leq +T$ , this implies that  $P$  will move in  $E$ 's direction along the smallest angle. Substituting  $u_1^*$  and  $u_2^*$  into (8.36), we obtain  $\sin^2 \theta^*(t) = \frac{R^2}{r(t)}$ . (8.38)  $r(t)$  Two observations can be made at this stage. First, (8.38) is valid as long as  $r(t) > R^2$ ; for  $r(t) < R^2$  the outlined derivation does not hold. Second,  $E$  swims along a straight line in real space, tangent to a circle of radius  $R$  (see Fig. 8.13a). This result could also have been obtained geometrically (see Fig. 8.13b). In the relative coordinate system,  $E$ 's velocity  $v_r$  is made up of two components: (i)  $E$ 's own velocity vector in real space,  $v_2$ , and (ii) a vector  $v_i$ , obtained by rotation around the center  $C$ , opposite to  $P$ 's direction of motion (recall that  $P$  tries to reduce  $J(\theta, r)$ ), and with magnitude  $r(t)/R$ . The best policy for  $E$  is to make  $v_2$  and the resulting velocity vector  $v_r$  perpendicular to each other. For any other choice of  $u_2$  (i.e., direction of  $v_2$  in Fig. 8.13b), the angle between  $CE$  and  $v_r$  will become smaller, or, in other words,  $E$  will move faster into  $P$ 's direction, which is precisely what  $E$  tries to avoid. Therefore,  $E$ 's equilibrium strategy is to keep the angle between  $v_2$  and  $v_r$  at  $\pi/2$  radians, or equivalently,  $\sin^2 \theta^* = R^2/r(t)$ . <sup>94</sup>Note that, although the open-loop representation of the feedback



saddle-point solution (if it exists) may exist, an open-loop saddle-point solution surely does not exist for this pursuit evasion game. Whether a mixed open-loop saddle-point equilibrium exists is an altogether different question which we do not address here.

450 T. BAŞAR AND G. J. OLSDER (a) (b) Figure 8.13: E's equilibrium strategy for the "lady in the lake". This way of constructing E's optimal strategy fails for  $r(t) < Rv_2$  in which case E can achieve a larger angular velocity (with respect to the center C) than P, and therefore she can always maneuver herself into the position  $\theta(t) = \theta_0 - \omega t$ , i.e., to a position diametrically opposite from P. Note that, in order to maintain  $\dot{\theta}(t) = \dot{\theta}_0$ , E would have to know P's current action. If she does not have access to this (which is a more realistic situation), she can stay arbitrarily close to  $\theta(t) = \theta_0$  as she moves outward towards the circle with radius  $Rv_2$ . In case E would know P's current action, she can reach the point  $(\theta = \theta_0, r = Rv_2)$  in finite time (see Problem 3 in Section 8.8); if she does not know P's current action, she can get arbitrarily close to this point, also in finite time.<sup>95</sup> From the position  $(\theta = \theta_0, r = Rv_2)$ , E moves right or left along the tangent. More precisely, E will first swim outward to a position  $r > Rv_2$  which is sufficiently close to  $r = Rv_2$ . Then P has to make a decision as to whether to run clockwise or anti-clockwise around the pond (P cannot wait since, as E moves outward, his pay-off will then become worse). Once E knows which direction has been chosen by P, she will swim "away" from P, along the tangent line just described. In the region  $r(t) > Rv_2$ , P's angular velocity is larger, and therefore he will continue to run in the direction he had chosen before. The outcome of the game is readily calculated to be  $\theta(T) = \theta_0 + \arccos \frac{v_2}{V} (1 - \frac{v_2}{V})$ , (8.39)  $V > v_2$  which holds true for all initial positions inside the circle of radius  $Rv_2$ . The lady can escape from the man if  $\theta(T) > 0$ , which places a lower bound on E's speed:  $V > 0.21723$  From all initial positions inside the pond, E can always first swim to the center and then abide by the strategy just described, resulting in the outcome (8.39). From some initial positions she can do better, namely, the positions in the shaded area in Fig. 8.14 which is bounded by the pond and <sup>95</sup>This would inevitably involve some delay in the current state information of E and thereby some memory in her information. Restriction to only feedback strategies would lead to some subtle measurability questions.

PURSUIT-EVASION GAMES 451 Figure 8.14: Optimal trajectories in the

relative space. the two equilibrium trajectories which, constructed in retrograde time, end at  $(r = Rv_2, \theta = \theta_1)$ . Within the shaded area the optimal strategies are determined by (8.37). The backward construction is only valid until  $\theta(t) = \theta_1$ ; the line segment  $(\theta = \theta_1, Rv_2 < r < R)$  obviously forms a dispersal line on which it is  $P$  who decides whether to go "to the left" or "to the right". In the non-shaded area of the pond in Fig. 8.14, the value function is constant and equals  $V_0(T)$  given by (8.39). Properly speaking, this value can only be obtained if there is an informational advantage to  $E$  as explained above; see also Problem 9 in Section 8.8. If this is not the case and only feedback strategies are allowed, only an approximation to this value function can be obtained. Therefore, in such a case, a saddle point will not exist, but an e-saddle-point equilibrium will (cf. Section 4.2, specifically Def. 4.2). The solution obtained for the "lady in the lake" has a very special feature. If  $E$  would accidentally choose a wrong maneuver, she can always return to the center of the lake and start all over again. The point  $E_0$  in the relative coordinate system (see Fig. 8.14), which plays a special role, is sometimes referred to as a decision point.

### 8.6 An Application in Maritime Collision Avoidance

In this section, collision avoidance during a two-ship encounter in the open sea will be treated as a problem in the theory of differential games. Given two ships close to each other in the open sea, a critical question to be addressed is whether a collision can be avoided. If, for instance, we assume that the helmsman of ship 1 ( $P_1$ ) has complete information on the state of ship 2, whereas the helmsman of ship 2 ( $P_2$ ) is not aware of the presence of the first ship, this lack of information on the part of  $P_2$  may lead to a hazardous outcome. It is quite possible that  $P_2$  may perform a maneuver leading to a collision which would not have occurred if  $P_2$  had had full knowledge of  $P_1$ 's position. Hence, in such a situation, it would be reasonable to undertake a worst-case analysis, by

452 T. BAŞAR AND G. J. OLSDER assuming that one of the ships ( $P_1$ ) tries to avoid and the other one ( $P_2$ ) tries to cause a collision. One can envisage other situations, such as the case in which both ships are aware of each other's position and they both try to prevent a collision, i.e., they cooperate. This case, although not a differential game in the proper sense, can be solved by similar techniques. Roughly speaking, the minmax operation in the differential game case must be replaced by the maxmax operation in the cooperative case. Still another situation arises when one ship holds course and speed (is "standing

on") because of the "rules of the road", whereas the other ship must evade. The dynamic system comprises two ships,  $P_1$  and  $P_2$ , maneuvering on a sea surface which is assumed to be homogeneous, isotropic, unbounded and undisturbed. The kinematic equations to be used are identical to those of the two-cars model (see Example 8.3):  $\dot{x}_i = v_i \cos \theta_i$ ,  $\dot{y}_i = v_i \sin \theta_i$ ,  $\dot{\theta}_i = u_i$ ,  $|u_i| \leq 1$  (8.40)  $r = \sqrt{x^2 + y^2}$ . This simple model may not be (too) unreasonable if only short duration maneuvers are considered, during which the ships cannot change their speeds markedly. Each ship is characterized by two parameters: the maximum angular velocity  $u_i$  and the constant forward speed  $v_i$ . The control of  $P_i$  is  $u_i$ , which is the rudder angle, and which is bounded by  $|u_i(t)| \leq 1$ . Extensions of this model are possible, such as, for instance, the case of variable forward speed (the engine setting will then become a control variable as well), but these would make the analysis to follow much more complicated. Some other effects (hydrodynamic and inertial) have been ignored in the model (8.40), such as the  $u_i$ -dependence of  $v_i^*$ . The cost function is described in terms of the closest point of approach (CPA), the distance corresponding to it is called the miss distance. The terminal condition is given by the CPA at first pass, characterized by  $\frac{dr}{dt} \leq 0$  where  $r = \sqrt{x^2 + y^2}$ , and the terminal range  $r(T)$  is the cost function. This constitutes a game of degree. In this section we shall consider, instead, the directly related game of kind, characterized by a given number  $r_m$ , the minimum range or minimum miss distance. If  $r(T) > r_m$ , no collision takes place, and for  $r(T) < r_m$ , a collision will take place. In the three-dimensional state space, the target set is described by the cylinder  $x^2 + y^2 = r_m^2$ . The UP, i.e., the set of points on this cylinder at which a collision can take place, is determined by  $(v_1 \sin \theta_1 + v_2 \cos \theta_2)^2 - v_2^2 < 0$ .

PURSUIT-EVASION GAMES 453 The terminal condition wherefrom a barrier can be constructed is determined from this relation with the inequality sign replaced by an equality sign. Substituting  $x = r \sin \alpha$  and  $y = r \cos \alpha$ , where  $\alpha$  is the bearing, into this equality, we obtain, at the final time,  $r(T) = \frac{v_1 \sin \alpha(T) + v_2 \cos \alpha(T)}{v_2 \cos \theta_2(T) - v_1 \sin \theta_1(T)}$ , (8.41) where  $w = \sqrt{v_1^2 + v_2^2 - 2v_1v_2 \cos \theta_2(T)}$ , and  $e = \pm 1$ . For each  $\theta_2(T)$  we obtain two values for  $\alpha(T)$ , one corresponding to  $e = +1$  and the other one to  $e = -1$ . These two values correspond to a right and left barrier, respectively;  $P_2$  can just miss  $P_1$ , either on the right or on the left side. From the final conditions thus obtained, we must construct equilibrium trajectories backward in time in order to obtain the barrier which separates the points from where avoidance is possible from the

points from where it is not. The barrier—actually composed of two parts, each corresponding to a different value of  $e$  in (8.41)—is depicted in Fig. 8.15. In this figure the two parts of the barrier intersect, which indicates that the enclosed points are guaranteed collision points for P2 (provided that the intersection of the two parts does not leak). target set Figure 8.15: A schematic picture of the barrier. We now turn to the actual construction of the barriers. The optimal  $u_1$  and  $v_r$  are the arguments of the corresponding Isaacs equation, thus leading to  $u_1^* = \text{sgn}(x\{VX_2 - x_2VX_1 - Ve)$ ,  $u_2^*$  Furthermore, the costate variables satisfy  $\text{sgn}(-Ve) - VX_1 = -u_1|u_1VX_2$ ,  $VX_1(T) = \sin a(T)$ ;  $VX_2 = u_3|u_1Vx^{\wedge}$ ,  $VX_2(T) = \cos a(T)$ ;  $Ve = v_2(VX_2\sin\theta - VX_1\cos\theta)$ ,  $Ve(T) = 0$ . (8.42) (8.43) As long as the arguments of the  $\text{sgn}$ -relations in (8.42) are nonzero, it is simple to obtain the solution of (8.40) and (8.43); but, substitution of the final values

454 T. BAŞAR AND G. J. OLSDER for  $VX_1$ ,  $VX_2$ ,  $Ve$  into (8.43) shows that the arguments of the  $\text{sgn}$ -relations are zero at  $t = T$ . The situation can, however, be saved by replacing the arguments with their retrograde time derivatives, which leads to  $u_1(T) = e \text{sgn}(-\dot{VX}_1(T))$ ,  $u_2(T) = -e \text{sgn}(-\dot{VX}_2(T))$  where  $e = \pm 1$ . Hence, the equilibrium strategies at the terminal situation are determined for almost all values of  $\theta(T)$ . For  $t^{\wedge}/i < 1$ , we can expect a singular control for P2 if  $\theta(T) = \pm \arccos(w_2/i)$  for the last part of the maneuver. Similarly,  $i/v_2 < 1$ , P1 may have a singular control. Figure 8.16: The trajectories which make up the right barrier, projected in the  $(x_2, 0)$  plane. Let us proceed with the case  $V_2 < V < 1$ . Closer scrutiny reveals that there are no equilibrium paths leading to the singular point  $\theta(T) = \arccos(v_2A/i)$  but, in its stead, a dispersal line ends at this point. The other point,  $\theta(T) = -\arccos(w_2/w_1)$ , however, is the end point of a singular line (with  $u_2 = 0$ ), which is a universal line. Integrating backward in time from the final condition, we first start with the singular control  $u_2 = 0$  which may be terminated at any time (switching time) we desire; and going further backward in time we obtain a nonsingular  $u_2$ . In Fig. 8.16, it is shown how the right barrier ( $e = +1$ ) can be constructed by determining the optimal trajectories backward in time. In the figure, the projections of the trajectories on the  $(x_2, 0)$  plane have been drawn. In order to construct these trajectories,  $\theta(T)$  can be taken as a parameter, and for  $\theta(T) = -t \arccos(v_2/i)$  we have another parameter, viz. the time instant at which the singular control switches to a nonsingular one (there are two options:  $u_2 = +1$  and  $u_2 = -1$ ) in retrograde time. The dispersal line is found

as the intersection of two sets of trajectories obtained through backward integration, namely that set of trajectories which have  $\theta(T)$  smaller than  $\arccos(w_2/w_1)$ ,

PURSUIT-EVASION GAMES 455 modulo  $2\pi$ , and the set of trajectories with  $\theta(T)$  greater than  $\arccos(w_2/w_1)$ . The left barrier can be constructed analogously. If the two barriers intersect over the whole range  $0 < \theta < 2\pi - k$ , and if the line of intersection does not leak, then the enclosed points are guaranteed collision points for P2. Also a tacit assumption which makes this whole construction work is that the two parts which define the dispersal line (for both the right and left barriers) do indeed intersect. If they do not intersect, then there would not be a dispersal line; instead, there would be a hole in the surface that encloses the points we described above. Consequently, this set of points would not be enclosed anymore, and P1 might be able to force escape by steering the state through this hole. As yet, no sufficiency conditions are known which ensure enclosure; hence, for each problem one has to verify numerically that no holes exist. For this reason, the method can be used conveniently only for systems with state space dimension not greater than three. Figure 8.17: Intersection of the semipermeable surfaces with the plane  $\theta = \theta_0$ . A possible and appealing way of constructing the barriers which enclose the points wherefrom collision is unavoidable, is to consider cuts in the  $(x_1, x_2, \theta)$  space for which  $\theta$  is a constant, say  $\theta = \theta_0$ . For each parameter value  $\theta(T)$  and also for the retrograde time parameter when  $\theta(T) = -\arccos(w_2/w_1)$  one calculates the equilibrium trajectories until they hit the plane  $\theta = \theta_0$ —this being performed for both the right and left barriers. For different values of the parameters, one obtains, in general, different points in the plane  $\theta = \theta_0$ , with the ultimate picture looking like the one sketched in Fig. 8.17. In this approach existence of a dispersal line cannot be detected; one simply "by-integrates" it. The semipermeable surface (line)  $m$  in Fig. 8.17 is the result of such a procedure. However, this is not a serious problem, since what one seeks in the plane  $\theta = \theta_0$  is a region completely surrounded by semipermeable surfaces; the connecting corners must not leak. Hence, in a picture like Fig. 8.17, one has to check whether point A leaks or not. If it does not, then one can disregard the semipermeable line  $m$ . If it leaks, however, then this is an indication that either the problem is ill-posed (i.e., no enclosed region exists from where collision can be guaranteed), or there must be other semipermeable surfaces which will define, together with (parts of) the already existing semipermeable surfaces, a

456 T. BAŞAR AND G. J. OLSDER Figure 8.18: Initial points  $(x_1, x_2)$ , with  $\theta = \pi/2$ , for which distance of CPA is new enclosed region. The next figure, Fig. 8.18, taken from Miloh and Sharma (1976), depicts several enclosed areas for different  $r_m$ -values. The figure shows the barrier cross-sections at  $\theta = \pi/2$ . The numerical data are as follows:  $v_1 = 1$ ,  $v_2 = 5$ ,  $u = 1$ . This means that the minimum turn radii of P1 and P2 are 1 and  $\frac{1}{5}$ , respectively. By constructing similar pictures for different sets of parameter values (different speed ratios and different maneuvering capabilities), it is possible to gain insight into collision sensitivity with respect to these parameters. As a final remark, collision was defined here as "two ships (considered as points in the plane) approaching each other closer than a given distance". If ships would actually collide this way, then they must have the shape of round disks, as viewed from above. The construction presented can, however, be applied to more realistic shapes, which is though slightly more complicated; such an example will be presented in Section 8.7. Problem 5 (Section 8.8) presents yet another slightly different problem in which the target set is still a round disk, but it is not centered at the origin.

### 8.7 Role Determination and an Application in Aeronautics

Heretofore we have considered the class of pursuit-evasion games wherein the roles of the players are prescribed at the outset of the problem. These problems are sometimes called one-target games which model situations such as a missile chasing an airplane. However, in a "dogfight" that takes place between two planes or ships which are both armed and capable of destroying their opponents, it may not be apparent at the outset who pursues whom. Accordingly, we now introduce two-target games, where either player may be the pursuer or

PURSUIT-EVASION GAMES 457 the evader, depending on their current configuration, and each target set is determined by the shooting range of the respective player. Each player's task is to destroy his opponent (i.e., maneuver the opponent into his target set) and to avoid destruction by his opponent. The solution method to two-target games comprises essentially two stages: (i) Given the initial conditions, determine the relative roles of the players; who pursues whom? It is clear that in a deterministic differential game the roles of the players are determined completely by the initial state and will not change during the course of the game—provided that only pure strategies are permissible and that the players act rationally. (ii) With the roles of the players determined, what are the optimal, e.g., time- optimal, strategies for the

players? Note that the question of who pursues whom has nothing to do with the cost function. Here we shall be concerned with the first stage. The state space consists of three mutually disjoint regions: two regions corresponding to victory by one or the other player, and the third region corresponding to a draw, i.e., neither player is capable of destroying his opponent. The states within the draw region may correspond to different situations. A particular possibility is a stalemate: both players play certain strategies such that neither player will be destroyed and, moreover, a deviation from his equilibrium strategy by one of the players will lead to his own destruction. Another possibility is that the faster player can escape—tacitly assuming that this player cannot win the game in finite time. The region  $R_i$  of the state space, which corresponds to victory by PI, may be expected to be bounded by a surface comprising a part of the surface of  $A_i$ , the target set for  $P_i$ , together with a semipermeable surface  $S_i$  which prevents the state from leaving  $R_i$  provided that PI plays optimally in the neighborhood of  $E_i$ . The interior of  $R_x$  does not contain any points of P2's target set  $A_2$ , and thus, assuming that the game terminates in finite time (i.e., a stalemate cannot arise), victory for PI is guaranteed for initial states inside  $R_i$  (see Fig. 8.19a). That portion of  $A_x$ , which forms part of the boundary of  $R_i$ , is permeable only Figure 8.19: Regions  $R_i$  and  $R_2$  corresponding to victory by PI and P2, respectively.

458 T. BAŞAR AND G. J. OLSDER in the direction away from  $R_i$ . In a similar way we can construct  $R_2$ —the region corresponding to victory by P2. If  $R_i$  and  $R_2$  do not fill up the whole state space, then points belonging to neither  $R_i$  nor  $R_2$  belong to  $R_3$ , the draw region. Such a situation is illustrated schematically in Fig. 8.19b. Victory by P2 in region  $R_2$  is guaranteed if we assume that no stalemate is possible, so that the target set  $A_2$  is reached in finite time. States belonging to both  $S_i$  and  $S_2$  (such as I in Fig. 8.19b) will terminate on both  $A_i$  and  $A_2$ , thereby giving rise to "victory" by both players, or, what is essentially the same, to a simultaneous confrontation. States belonging to  $S_2$  and not to  $S_i$ , or belonging to  $S_i$  and not to  $S_2$ , give rise to near miss situations. We now present an example of a two-target game which has the additional feature that the two target sets are nonsmooth. Consider an aerial duel or dogfight, wherein each combatant wishes to destroy his opponent without himself being destroyed. In order to keep the analysis at a reasonable level, we take the dimension of the state space to be three, and assume that the two players move in a single horizontal plane. Equations of motion within this plane are the same

as those in the ship collision avoidance problem, and they are given by (8.40). The constraints on the controls are also the same. We assume that  $V_1 > v_2 > 0$  and  $0 < u_1 < u_2$ , i.e.,  $P_1$  has the advantage of having a higher speed, whereas  $P_2$  is more maneuverable.  $P_1$  has a confrontation range (which is his target set  $A_1$ ) consisting of a line segment of length  $l_1$  in the direction of his velocity vector. If  $P_1$  crosses the confrontation range of  $P_2$ , or vice versa, the game ends. For the analysis to follow, we make the additional assumption  $0 < \epsilon_1 < \epsilon_2$ . We start with the construction of the barrier  $S$  which is the composition of  $S_1$  and  $S_2$  introduced earlier. Each portion of  $S$ , whether it leads to simultaneous confrontation or to near miss, is obtainable by solving a particular "local" differential game, with a terminal cost function defined only in the immediate neighborhood of the simultaneous confrontation or near miss. If we adopt the convention that the local game has a positive (respectively, negative) value when it ends in favor of  $P_2$  (respectively,  $P_1$ ), the equilibrium controls are determined from  $\min \max (V_{x_1} \pm i + V_{x_2} x_2 + V_s \delta) = 0$ ,  $u_1 \ u_2$  which leads to  $u_1^* = \text{sgn}(x_2 V_{x_1} - x_1 V_{x_2} + V_s)$ ,  $u_2^* = \text{sgn}(V_s)$ . The costate equations are the same as in (8.43), but in this case the terminal conditions are different as will be seen shortly. In fact, the terminal conditions will be different for each local game. From (8.44) it follows that, if  $u_1$  remains constant for a duration of at least  $r$  units of time before the final time  $T$  is reached, then, apart from an insignificant constant multiplier, assumed henceforth to be unity, the costates at time  $r = (8.44)$

PURSUIT-EVASION GAMES 459  $T - t$  satisfy  $V_{x_1}(t) = \cos(\theta - \omega_1 t)$ ,  $V_{x_2}(t) = \sin(\theta - \omega_1 t)$ , (8.45) where  $\theta$  is determined by the final conditions. From (8.44) and (8.43) it further follows that  $V_s(t) = -v_2 \cos(\theta + \theta - \omega_1 t) = -v_2 \cos(\theta(T) + \theta - \omega_1 t)$ , (8.46) so that, if  $u_2 \neq 0$ ,  $V_s(r) = V_s(T) - \int_r^T \omega_1 V_s(t) dt$   $\{ \sin(\theta(T) + \theta - \omega_1 t) - \sin(\theta(T) + \theta) \}$  By introducing  $A = x_2 V_{x_1} - x_1 V_{x_2} + V_s$ , we also find  $A = -V_1 V_{x_1}$ , so that, if  $u_1 \neq 0$ , (8.47)  $A(t) = A(T) + \int_t^T \omega_2 (V_{x_2}(T) - V_{x_2}(t)) dt$ . Since many of the local games will include singular controls, we now investigate their properties. From (8.44) it follows that a singular control for  $P_2$  is only possible for  $V_s = 0$ , which results in  $V_g = 0$ . Equation (8.46) now leads to the conclusion that the only possible singular control for  $P_2$  is  $u_2 = 0$ . Similarly, it can be shown that the only possible singular control for  $P_1$  is  $u_1 = 0$ . Returning to the actual construction of  $E$ , we shall first deal with simultaneous confrontation and thereafter with the near miss situations. For each situation, several types may be distinguished, since



simultaneous confrontation and near miss occur in different ways.

Simultaneous confrontation Consider the situation depicted in Fig. 8.20, where the trajectories of P1 and P2 in real space have been sketched, together with the relative coordinate system at terminal time T. In the figure, P1 turns to the right as fast as possible, and  $y^T > 0$ . Figure 8.20: A situation of simultaneous confrontation.

460 T. BAŞAR AND G. J. OLSDER P2 does likewise, but to the left. The final state is given as follows:  $X(T) = 0$ ,  $\theta(T) = \theta_k$  and  $x_2(T)$  is a variable that must satisfy  $0 < x_2(T) < \min(i_1, i_2)$  (8.48). The equilibrium strategies are  $u_1^* = +1$ ,  $u_2^* = -1$ . Initial conditions associated with these strategies and terminal conditions can easily be found by backward integration of the state equations. The value of the game corresponding to these initial conditions is neither positive (victory for P2) nor negative (victory for P1) and is therefore determined as  $V = 0$ . It should be clear from Fig. 8.20 that  $u_1^*$  and  $u_2^*$  are indeed in equilibrium since a deviation from either strategy will lead to a destruction of the deviating player. This is also confirmed by application of the necessary conditions. Let us assume for the moment that the terminal values of the costate variables satisfy the requirements  $V_{X1}(T) > 0$ ,  $V_{X2}(T) = 0$ ,  $V_\theta(T) < 0$ . (8.49) Then, it follows that  $\dot{V} = 0$  in (8.45), which, together with (8.44) and (8.47), indeed leads to  $u_1^* = +1$ ,  $u_2^* = -1$ . But why should (8.49) be a requirement? Let us consider the case of  $V_{X1}(T)$ : a small deviation of P2's position in the positive  $X$ -direction in the final situation (keeping  $x_2(T) = 0$ ,  $\theta(T) = \theta_k$ ) will be to the advantage of P2, since he will be able to win the game from the new relative position. Hence, after this deviation, we have  $V > 0$  instead of  $V = 0$ , and therefore  $V_{X1}(T) > 0$ . A similar argument verifies the sign of  $V_\theta(T)$ . Furthermore, since a change in  $x_2(T)$  does not change the outcome of the game (it remains  $V = 0$ ), we have  $V_{X2}(T) = 0$ . Let us now investigate how this simultaneous confrontation affects the barrier  $\mathcal{L}$  in the  $(x, x_2, \theta)$  space, which separates the initial points where P1 can win from those corresponding to P2's victory. Since  $x_2(T)$  is a parameter, subject to the constraint (8.48), a whole family of equilibrium trajectories exists, all leading to a simultaneous confrontation of the type described. This family forms a semipermeable surface in the  $(x, x_2, \theta)$  space. For the actual construction we shall consider the intersection of this surface with the planes  $\theta = \text{constant}$ , say  $\theta = \theta_0$ . Equations (8.40), (8.43), (8.44) and (8.49) are integrated backwards in time until  $\theta = \theta_0$  is reached and for varying  $x_2(T)$  we obtain different points in

the  $\theta = \theta_0$  plane, which make up a line segment. In Fig. 8.21, the solution of a specific example has been sketched, for which the characteristics of the airplanes are  $v_1 = u_1 = 1$ ,  $v_2 = -1$ ,  $u_2 = 3$ ,  $i_1 = 2$  and  $i_2 = 00$ . (8.50) In the figure,  $\theta_0 = 7r/3$ . The line segment which separates the states corresponding to victory by  $P_1$  from those corresponding to  $P_2$ 's victory, according to the simultaneous confrontation described above, has number 1 attached to it. From initial positions slightly above this line,  $P_1$  will win, and from initial positions slightly below,  $P_2$  will win.

PURSUIT-EVASION GAMES 461 Figure 8.21: Intersection of semipermeable surfaces with the plane  $\theta = 7r/3$ . Different types of simultaneous confrontation exist, each one corresponding to a different maneuver. These different types are shown in Fig. 8.22. Type 1 has already been discussed extensively. Type 2 is similar, but now  $P_2$  turns right as fast as possible;  $X_2(T)$  is again the parameter. Types 3 and 4 are different in the sense that destruction is not caused by the confrontation range, but instead by collision; the angle  $\theta(T)$  is the parameter. Type 3 is a collision with  $P_1$  turning right and  $P_2$  turning left. In type 4, both turn right. Type 5 calls for an S-maneuver by  $P_1$  to get within the range of  $P_2$ , who has the longer confrontation range, before  $P_2$ , who is turning right, faces directly towards  $P_1$ , at a distance of  $X_2(T) = l$ ; the turn angle  $\alpha$  of the final bend of  $P_1$  is the parameter. Type 6 is an extension of the S-maneuver of  $P_1$  in type 5, in which the final bend with a fixed angle  $\alpha$  is preceded by a straight-line segment (singular arc) whose length functions as a parameter. This singular arc itself is preceded by a left turn. Type 7 is similar to type 6, the only difference being that now  $P_1$  starts with a right turn. In addition to these 7 types, there are 7 other types, indicated by  $1', 2', \dots, 7'$ , obtained from the first 7 types by left-right reflections. For instance, in case  $1'$ ,  $P_1$  turns left and  $P_2$  turns right. In all these types, the parameter is subject to certain constraints. We will not treat detailed analyses of all these types here; they are somewhat technical and not enlightening for the results to be obtained. The details can be found in Olsder and Breakwell (1974). Those types which give rise to curves lying in the  $\theta = 7r/3$  plane have been indicated in Fig. 8.21. A question that comes into mind now is whether these fourteen different types of simultaneous confrontation constitute the whole picture. In fact, we do not know for sure whether there are other types or not. The ones that have

2 type 3  $\lambda(T)$  type 4 type 5 type 6 type 7 Figure 8.22: Different maneuvers, all leading to simultaneous confrontation. been depicted in Fig. 8.21 have been obtained using intuitive reasoning, and the analysis indeed showed that they give rise to semipermeable surfaces. However, together with the semipermeable surfaces corresponding to near miss situations, for which the discussions will follow shortly, they separate (and enclose) the states where from  $P_1$  can guarantee victory from the other states, i.e., there is no hole in  $\mathcal{E}$  (cf. Section 8.6). If, however, another simultaneous confrontation maneuver (or another yet-unknown near miss maneuver) is discovered, it may lead to either a larger or a smaller  $R_i$ . Essentially ten different types of near miss situations have been discovered so far, which are labeled as a, b, ..., j in Fig. 8.23, which depicts the corresponding trajectories in real space. Another set of ten types exists, obtained by interchanging left and right, as in the case of simultaneous confrontation. We now give a brief rundown of the first ten types. In type a both players move to the right as fast as possible;  $P_2$ 's trajectory in relative space looks as sketched in Fig. 8.24, and its characteristic feature is that it touches the boundary of  $P_1$ 's confrontation range, but does not cross it—hence the name "near miss". The parameter can be taken as  $\lambda(T)$ , which

PURSUIT-EVASION GAMES 463 type j Figure 8.23: Different maneuvers, all leading to near misses. satisfies the constraint  $0 < x_2(T) < \min\{\lambda, v_2/u_j\}$  and determines  $\theta(T)$  through  $\sin\theta(T) = u_j x_2(T)/v_2$ . The terminal conditions for the costate equations are  $V_{x_1}(T) > 0$ ,  $V_{x_2}(T) = V_e(T) = 0$ . Type b resembles type a, the difference being that  $u_2^* = -1$ . Types c and d have  $\theta(T) = \pi/2$  and end with a singular arc by  $P_2$ . The length of this arc is the parameter. These two types only occur if  $W_i^i > v_2$ . For  $(J_{iii} < v_2$ , the types c and d do not arise, but instead we have the types e, f, g and h, for which  $x_2(T) = \lambda$ . For types e and h the parameter is  $\theta(T)$ , for f and g it is the length of the singular arc of  $P_2$ 's trajectory. Types i and j resemble e and h, respectively, but now  $P_1$  is almost destroyed at the far end of  $P_2$ 's confrontation range. In e and h it is  $P_2$  who can barely escape the far end of  $P_1$ 's confrontation range. Note that in the figures corresponding to i and j, the relative coordinate system at  $t = T$  has been drawn with respect to  $P_2$  (instead of  $P_1$ , as in all other types). In these two latter types,  $P_2$  has a

T. BAŞAR AND G. J. OLSDER V Figure 8.24: The trajectory corresponding to type a in Fig. 8.23, now in relative space. singular arc, the length of which is

the parameter. In Fig. 8.21, we have drawn the line segments corresponding to the applicable near miss cases for the example with characteristics (8.50), which are the types a, b, c and d and their mirror images a', b', d', d'. The crucial question now is whether all the line segments (both simultaneous confrontation and near miss) enclose a region, together with  $A_i$ . It should be clear from the figure that two such regions exist: an "inner one" and an "outer one". Closer scrutiny reveals that some corners of the outer region leak, and the semipermeable surfaces which enclose the inner region constitute the barrier which separates  $R_i$  from  $R_2$ . To summarize, we have given a complete characterization of the regions leading to victory for  $P_1$  and  $P_2$ , and have arrived at the qualitative result that if  $P_1$  plays optimally along the boundary of his region of victory ( $R_i$ ), then  $P_2$  is captured within  $R_i$  (and likewise with the roles of  $P_1$  and  $P_2$  interchanged).<sup>96</sup> If he cannot stay within  $R_i$  forever, then he has to leave this region via  $A_i$  which leads to termination of the game (with  $P_1$ 's victory).

8.8 Problems 1. Instead of the pursuit-evasion game of Example 8.2, consider the equivalent problem described by  $\dot{x}_i = u^i + u^2$ ,  $x_i(0) > 0$ ,  $\pm 2 = 1 - X_i$ ,  $\dot{x}_2(0) = 0$ ,  $L = x_2(T)$ , where  $T = \min\{t : x_i(t) = 0\}$ ,  $-1 < u^* < 0$ ,  $i = 1, 2$ . Obtain the final condition for the costate equation by making use of (8.14), and show that the result is in complete agreement with the corresponding condition derived in Example 8.2. <sup>96</sup>Note that since  $\lambda_1 = 0$ ,  $R_3$  is empty.

PURSUIT-EVASION GAMES 465

2. Calculate the value function  $V(x)$  for Example 8.5, and show that both  $V$  and  $dV/dx$  are continuous along the universal surface.

3. Consider the "lady in the lake" with an instantaneous informational advantage to  $E$ , i.e., the strategies are of the form  $u^x(t) = \begin{cases} y_1(x(t)), & u^2(t) = -f_2(x(t), \lambda'(t)) \end{cases}$ .  $E$  starts at the center of the pond and spirals outward as fast as possible, while maintaining  $\dot{\theta}(t) = \dot{r}$ . Show that the circle  $r = Rv^2$  will be reached in finite time. Figure 8.25: Geometrical derivation of the saddle-point solution.

4. This problem addresses the derivation of the equilibrium solution of the "lady in the lake" geometrically, without utilizing Thm. 8.2. Start from the decision point  $E_q$  in Fig. 8.25, and suppose that  $P$  chooses  $u^1 = -1$  (i.e., he moves such that the pond is to his left in real space); the case  $u^1 = +1$  can be dealt with analogously.  $E$  will then swim in a straight line to the shore. The question is: in which direction will she swim? Suppose that she swims in the direction (3) (see Fig. 8.25). (i) Why can  $\theta$  be restricted to  $\theta < \pi/2$ ? (ii) Show geometrically that swimming towards  $E_3$  is better for  $E$  than to swim to  $E_2$ , where  $E_3$  is only an  $e$ -distance away from  $E_2$ . (Show that

the time needed for P to cover the distance  $E_2 E_3$  is greater than the time needed for E to cover the distance  $E_2 E_3$ , where  $E_2$  is the projection of  $E_2$  on  $E_0 E_3$ .) (iii) Show that swimming towards  $E_4$  is the best E can do. (Make use of the fact that the circle through the points  $E_4$ ,  $E_q$  and C lies inside the pond.)

5. Consider the following two-cars model:  $\dot{x}_i = v_i \sin \theta_i - a_i u_i$ ,

466 T. BAŞAR AND G. J. OLSDER  $\dot{x}_i = (v_i/2) \cos \theta_i + X_i u_i - 1$ ,  $\dot{\theta}_i = -u_i + 2v_i u_i$ , together with the constraints  $|u_i| < 1$ ,  $i = 1, 2$ . The target set A is defined by  $x_2^2 + (x_2 + R)^2 - R^2 = 0$ , where  $R = \sqrt{2}/8$ .  $P_1$  tries to avoid A, whereas  $P_2$  would like to have the state enter A. (i) Determine the barrier(s) which separate(s) points from where avoidance is possible from those from where it is not, and show that each one lies partly inside A. Therefore, the semipermeable surface determined cannot be the barrier(s) sought. (ii) Determine the actual barrier, in a way similar to that employed in the dolichobrachistochrone problem (see Chigir (1976) or the first edition of Başar and Olsder (1982)). Consider the system described by  $\dot{x}_i = X_2 + 1 + 2\sin u_2$ ,  $\dot{\theta}_i = -3u_1 + 2\cos u_2$ , with  $|u_i| < 1$  and no constraints imposed on  $u_2$ . The target set is the half line  $(x_1 > 0, X_2 = 0)$ .  $P_1$  wants to steer the system to the target set as soon as possible, whereas  $P_2$  wants to do the opposite. The game is played in the half plane  $X_2 > 0$ . (i) Show that a barrier starts at the origin  $(x_1 = 0, X_2 = 0)$  and ends at  $X_2 = 1$ . (ii) Show that the continuation of the barrier, for  $X_2 > 1$ , is an equivocal line. The differential game addressed in this problem is the same as that of Problem 6, except that the dynamics are replaced by  $\dot{x}_i = X_2 + 1 + u_i + 2\sin u_2$ ,  $\dot{\theta}_i = -3u_1 + 2\cos u_2$ , where  $P_1$  chooses  $u_1$  and  $u_2$ , subject to the constraint  $(u_1)^2 + (u_2)^2/e < 1$ , where  $e$  is a small positive number.  $P_2$  chooses, as before,  $u_2$ , without any restrictions. For  $e = 0$  this problem reduces to Problem 6. It is conjectured that the present more general version features a switching envelope. Prove or disprove this conjecture, and also investigate the limiting behavior of the solution when  $e \rightarrow 0$ . The kinematics of this problem are described by (8.22) with  $V_2 = \sqrt{2}$ . The only restriction on the controls is  $|\dot{\theta}_i| < 1$ . The target set of  $P_1$  is given by  $A_1 = \{(x_1, X_2) : x_2^2 + x_1^2 < 4 \text{ and } \arctan \frac{x_2}{x_1} < \frac{\pi}{3}\}$

PURSUIT-EVASION GAMES 467 and the target set of  $P_2$  is given by  $A_2 = \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$ .  $P_1$ 's objective is to force  $(x_1, x_2)$  into  $A_1$  without passing through  $A_2$ , whereas  $P_2$  wants to have the state enter  $A_2$  without passing through  $A_1$ . Show that the region of victory  $R_j$  for  $P_i$  is as indicated in the

following figure: Check also that the semipermeable curve DE lies completely outside A2 and that the corner E does not leak. 9. We are given a zero-sum differential game for which the Isaacs condition does not hold everywhere, but instead the Isaacs equation admits two different solutions depending on the order of the min and max operations. Prove that the associated differential game does not admit a saddle-point solution in feedback strategies, but it admits a saddle point when one or the other player is given an informational advantage on the current action of his opponent. Make use of this result in providing a possible verification for the existence of a saddle point in the differential game "lady in the lake." 8.9 Notes Section 8.2. The theory of deterministic pursuit-evasion differential games was single-handedly created by Isaacs in the early 1950s, which culminated in his book (Isaacs, 1975; first edn. 1965). Blaquiere and Leitmann, independently of Isaacs, also obtained Thms. 8.1 and 8.2 (see Blaquiere et al., 1969). Their geometric approach is essentially the same as that of Isaacs, but it is stated in a mathematically more precise language. In this chapter, we follow essentially Isaacs' approach. Starting in the 1960s, many researchers have worked on a rigorous establishment of the validity of the Isaacs equation. The starting point is often a time-discretization and a carefully defined information structure. Depending on this information structure, lower and upper values of the game are defined, and then appropriate limiting arguments are incorporated, under which these upper and lower values approach each other; the resulting limiting value, in each case, is declared to be the value of the game. References in this connection are Fleming (1961, 1964), Varaiya and Lin (1969), Roxin (1969),

468 T. BAŞAR AND G. J. OLSDER Friedman (1971), Elliott and Kalton (1972), Elliott (1977) and Elliott et al. (1973). The relationship with viscosity solutions is given in Evans and Souganidis (1984) and in Barron, Evans and Jensen (1984). These two references contain results for fixed terminal time problems only; see also Elliott (1987). For results on viscosity solutions for problems with variable terminal time the reader is referred to Bardi and Soravia (1991a, b). For extended solution concepts of ordinary differential equations, in which, for instance, the  $f$ -function in the state equation is not Lipschitz-continuous, the reader is referred to Krasovskii and Subbotin (1988), which also includes a rigorous definition of strategy, and also to Hajek (1979). The notions of playability and nontermination aided in the understanding of the so-called "bang-bang-bang surface" introduced in Isaacs (1969); see also

Ciletti (1970) and Lewin (1976). Section 8.3. The concept of capturability also originates in Isaacs' works. A mathematically more rigorous approach to capturability is given in Krasovskii and Subbotin (1988), leading to the so-called "Theorem of the Alternative". A different approach to capturability has been developed by Pontryagin (1967) and subsequently by Hajek (1975). In this set-up the pursuer knows the current action of the evader, which he "neutralizes" by means of his own control. The left-over power (if any) is used by the pursuer to steer the system to the target set. The famous "homicidal chauffeur game" is very versatile in that its solution features many singular surfaces. We do not provide here the complete solution to this game; it was partly solved in Isaacs (1975). The complete solution is spread out over Merz (1971) and Breakwell (1973). The notion of a leaking corner was introduced in Bernhard (1971). Section 8.4. The transition, dispersal, universal and equivocal lines were first introduced in Isaacs (1975). The focal line and the switching envelope appeared for the first time in Merz (1971) and Breakwell (1973). Introduction of noise in the systems equations tends to "smoothen" the singular surfaces which will therefore no longer exist as such. The addition of noise to the system dynamics is the starting point for viscosity solutions; see Chapter 5. For a "noisification" of the homicidal chauffeur game, see Pachter and Yavin (1979). Section 8.5. The "lady in the lake" appeared in the Russian translation of Isaacs' book; see also Breakwell (1977). Another well-known differential game, which has a safe contact in its solution, is the dolichobrachistochrone problem, stated and partly solved in Isaacs (1975). The complete solution was given later in Breakwell (1971) and Chigir (1976). In the survey article (Ho and Olsder, 1983), a rather exhaustive list of other known solved or partly solved zero-sum differential games is given, including such games as "lion and man", "obstacle tag", "conic surveillance evasion" and the "suicidal pedestrian". References to extensions with "two pursuers and one evader" or "one pursuer and two evaders" are also given there. A recent book in the vein of Isaacs is Lewin (1994); it contains new developments and examples. Section 8.6. The application in maritime ship collision avoidance follows the works of Sharma (1976), and Vincent and Peng (1973). See also Merz (1973) and Olsder and Walter (1978).

PURSUIT-EVASION GAMES 469 Section 8.7. Role determination and its application in aeronautics were first discussed in Olsder and Breakwell (1974). Similar problems, but with different target sets, have been considered

in Merz (1976). For other aeronautical applications see Peng and Vincent (1975) and Vincent et al. (1976). Section 8.8. Problem 5 is from Vincent and Peng (1973). Problem 6 is a special version of a problem considered in Isaacs (1975), Problem 8 is from Getz and Pachter (1980).

This page intentionally left blank

Appendix A Mathematical Review This appendix provides some background material on those aspects of real analysis and optimization which are frequently used in the text; it also serves to introduce the reader to our notation and terminology. For a more detailed exposition on the topics covered here, a standard reference is Luenberger (1969).

A.1 Sets A set  $S$  is a collection of elements. If  $s$  is a member (element) of  $S$ , we write  $s \in S$ ; if  $s$  does not belong to  $S$ , we write  $s \notin S$ . If  $S$  contains a finite number of elements, it is called a finite set, otherwise it is called an infinite set. If the number of elements of an infinite set is countable (i.e., if there is a one-to-one correspondence between its elements and positive integers), we say that it is a denumerable (countable) set, otherwise it is a nondenumerable set. A set  $S$  with some specific structure attached to it is called a space, and it is called a linear (vector) space if this specific structure is of algebraic nature with certain well-known properties which we assume that the reader is familiar with. If  $S$  is a vector space, a subset of  $S$  which is also a vector space is called a subspace. An example of a vector space is the  $n$ -dimensional Euclidean space (denoted by  $R^n$ ) each element of which is determined by  $n$  real numbers. An  $iSR^n$  can either be written as a row vector  $x = (x_1, \dots, x_n)$ , where  $x_1, \dots, x_n$  are real numbers and denote the components of  $x$ , or as a column vector which is the "transpose" of  $(x_1, \dots, x_n)$  (written as  $x = (x_i, \dots, x_n)^T$ ). We shall adopt the latter convention in this text, unless indicated otherwise.

Linear independence Given a finite set of vectors  $S_1, \dots, S_n$  in a vector space  $S$ , we say that this set of vectors is linearly independent if the equation  $\sum_{i=1}^n a_i S_i = 0$  implies  $a_i = 0, i = 1, \dots, n$ . Furthermore, if every element of  $S$  can be written as a linear combination of these vectors, we say that this set of vectors generates  $S$ . Now, if  $S$  is generated by such a linearly independent finite set (say,  $X$ ), it is said to

472 T. BAŞAR AND G. J. OLSDER be finite dimensional with its unique "dimension" being equal to the number of elements of  $X$ ; otherwise,  $S$  is infinite dimensional. A. 2 Normed Linear (Vector) Spaces A normed linear



vector space is a linear (vector) space  $S$  which has some additional structure of topological nature. This structure is induced on  $S$  by a real-valued function which maps each element  $u \in S$  into a real number  $\|u\|$  called the norm of  $u$ . The norm satisfies the following three axioms: (1)  $\|u\| > 0 \forall u \in S$ ;  $\|u\| = 0$  if, and only if,  $u = 0$ - (2)  $\|u + v\| \leq \|u\| + \|v\|$  for each  $u, v \in S$ - (3)  $\|a u\| = |a| \cdot \|u\| \forall a \in \mathbb{R}$  and for each  $u \in S$ . Convergent sequences, Cauchy sequence An infinite sequence of vectors  $\{s_i, s_2, \dots, \text{st...}\}$  in a normed vector space  $S$  is said to converge to a vector  $s$  if, given an arbitrary  $\epsilon > 0$ , there exists an  $N$ , which may depend on  $\epsilon$ , such that  $\|s - s_j\| < \epsilon$  for all  $i > N$ . In this case, we write  $s_i \rightarrow s$ , or  $\lim_{i \rightarrow \infty} s_i = s$ , and call  $s$  the limit point of the sequence  $\{s_i\}$ . More generally, a point  $s$  is said to be a limit point of an infinite sequence  $\{s_j\}$  if the latter has an infinite subsequence  $\{s_{i_k}\}$  that converges to  $s$ . An infinite sequence  $\{s_j\}$  in a normed vector space is said to be a Cauchy sequence if, given an  $\epsilon > 0$ , there exists an  $N$  such that  $\|s_n - s_m\| < \epsilon$  for all  $n, m > N$ . A normed vector space  $S$  is said to be complete, or a Banach space, if every Cauchy sequence in  $S$  is convergent to an element of  $S$ - Open, closed and compact sets Let  $S$  be a normed vector space. Given an  $s \in S$  and an  $\epsilon > 0$ , the set  $N_\epsilon(s) = \{x \in S : \|x - s\| < \epsilon\}$  is said to be an  $\epsilon$ -neighborhood of  $s$ . A subset  $X$  of  $S$  is open if, for every  $x \in X$ , there exists an  $\epsilon > 0$  such that  $N_\epsilon(x) \subset X$ . A subset  $X$  of  $S$  is closed if its complement in  $S$  is open; equivalently,  $X$  is closed if every convergent sequence in  $X$  has its limit point in  $X$ . Given a set  $X \subset S$ , the largest subset of  $X$  which is open is called the interior of  $X$  and denoted as  $\text{int } X$ . A subset  $X$  of a normed vector space  $S$  is said to be compact if every infinite sequence in  $X$  has a convergent subsequence whose limit point is in  $X$ . If  $X$  is finite dimensional, compactness is equivalent to being closed and bounded. Transformations and continuity A mapping  $f$  of a vector space  $S$  into a vector space  $T$  is called a transformation or a function, and is written symbolically  $f : S \rightarrow T$  or  $y = f(x)$ , for  $x \in S, y \in T$ ;  $f$  is said to be a functional if  $T = \mathbb{R}$ .

APPENDIX A 473 Let  $f : S \rightarrow T$ , where  $S$  and  $T$  are normed linear spaces. The mapping  $f$  is said to be continuous at  $x_0 \in S$  if, for every  $\epsilon > 0$ , there exists a  $\delta > 0$  such that  $f(x) \in N_\epsilon(f(x_0))$  for every  $x \in N_\delta(x_0)$ . If  $f$  is continuous at every point of  $S$  it is said to be continuous everywhere or, simply, continuous.

A. 3 Matrices An  $(m \times n)$  matrix  $A$  is a rectangular array of numbers, called elements or entries, arranged in  $m$  rows and  $n$  columns. The element in the  $i$ th row and  $j$ th column of  $A$  is denoted by a subscript  $ij$ , such as  $a_{ij}$  or  $[A]_{ij}$ , in

which case we write  $A = \{a^{ij}\}$ . A matrix is said to be square if it has the same number of rows and columns; an  $(n \times n)$  square matrix  $A$  is said to be an identity matrix if  $a_{ii} = 1, i = 1, \dots, n$ , and  $a_{ij} = 0, i \neq j, i, j = 1, \dots, n$ . An  $(n \times n)$  identity matrix will be denoted by  $I_n$  or, simply, by  $I$  whenever its dimension is clear from the context. The transpose of an  $(m \times n)$  matrix  $A$  is the  $(n \times m)$  matrix  $A'$  with elements  $a'_{ij} = a_{ji}$ . A square matrix  $A$  is symmetric if  $A = A'$ ; it is nonsingular if there is an  $(n \times n)$  matrix, called the inverse of  $A$ , denoted by  $A^{-1}$ , such that  $A^{-1}A = I = AA^{-1}$ . Eigenvalues and quadratic forms If  $A$  is a square matrix, a scalar  $\lambda$  and a nonzero vector  $x$  satisfying the equation  $Ax = \lambda x$  are said to be, respectively, an eigenvalue and an eigenvector of  $A$ . A square symmetric matrix  $A$  whose eigenvalues are all positive (respectively, nonnegative) is said to be positive definite (respectively, nonnegative definite or positive semidefinite). An equivalent definition is as follows. A symmetric  $(n \times n)$  matrix  $A$  is said to be positive definite (respectively, nonnegative definite) if  $x'Ax > 0$  (respectively,  $x'Ax \geq 0$ ) for all nonzero  $x \in \mathbb{R}^n$ . The matrix  $A$  is said to be negative definite (respectively, nonpositive definite) if the matrix  $(-A)$  is positive (respectively, nonnegative) definite. We symbolically write  $A > 0$  (respectively,  $A \geq 0$ ) to denote that  $A$  is positive (respectively, nonnegative) definite.

**A.4 Convex Sets and Functionals** A subset  $C$  of a vector space  $S$  is said to be convex if for every  $u, v \in C$  and every  $a \in [0, 1]$ , we have  $au + (1 - a)v \in C$ . A functional  $f : C \rightarrow \mathbb{R}$  defined over a convex subset  $C$  of a vector space  $S$  is said to be convex if, for every  $u, v \in C$  and every scalar  $a \in [0, 1]$ , we have  $f(au + (1 - a)v) \leq af(u) + (1 - a)f(v)$ . If this is a strict inequality for every  $a \in (0, 1)$ , then  $f$  is said to be strictly convex. The functional  $f$  is said to be concave if  $(-f)$  is convex, and strictly concave if  $(-f)$  is strictly convex. A functional  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be differentiable if, with  $x = (x_1, \dots, x_n)' \in \mathbb{R}^n$ , the partial derivatives of  $f$  with respect to the components of  $x$  exist, in

474 T. BAŞAR AND G. J. OLSDER which case we write  $\nabla f(x) = [df(x)/dx_1, \dots, df(x)/dx_n]$ .  $\nabla f(x)$  is called the gradient of  $f$  at  $x$  and is a row vector. We shall also use the notation  $f_x(x)$  or  $df(x)/dx$  to denote the same quantity. If we partition  $x$  into two vectors  $y$  and  $z$  of dimensions  $n_1$  and  $n_2 - n_1$ , respectively, and are interested only in the partial derivatives of  $f$  with respect to the components of  $y$ , then we use the notation  $\nabla_y f(y, z)$  or  $df\{y, z\}/dy$  to denote this partial gradient. Let  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  be a vector-valued function whose components are differentiable with respect to the components of  $x \in \mathbb{R}^n$ .

Then, we say that  $g(x)$  is differentiable, with the derivative  $dg(x)/dx$  being an  $(m \times n)$  matrix whose  $ij$ th element is  $dg_i(x)/dx_j$ . (Here  $\langle j \rangle$  denotes the  $i$ th component of  $g$ .) The gradient  $\nabla f(x)$  being a vector, its derivative (which is the second derivative of  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ) will thus be an  $(n \times n)$  matrix, assuming that  $f(x)$  is twice continuously differentiable in terms of the components of  $x$ . This matrix, denoted by  $\nabla^2 f(x)$ , is symmetric, and is called the Hessian matrix of  $f$  at  $x$ . This Hessian matrix is nonnegative definite for all  $x \in \mathbb{R}^n$  if, and only if,  $f$  is convex.

**Separating hyperplane theorem** Given a convex set  $S$ , one can always find a (hyper)plane such that  $S$  lies on one side of it. This result, which should be intuitive in finite-dimensional spaces, also holds in infinite-dimensional linear vector spaces, and is referred to as support theorem or separating hyperplane theorem. Toward giving a precise statement for the most general case, let us first define what a hyperplane is: A hyperplane  $H$  in a linear vector space  $S$  is a linear variety  $H$  (that is, translation of a subspace of  $S$ ), with the property that  $H \cap S = H$ , and if  $V$  is any linear variety containing  $H$ , then either  $V = S$  or  $V = H$ . Then the separating hyperplane theorem says that if  $C$  is a convex set (as a subset of  $S$ ) with nonempty interior, and  $x$  a point in  $S$  and not an interior point of  $C$ , there is a closed hyperplane  $H$  containing  $x$  such that  $H$  lies on one side of  $C$ . The point  $x$  can of course be chosen as a boundary point of  $C$ .

**A.5 Optimization of Functional** Given a functional  $f : S \rightarrow \mathbb{R}$ , where  $S$  is a vector space, and a subset  $X \subset S$ , by the optimization problem minimize  $f(x)$  subject to  $x \in X$  we mean the problem of finding an element  $x^* \in X$  (called a minimizing element or an optimal solution) such that  $f(x^*) \leq f(x) \forall x \in X$ . If such an  $x^* \in X$  exists, then we use the notation  $x^* = \arg \min_{x \in X} f(x)$ . This is sometimes also referred to as a globally minimizing solution, in order

APPENDIX A 475 to differentiate it from the other alternative—a locally minimizing solution. An element  $x^0 \in X$  is called a locally minimizing solution if we can find an  $\epsilon > 0$  such that  $f(x^0) \leq f(x) \forall x \in N_\epsilon(x^0) \cap X$ , i.e., we compare  $f(x^0)$  with values of  $f(x)$  in that part of a certain  $\epsilon$ -neighborhood of  $x^0$ , which lies in  $X$ . For a given optimization problem, it is not necessary that an optimal solution exists; an optimal solution will exist if the set of real numbers  $\{f(x) : x \in X\}$  is bounded below and there exists  $x^* \in X$  such that  $\inf\{f(x) : x \in X\} = f(x^*)$ , in which case we write  $f(x^*) = \inf_{x \in X} f(x)$ . If such an  $x^*$  cannot be found, even though  $\inf\{f(x) : x \in X\}$  is finite, we simply say that an optimal solution does not exist; but we declare the quantity  $\inf\{f(x) : x \in X\}$  or  $\inf_{x \in X} f(x)$  as the optimal value of the optimization problem. If  $\{f(x) : x \in X\}$  is not

bounded below, i.e.,  $\inf_{x \in X} f(x) = -\infty$ , then neither an optimal solution nor an optimal value exists. An optimization problem which involves maximization instead of minimization may be converted into a minimization problem by simply replacing  $f$  by  $-f$ . Any optimal solution of this minimization problem is also an optimal solution for the initial maximization problem, and the optimal value of the latter, denoted  $\sup_{x \in X} f(x)$ , is equal to minus the optimal value of the former. If a maximizing element  $x^* \in X$  exists, then  $\sup_{x \in X} f(x) = \max_{x \in X} f(x) = f(x^*)$ . Existence of optimal solutions In the minimization problem formulated above, an optimal solution exists if  $X$  is a finite set, since then there is only a finite number of comparisons to make. If  $X$  is not finite, however, existence of an optimal solution is not always guaranteed; it is guaranteed if  $f$  is continuous and  $X$  is compact—a result known as the Weierstrass theorem. For the special case when  $X$  is finite dimensional, we should recall that compactness is equivalent to being closed and bounded. Necessary and sufficient conditions for optimality Let  $S = \mathbb{R}^n$ , and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a differentiable function. If  $X$  is an open set, a first-order necessary condition for an optimal solution to satisfy is  $\nabla f(x^*) = 0$ . If, in addition,  $f$  is twice continuously differentiable on  $\mathbb{R}^n$ , a second-order necessary condition is  $\nabla^2 f(x^*) > 0$ .

476 T. BARSAN AND G. J. OLSDER The pair of conditions  $\{\nabla f(x^*) = 0, \nabla^2 f(x^*) > 0\}$  is sufficient for  $x^* \in X$  to be a locally minimizing solution. These conditions are also sufficient for global optimality if, in addition,  $X$  is a convex set and  $f$  is a convex functional on  $X$ . These results, by and large, hold also for the case when  $S$  is infinite dimensional, but then one has to replace the gradient vector and the Hessian matrix by first and second Gateaux (or Frechet) derivatives, and the positive-definiteness requirement by "strong positiveness" of an operator. See Luenberger (1969) for these extensions.

Appendix B Some Notions of Probability Theory This appendix briefly presents some notions of probability theory which are used in the text. For more complete exposition the reader should consult standard texts on probability theory, such as Feller (1971), Papoulis (1965), Ash (1972) and Loeve (1963). In two of the sections of this book (viz. Sections 5.3 and 6.7) we have used some material on stochastic processes and stochastic differential equations which is, however, not covered in this appendix; for this, the reader should consult Wong and Hajek (1985), Gikhman and Skorohod (1972),

Fleming and Rishel (1975), Fleming and Soner (1993) and the references cited in Section 6.7. B.1 Ingredients of Probability Theory Let  $Q$  denote a set whose elements are the outcomes of a random experiment. This random experiment might be the toss of a coin (in which case  $Q$  has only two elements) or selection of an integer from the set  $[0, \infty)$  (in which case  $Q$  is countably infinite) or the continuous roulette wheel which corresponds to a nondenumerable  $Q$ . Any subset of  $Q$  on which a probability measure can be defined is known as an event. Specifically, if  $F$  denotes the class of all such events (i.e., subsets of  $Q$ ), then it has the following properties: (1)  $\emptyset \in F$ . (2) If  $A \in F$ , then its complement  $A^c = \{\omega \in Q : \omega \notin A\}$  also belongs to  $F$ . (The empty set,  $\emptyset$ , being the complement of  $Q$ , also belongs to  $F$ .) (3) If  $A_1, A_2 \in F$ , then  $A_1 \cap A_2$  and  $A_1 \cup A_2$  also belong to  $F$ . (4) If  $A_1, A_2, \dots, A_n, \dots$  denote a countable number of events, the countable intersection  $\bigcap_{j=1}^{\infty} A_j$  and the countable union  $\bigcup_{j=1}^{\infty} A_j$  are also events (i.e., they belong to  $F$ ). 477

478 T. BASAR AND G. J. OLSDER The class  $F$ , thus defined, is called a sigma algebra ( $\sigma$ -algebra) and a probability measure  $V$  is a nonnegative functional defined on the elements of this  $\sigma$ -algebra. The probability measure  $V$  also satisfies the following axioms: (1) For every event  $A \in F$ ,  $0 \leq V(A) \leq 1$ , and  $V(Q) = 1$ . (2) If  $A_1, A_2 \in F$  and  $A_1 \cap A_2 = \emptyset$  (i.e.,  $A_1$  and  $A_2$  are disjoint events),  $V(A_1 \cup A_2) = V(A_1) + V(A_2)$ . (3) Let  $\{A_i\}$  denote a (countably) infinite sequence in  $F$ , with the properties  $A_{i+1} \subset A_i$  and  $\bigcap_{j=1}^{\infty} A_j = \emptyset$ . Then, the limit of the sequence of real numbers  $\{V(A_i)\}$  is zero (i.e.,  $\lim_{i \rightarrow \infty} V(A_i) = 0$ ). The triple  $(Q, F, V)$  defined above is known as a probability space, while the pair  $(\Omega, F)$  is called a measurable space. If  $\Omega = \mathbb{R}^n$ , then its subsets of interest are the  $n$ -dimensional rectangles, and the smallest  $\sigma$ -algebra generated by these rectangles is called the  $n$ -dimensional Borel  $\sigma$ -algebra and is denoted  $B_n$ . Elements of  $B_n$  are Borel sets, and the pair  $(\mathbb{R}^n, B_n)$  is a Borel (measurable) space. A probability measure defined on this space is known as a Borel probability measure. Finite and countable probability spaces If  $\Omega$  is a finite set (say,  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ ), we can assign probability weights on individual elements of  $\Omega$ , instead of on subsets of  $\Omega$ , in which case we write  $p_j$  to denote the probability of the single event  $\{\omega_j\}$ . We call the  $n$ -tuple  $(p_1, p_2, \dots, p_n)$  a probability distribution over  $\Omega$ . Clearly, we have the restriction  $0 < p_i < 1$ ,  $\sum_{i=1}^n p_i = 1$ , and furthermore as the elements of  $Q$  are all disjoint events, we have the property  $\sum_{j=1}^n p_j = 1$ . The same convention applies when  $\Omega$  is a countable set (i.e.,  $\Omega = \{\omega_1, \omega_2, \dots, \omega_j, \dots\}$ ), in which case we simply replace

n by  $\infty$ . B.2 Random Vectors Let  $(\Omega_1, \mathcal{F}_1)$  and  $(\Omega_2, \mathcal{F}_2)$  be two measurable spaces and  $f$  be a function defined from the domain set  $\Omega_1$  into the range set  $\Omega_2$ . If for every  $A \in \mathcal{F}_2$  we have  $f^{-1}(A) = \{\omega \in \Omega_1 : f(\omega) \in A\} \in \mathcal{F}_1$ , then  $f$  is said to be a measurable function, or a measurable transformation from  $(\Omega_1, \mathcal{F}_1)$  into  $(\Omega_2, \mathcal{F}_2)$ . If the latter measurable space is a Borel space, then  $f$  is said to be a Borel function, in which case we denote it by  $x$ . For the special case when the Borel space is  $(\Omega_2, \mathcal{F}_2) = (\mathbb{R}^n, \mathcal{B}_n)$ , the Borel function  $x$  is called a random variable. For the case when  $(\Omega_2, \mathcal{F}_2) = (\mathbb{R}^n, \mathcal{B}_n)$ ,  $x$  is known as an  $n$ -dimensional random vector. If there is a probability measure  $V$  defined on  $(\Omega_1, \mathcal{F}_1)$ —which we henceforth write simply as  $(\Omega, \mathcal{F})$ —then the random vector  $x$  will induce a probability measure  $P_x$  on the Borel space  $(\mathbb{R}^n, \mathcal{B}_n)$ , so that for every  $B \in \mathcal{B}_n$  we have  $P_x(B) = V(x^{-1}(B))$ . Since every element of  $\mathcal{B}_n$  is an  $n$ -dimensional rectangle, the arguments of  $P_x$  are in general infinite sets; however, considering the

APPENDIX B 479 collection of sets  $\{x \in \mathbb{R}^n : a_i < x_i \leq b_i, i = 1, \dots, n\}$  in  $\mathcal{B}_n$  where  $a_i, b_i (i = 1, \dots, n)$  are real numbers, restriction of  $P_x$  to this class is also a probability measure whose argument is now a finite set. We denote this probability measure by  $P_x(a, b) = P_x(a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n)$  and call it a probability distribution function of the random vector  $x$ . Note that  $P_x(a_1, a_2, \dots, a_n) = V\{\omega : x_1(\omega) < a_1, x_2(\omega) < a_2, \dots, x_n(\omega) < a_n\}$ , where  $x_i$  is a random variable denoting the  $i$ th component of  $x$ . Whenever  $n > 1$ ,  $P_x$  is sometimes also called the cumulative (joint) probability distribution function. It is a well-established fact that there is a one-to-one correspondence between  $P_x$  and  $V_x$  and the subspace on which  $P_x$  is defined can generate the whole  $\mathcal{B}_n$  (cf. Loeve, 1963). Independence Given the probability distribution function of a random vector  $x = (x_1, \dots, x_n)'$  the (marginal) distribution function of each random variable  $x_i$  can be obtained from  $P_{x_i}(a_i) = \lim_{b_i \rightarrow \infty} P_x(a_i, \dots, a_n, b_1, b_2, \dots, b_n)$ . The random variables  $x_1, \dots, x_n$  are said to be (statistically) independent if  $P_x(a_1, \dots, a_n) = P_{x_1}(a_1)P_{x_2}(a_2) \dots P_{x_n}(a_n)$ , for all scalars  $a_1, \dots, a_n$ . Probability density function A measure defined on subintervals of the real line and which equals the length of the corresponding subinterval(s) is called a Lebesgue measure. It assigns zero weight to countable subsets of the real line, and its definition can readily be extended to  $n$ -dimensional rectangles in  $\mathbb{R}^n$ . Let  $V$  be a Borel probability measure on  $(\mathbb{R}^n, \mathcal{B}_n)$  such that any element of  $\mathcal{B}_n$  which has a Lebesgue measure of zero has also a  $P$ -measure of zero; then we say that  $V$  is absolutely continuous with respect to the Lebesgue measure. Now, a well-established

result of probability theory says that (cf. Loeve, 1963), if  $x : (Q, F, P) \rightarrow (R^n, B_n, V_x)$  is a random vector and if  $V_x$  is absolutely continuous with respect to the Lebesgue measure, there exists a nonnegative Borel function  $p_x(-)$  such that, for every  $A \in B_n$ ,  $V_x\{A\} = \int_A p_x(\xi) d\xi$ . Such a function  $p_x(-)$  is called the probability density function of the random vector  $x$ . In terms of the distribution function  $P_x$ , the preceding relation can be written as  $\int_{-\infty}^{\infty} p_x(\xi) d\xi = 1$  for every scalar  $a_1, \dots, a_n$ .

480 T. BAŞAR AND G. J. OLSDER B.3 Integrals and Expectation Let  $x : (Q, F, V) \rightarrow (R^n, B_n, V_x)$  be a random vector and  $f : (R^n, B_n) \rightarrow (R^m, B_m)$  be a nonnegative Borel function. Then,  $f \circ x$  can also be considered as a random vector from  $(Q, F)$  into  $(R^m, B_m)$ , and its average value (expected value) is defined either by  $\int_Q f(x(u)) V(du)$  or by  $\int p_x(\xi) f(\xi) d\xi$  depending on which interpretation we adopt. Both of these integrals are well defined and are uniquely equal in value. If  $f$  changes signs, then we take  $f = f^+ - f^-$  where both  $f^+$  and  $f^-$  are nonnegative, and write the expected value of  $f \circ x$  as  $E[f(x)] = \int_Q f(x(u)) V(du) - \int_Q f^-(x(u)) V(du)$ , provided that at least one of the pair  $E[f^+(x)]$  and  $E[f^-(x)]$  is finite. Since, by definition,  $V_x(dQ) = V_x(\xi + d\xi) - V_x(\xi)$ , this integral can further be written as  $E[f(x)] = \int f(\xi) p_x(\xi) d\xi$  which is a Lebesgue-Stieltjes integral and which is the convention that we shall adopt. For the special case when  $f(x) = x$  we have  $E[x] = \int \xi dV_x(\xi) = \int \xi p_x(\xi) d\xi$  which is known as the mean (expected) value of  $x$ . The covariance of the  $n$ -dimensional random vector  $x$  is defined as  $E[(x - E[x])(x - E[x])'] = \int (x - E[x])(x - E[x])' dV_x(x) = \text{cov}(x)$  which is a nonnegative definite matrix of dimension  $(n \times n)$ . Now, if  $V_x$  is absolutely continuous with respect to the Lebesgue measure,  $E[f]$  can equivalently be written, in terms of the corresponding density function  $p_x$ , as  $E[f] = \int f(x) p_x(x) dx$ . If  $Q$  consists only of a finite number of disjoint events  $W_1, W_2, \dots, W_n$ , then the integrals are all replaced by the single summation  $\sum_{j=1}^n f(x_j) P_j$  where  $P_j$  denotes the probability of occurrence of event  $W_j$ . For a countable set  $\{u_i\}$ , we have the counterpart:  $E[f(x(u))] = \sum_{i=1}^{\infty} f(x(u_i)) P_i$ .

APPENDIX B 481 B.4 Norms and the Cauchy-Schwarz Inequality Given a random vector  $x : (Q, F, V) \rightarrow (R^n, B_n, V_x)$ , the quantity defined by  $\|x\|_2 = \{E[x'x]\}^{1/2}$  provided that it is finite, is called the  $(\xi^2)$  norm of  $x$ , and indeed satisfies all the properties of a norm introduced in the previous appendix. A random vector with a finite  $\xi^2$ -norm is called a second-order random vector. Note that a second-order random vector has a well-defined covariance. If  $x$

and  $y$  are two  $n$ -dimensional second-order random vectors defined on the same probability space, then we have the following useful inequality, known as the Cauchy-Schwarz inequality:  $E\{x'y\} \leq \|x\| \|y\|$ . The inequality here is an equality if, and only if,  $x = Ay$ , w.p. 1 for some scalar  $A$ , or  $y = 0$  w.p. 1.

This page intentionally left blank

Appendix C Fixed Point Theorems In this appendix we give, without proof, three theorems (on fixed points) which have been used in Chapters 3 and 4 of the text. Proofs of these theorems are rather lengthy and can be found in the references cited. Some general references for these as well as other results on fixed points are (Istratescu, 1981), (Joshi and Bose, 1985), and (Smart, 1974). Fixed point theorems Theorem C.1 (Brouwer fixed point theorem) If  $S$  is a compact and convex subset of  $R^n$  and  $f$  is a continuous function mapping  $S$  into itself, then there exists at least one  $x \in S$  such that  $f(x) = x$ . Several proofs exist for this fixed point theorem, one of the most elementary ones being given by Kuga (1974). Its original version has appeared in 1910 (Brouwer, 1910). A generalization of this theorem is the Kakutani fixed point theorem (Kakutani, 1941) given below. Definition C.1 (Upper semicontinuity) Let  $f$  be a function defined on a normed linear space  $X$ , and associating with each  $x \in X$  a subset  $f(x)$  of some (other) normed linear space  $Y$ . Then,  $f$  is said to be upper semicontinuous (use) at a point  $x_0 \in X$  if, for any sequence  $\{x_i\}$  converging to  $x_0$  and any sequence  $\{y_i \in f(x_i)\}$  converging to  $y_0$ , we have  $y_0 \in f(x_0)$ . The function  $f$  is upper semicontinuous if it is use at each point of  $X$ . Theorem C.2 (Kakutani) Let  $S$  be a compact and convex subset of  $R^n$ , and let  $f$  be an upper semicontinuous function which assigns to each  $x \in S$  a closed and convex subset of  $S$ . Then there exists some  $x \in S$  such that  $x \in f(x)$ . Another generalization of the Brouwer fixed point theorem is due to Schauder (1930), and involves the space of real-valued bounded continuous functions on a subset  $S$  of  $R^n$ , to be denoted  $C(S)$ . Here, we will first need the definition of equicontinuity: 483

484 T. BAŞAR AND G. J. OLSDER Definition C.2 (Equicontinuity) Let the space of all real-valued bounded continuous functions on  $S$ , denoted  $C(S)$ , be endowed with the sup norm. A subset  $F$  of  $C(S)$  is equicontinuous if for every  $\epsilon > 0$  there exists a  $\delta > 0$  such that  $\|x - y\| < \delta$  implies  $\|f(x) - f(y)\| < \epsilon$ ,  $\forall f \in F$ . Theorem C.3 (Schauder) Let  $S$  be a bounded subset of  $R^n$ , and let  $C(S)$  be the space of real-valued bounded continuous functions on  $S$ , endowed with the sup



norm. Let  $F \subset C(S)$  be nonempty, closed, bounded and convex. Then if the mapping  $T : F \rightarrow F$  is continuous and the family  $T(F)$  is equicontinuous,  $T$  has a fixed point in  $F$ . A proof of this theorem can be found in (Istratescu, 1981) and (Hutson and Pym, 1980). This is a special case of a more general theorem (also due to Schauder) which says that "every real-valued continuous function mapping a convex compact subspace of a Banach space into itself has a fixed point." The space  $C(S)$  defined above is indeed one such space, and equicontinuity assures that convex compact subsets of  $C(S)$  are mapped into themselves. For interesting discussions on the Brouwer fixed point theorem and its various extensions, see Franklin (1980).

Bibliography S. Alpern, Games with repeated decisions, *SIAM Journal on Control and Optimization*, 26 (1988), pp. 468-477. S. Alpern, Cycles in extensive form perfect information games, *Journal of Mathematical Analysis and Applications*, 159 (1991), pp. 1-17. B. Anderson and J. B. Moore, *Optimal Control*, Prentice-Hall, Englewood Cliffs, NJ, 1989. R. B. Ash, *Real Analysis and Probability*, Academic Press, New York, NY, 1972. J. AUBIN, *Mathematical Methods of Game and Economic Theory*, North-Holland, Amsterdam, The Netherlands, 1980. R. J. Aumann, Borel structures for function spaces, *Illinois Journal of Mathematics*, 5 (1961), pp. 614-630. , Mixed and behavior strategies in infinite extensive games, in *Advances in Game Theory*, M. Dresher, L. S. Shapley, and A. W. Tucker, eds., Princeton University Press, Princeton, NJ, 1964, pp. 627-650. R. J. Aumann and M. Maschler, Some thoughts on the minimax principle, *Management Science*, 18 (1972), pp. 54-63. A. Bagchi and G. J. Olsder, Linear stochastic pursuit evasion games, *Journal of Applied Mathematics and Optimization*, 7 (1981), pp. 95-123. A. V. Balakrishnan, *Applied Functional Analysis*, Springer-Verlag, New York, NY, 1976. R. B. Bapat and T. Raghavan, *Nonnegative Matrices and Applications*, Cambridge University Press, Cambridge, England, 1997. M. Bardi, A boundary value problem for the minimum-time function, *SIAM J. Control and Optimization*, 27 (1989), pp. 776-785. M. Bardi and C. Sartori, Convergence results for Hamilton-Jacobi equations in variable domains, *Differential and Integral Equations*, 5 (1992), pp. 805-816. M. Bardi and P. SORAVIA, Hamilton-Jacobi equations with singular boundary conditions on a free boundary and applications to differential games, 485

325 (1991), pp. 205- 229. M. Bardi and V. Staicu, The Bellman equation for time-optimal control of non-controllable nonlinear systems, Tech. Report SISSA 132/91/M, SISSA- ISAS, Strada Costiera 11, 34014 Trieste, Italy, 1991. E. N. Barron, L. C. Evans, and R. Jensen, Viscosity solutions of Isaacs' equations and differential games with Lipschitz controls, *Journal of Differential Equations*, 53 (1984), pp. 213-233. T. BaşAR, On the relative leadership property of Stackelberg strategies, *Journal of Optimization Theory and Applications*, 11 (1973), pp. 655-661. , A counter example in linear-quadratic games: Existence of non-linear Nash solutions, *Journal of Optimization Theory and Applications*, 14 (1974), pp. 425-430. , Nash strategies for M-person differential games with mixed information structures, *Automatica*, 11 (1975), pp. 547-551. , On the uniqueness of the Nash solution in linear-quadratic differential games, *International Journal of Game Theory*, 5 (1976a), pp. 65-90. , Some thoughts on saddle-point conditions and information structures in zero-sum differential games, *Journal of Optimization Theory and Applications*, 18 (1976b), pp. 165-170. , Existence of unique equilibrium solutions in nonzero-sum stochastic differential games, in *Differential Games and Control Theory II*, E. O. Roxin, P. T. Liu, and R. Sternberg, eds., Marcel Dekker, Inc., 1977a, pp. 201-228. , Informationally nonunique equilibrium solutions in differential games, *SIAM Journal on Control and Optimization*, 15 (1977b), pp. 636-660. , Multicriteria optimization of linear deterministic and stochastic systems: A survey and some new results. Marmara Research Institute Publication, Applied Mathematics Division, no. 37, Gebze Kocaeli, Turkey, 1977c. , Two general properties of the saddle-point solutions of dynamic games, *IEEE Transactions on Automatic Control*, AC-22 (1977d), pp. 124-126. , Decentralized multicriteria optimization of linear stochastic systems, *IEEE Transactions on Automatic Control*, AC-23 (1978), pp. 233-243. , Hierarchical decision making under uncertainty, in *Dynamic Optimization and Mathematical Economics*, P. T. Liu, ed., Plenum Press, New York, NY, 1979a, pp. 205-221. , Information structures and equilibria in dynamic games, in *New Trends in Dynamic System Theory and Economics*, M. Aoki and A. Marzollo, eds., Academic Press, New York and London, 1979b, pp. 3-55. , Stochastic stagewise Stackelberg strategies for linear-quadratic systems, in *Stochastic Control Theory and Stochastic Differential Systems*, M. Kohlmann and W. Vogel, eds., Lecture Notes in Control and Information Sciences, Springer-Verlag, New York, 1979c, ch. 16, pp. 264-276.

BIBLIOGRAPHY 487 —, Memory strategies and a general theory for Stackelberg games with partial dynamic information, in Proceedings of the 4th International Conference on the Analysis and Optimization of Systems, Versailles, France, December 1980a, Springer-Verlag, New York, pp. 397-415. —, On the existence and uniqueness of closed-loop sampled-data Nash controls in linear-quadratic stochastic differential games, in Optimization Techniques, K. Iracki et al., eds., Lecture Notes in Control and Information Sciences, Springer-Verlag, New York, 1980b, ch. 22, pp. 193-203. —, Equilibrium strategies in dynamic games with multi-levels of hierarchy, *Automatica*, 17 (1981a), pp. 749-754. —, A new method for the Stackelberg solution of differential games with sampled-data state information, in Proceedings of the 8th IFAC World Congress, Kyoto, Japan, Pergamon Press, New York, August 1981b, pp. 139-144. —, On the saddle-point solution of a class of stochastic differential games, *Journal of Optimization Theory and Applications*, 33 (1981c), pp. 539-556. —, A general theory for Stackelberg games with partial state information, *Large Scale Systems*, 3 (1982), pp. 47-56. —, Affine incentive schemes for stochastic systems with dynamic information, *SIAM Journal on Scientific and Statistical Computing*, 22 (1984), pp. 199-210. —, ed., *Dynamic Games and Applications in Economics*, vol. 265 of *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin, 1986a. —, A tutorial on dynamic and differential games, in *Dynamic Games and Applications in Economics*, T. Başar, ed., vol. 265 of *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, New York, 1986b, pp. 1-25. —, Relaxation techniques and asynchronous algorithms for on-line computation of noncooperative equilibria, *Journal of Economic Dynamics and Control*, 71 (1987), pp. 531-549. —, Stochastic incentive problems with partial dynamic information and multiple levels of hierarchy, *European J. Political Economy*, 5 (1989a), pp. 203-217. —, Time consistency and robustness of equilibria in noncooperative dynamic games, in *Dynamic Policy Games in Economics*, F. Van der Ploeg and A. de Zeeuw, eds., North-Holland, Amsterdam, 1989b, pp. 9-54. —, A dynamic games approach to controller design: Disturbance rejection in discrete time, *IEEE Transactions on Automatic Control*, AC-36 (1991a), pp. 936-952.

488 BIBLIOGRAPHY , Generalized Riccati equations in dynamic games, in *The Riccati Equation*, S. Bittanti, A. Laub, and J. C. Willems, eds., Springer-Verlag, Berlin, 1991b, pp.293-333. , On the application of differential game

theory in robust control design for economic systems, in *Dynamic Economic Models and Optimal Control*, G. Feichtinger, ed., North-Holland, Amsterdam, 1992, pp. 171-186. T. BaşAR and P. BernhaRD, eds., *Differential Games and Applications*, Lecture Notes in Control and Information Sciences, vol. 119, Springer-Verlag, Berlin, 1989. , *H<sup>∞</sup>-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, Birkhauser, Boston, MA, 2nd ed., 1995. T. BaşAR and A. HAURIE, *Feedback equilibria in differential games with structural and modal uncertainties*, vol. 1 of *Advances in Large Scale Systems*, J. B. Cruz, Jr., ed. JAI Press Inc., Connecticut, 1984, pp. 163-201. T. BaşAR AND A. HAURIE, eds., *Advances in Dynamic Games and Applications*, Birkhauser, Boston, MA, 1994. T. BaşAR, A. HAURIE, and G. RICCI, *On the dominance of capitalists leadership in a feedback Stackelberg solution of a differential game model of capitalism*, *Journal of Economic Dynamics and Control*, 9 (1985), pp. 101- 125. T. Başar and Y. C. Ho, *Informational properties of the Nash solutions of two stochastic nonzero-sum games*, *Journal of Economic Theory*, 7 (1974), pp. 370-387. T. Başar and S. Li, *Distributed algorithms for the computation of Nash equilibria in linear stochastic differential games*, *SIAM Journal on Control and Optimization*, 27 (1989), pp. 563-578. T. BaşAR AND M. Mintz, *On the existence of linear saddle-point strategies for a two-person zero-sum stochastic game*, in *Proceedings of the IEEE 11th Conference on Decision and Control*, New Orleans, LA, 1972, IEEE Computer Society Press, Los Alamitos, CA, pp. 188-192. , *A multistage pursuit-evasion game that admits a Gaussian random process as a maximum control policy*, *Stochastics*, 1 (1973), pp. 25-69. T. Başar and G. J. OLSDER, *Mixed Stackelberg strategies in continuous- kernel games*, *IEEE Transactions on Automatic Control*, AC-25 (1980a), pp. 307-309. , *Team-optimal closed-loop Stackelberg strategies in hierarchical control problems*, *Automatica*, 16 (1980b), pp. 409-414. , *Dynamic Noncooperative Game Theory*, Academic Press, London/New York, 1982. (Second printing 1989.) T. Başar and H. Selbuz, *Properties of Nash solutions of a 2-stage nonzero- sum game*, *IEEE Transactions on Automatic Control*, AC-21 (1976), pp. 48-54.

BIBLIOGRAPHY 489 , *Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems*, *IEEE Transactions on Automatic Control*, AC-24 (1979a), pp. 166-179. , *A new approach for derivation of closed-loop Stackelberg strategies*, in *Proceedings of the IEEE 17th*

Conference on Decision and Control, San Diego, CA, January 1979b, IEEE Computer Society Press, Los Alamitos, CA, pp. 1113-1118. T. BaşAR, S. J. TURNOVSKY, AND V. D'OREY, Optimal macroeconomic policies in interdependent economies: A strategic approach, in *Dynamic Games and Applications in Economics*, T. Başar, ed., *Lecture Notes in Economics and Mathematical Systems*, volume 265, Springer-Verlag, Berlin, 1986, pp. 134-178. R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1957. , *Introduction to Matrix Analysis*, McGraw-Hill, New York, NY, 2nd ed., 1970. A. BENSOUSSAN, Saddle points of convex concave functionals, in *Differential Games and Related Topics*, H. W. Kuhn and G. P. Szego, eds., North- Holland, Amsterdam, The Netherlands, 1971, pp. 177-200. , Points de Nash dans le cas de fonctionnelles quadratiques et jeux differentiels lineaires a N personnes, *SIAM Journal on Control and Optimization*, 12 (1974), pp. 237-243. , *Perturbation Methods in Optimal Control*, John Wiley, Gauthier-Villars, Chichester, England, 1988. L. D. Berkovitz, A variational approach to differential games, in *Advances in Game Theory*, M. Dresher, L. S. Shapley, and A. W. Tucker, eds., Princeton University Press, Princeton, NJ, 1964, pp. 127-174. , *Optimal Control Theory*, Springer-Verlag, Berlin, 1974. P. Bernhard, Condition de coin pour les jeux differentiels. Seminaire des 21-25 Juin, Les Jeux Differentiels, Centre d'Automatique de l'Ecole Nationale Superieure de Mines de Paris, 1971. , Singular surfaces in differential games: an introduction, in *Differential Games and Applications*, P. Hagedorn, H. W. Knobloch, and G. J. Olsder, eds., Springer-Verlag, Berlin, 1977, pp. 1-33. *Lecture Notes in Control and Information Sciences*, vol. 3. , Linear-quadratic two-person zero-sum differential games: Necessary and sufficient conditions, *Journal of Optimization Theory and Applications*, 27 (1979), pp. 51-69. D. P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models*, Prentice-Hall, Englewood Cliffs, NJ, 1987. D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

490 BIBLIOGRAPHY D. P. Bertsekas and J. N. Tsitsiklis, Convergence rate and termination of asynchronous iterative algorithms, in *Proceedings of the 1989 International Conference on Supercomputing*, Irakleion, Greece, 1989, pp. 561-470. , Some aspects of parallel and distributed iterative algorithms - a survey, *Automatica*, 27 (1991), pp. 3-21. K. BINMORE, *Fun and Games*, D. C. Heath and Company, Lexington, MA, 1992. D. BLACKWELL and M. A.

GIRSHICK, Theory of Games and Statistical Decisions, John Wiley and Sons, New York, NY, 1954. A. BlaQUIERE, ed., Topics in Differential Games, North-Holland, Amsterdam, The Netherlands, 1973. , Une generalisation du concept d'optimalite et des certaines notions geometriques qui's rattachent. Institute des Hautes Etudes de Belgique, Cahiers du centre d'etudes de recherche operationelle, vol. 18, no. 1-2, Bruxelles, pp. 49-61, 1976. , Differential games with piece-wise continuous trajectories, in Differential Games and Applications, P. Hagedorn, H. W. Knobloch, and G. J. Olsder, eds., Springer-Verlag, Berlin, 1977, ch. 3, pp. 34-69. A. Blaquiere, F. Gerard, and G. Leitmann, Quantitative and Qualitative Games, Academic Press, New York and London, 1969. V. G. BOLTYANSKI, Optimal Control of Discrete Systems, Hallsted Press, John Wiley, New York, NY, 1978. E. BOREL, The theory of play and integral equations with skew symmetrical kernels: On games that involve chance and skill of the players, On systems of linear forms of skew symmetric determinants and the general theory of play, trans, by L. J. Savage, Econometrica, 21 (1953), pp. 97-117. D. BRAESS, Uber eine Paradoxen aus der Verkehrsplanung, Unternehmensforschung, 12 (1968), pp. 258-268. J. V. Breakwell, Examples elementaires. Seminaire des 21-25 Juin, Les Jeux Differentiels, Centre d'Automatique de l'Ecole Nationale Superieure des Mines de Paris, 1971. , Some differential games with interesting discontinuities. Internal Report, Stanford University, Stanford, CA, 1973. , Zero-sum differential games with terminal payoff, in Differential Games and Applications, P. Hagedorn, H. W. Knobloch, and G. J. Olsder, eds., Springer-Verlag, Berlin, 1977, ch. 3, pp. 70-95. T. F. Bresnahan, Duopoly models with consistent conjectures, American Economic Review, 71 (1981), pp. 934-945. R. W. Brockett, Finite Dimensional Linear Systems, John Wiley and Sons, New York, NY, 1970.

BIBLIOGRAPHY 491 L. E. J. BROUWER, Uber abbildung von mannigfaltigkeiten, Math. Annalen., 71 (1910), pp. 97-115. G. W. BROWN, Iterative solutions of games by fictitious play, in Activity Analysis of Production and Allocation, T. C. Koopmans, ed., John Wiley and Sons, New York, NY, 1951, ch. 13, pp. 374-376. Cowles Commission Monograph. G. W. Brown and J. von Neumann, Solutions of games by differential equations, in Contributions to the Theory of Games, H. W. Kuhn and A. W. Tucker, eds., Princeton University Press, Princeton, NJ, 1950, pp. 73-79. Vol. I, no. 24. A. E. BRYSON, Jr. and Y. C. Ho, Applied Optimal Control, Hemisphere,

Washington, DC, 1975. E. BURGER, Einführung in die Theorie der Spiele, (2 Auflage), Walter de Gruyter and Company, Berlin, 1966. M. D. Canon, C. D. Cullum, Jr., and E. Polak, Theory of Optimal Control and Programming, McGraw-Hill, New York, NY, 1970. D. H. CanSEVER and T. BaŞAR, A minimum sensitivity approach to incentive design problems, Large Scale Systems, 5 (1983), pp. 233-244. , On stochastic incentive control problems with partial dynamic information, Systems & Control Letters, 6 (1985a), pp. 69-75. , Optimum/near optimum incentive policies for stochastic decision problems involving parametric uncertainty, Automatica, 21 (1985b), pp. 575-584. J. H. Case, Equilibrium Points of N-person differential games, PhD thesis, University of Michigan, Ann Arbor, MI, 1967. Department of Industrial Engineering, Tech. report no. 1967-1. , Toward a theory of many player differential games, SIAM Journal on Control and Optimization, 7 (1969), pp. 179-197. , Applications of the theory of differential games to economic problems, in Differential Games and Related Topics, H. W. Kuhn and G. P. Szego, eds., North-Holland, Amsterdam, The Netherlands, 1971, pp. 345-371. , Economics and the Competitive Process, New York University Press, New York, NY, 1979. D. A. CaStanon, Equilibria in Stochastic Dynamic Games of Stackelberg Type, PhD thesis, M.I.T. Electronics Systems Laboratory, Cambridge, MA, 1976. D. Chazan and W. MIRANKER, Chaotic relaxation, Linear Algebra and Applications, 2 (1969), pp. 199-222. C. I. Chen and J. B. Cruz Jr., Stackelberg solution for two-person games with biased information patterns, IEEE Transactions on Automatic Control, AC-17 (1972), pp. 791-798.

492 BIBLIOGRAPHY S. A. Chigir, The game problem on the dolichobrachistochrone, PMM, 30 (1976), pp. 1003-1013. M. D. Ciletti, On the contradiction of the bang-bang-bang surfaces in differential games, Journal of Optimization Theory and Applications, 5 (1970), pp. 163-169. S. Clemhout, G. Leitmann, and H. Y. Wan, Jr., A differential game model of oligopoly, Journal of Cybernetics, 3 (1973), pp. 24-39. E. A. CODDINGTON AND N. LEVINSON, Theory of Ordinary Differential Equations, McGraw-Hill, New York, NY, 1955. J. E. Cohen, The counterintuitive in conflict and cooperation, American Scientist, (1988). J. E. Cohen and P. Horowitz, Paradoxal behaviour of mechanical and electrical networks, Nature, 352 (1991), pp. 699-701. A. COURNOT, Recherches sur les principes mathematiques de la theorie des richesses. Hachette, Paris. (English edition published in 1960 by Kelley, New

York, under the title 'Researches into the Mathematical Principles of the Theory of Wealth', trans, by N. T. Bacon, 1838. M. Crandall and P. Lions, Viscosity solutions of Hamilton-Jacobi equations, Transactions of the American Mathematical Society, 277 (1983), pp. 1-42. M. G. Crandall, L. C. Evans, and P. L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations, Transactions of the American Mathematical Society, 282 (1984), pp. 487-502. J. B. CRUZ, Jr., Leader-follower strategies for multilevel systems, IEEE Transactions on Automatic Control, AC-23 (1978), pp. 244-255. G. B. Dantzig, Linear Programming and Extensions, Princeton University Press, Princeton, NJ, 1963. M. Davis, On the existence of optimal policies in stochastic control, SIAM Journal on Control and Optimization, 11 (1973), pp. 587-594. M. Davis and P. P. Varaiya, Dynamic programming conditions for partially observable stochastic systems, SIAM Journal on Control and Optimization, 11 (1973), pp. 226-261. M. A. DAVIS, Linear Estimation and Stochastic Control, Chapman and Hall, London, 1977. E. DOCKNER AND G. Feichtinger, Dynamic advertising and pricing in an oligopoly: A Nash equilibrium approach, in Dynamic Games and Applications in Economics, T. Başar, ed., Lecture Notes in Economics and Mathematical Systems, volume 265, Springer-Verlag, Berlin, 1986, pp. 238-251. J. Doyle, K. Glover, P. Khargonekar, and B. Francis, State-space solutions to standard  $H_2$  and  $jET^{\wedge}$  control problems, IEEE Transactions on Automatic Control, AC-34 (1989), pp. 831-847.

BIBLIOGRAPHY 493 P. Dubey AND M. Shubik, Information conditions, communications and general equilibrium, Mathematics of Operations Research, 6 (1981), pp. 186- 189. F. Y. EDGEWORTH, Paper Relating to Political Economy, vol. 1, Macmillan, London, 1925. A. Ehrenfeucht and J. Mycielski, Positional strategies for mean payoff games, International Journal of Game Theory, 8 (1979), pp. 109-113. T. ELSELE, Nonexistence and nonuniqueness of open-loop equilibria in linear- quadratic differential games, Journal of Optimization Theory and Applications, 37 (1982), pp. 443-468. R. J. Elliott, The existence of value in stochastic differential games, SIAM Journal on Control and Optimization, 14 (1976), pp. 85-94. , Feedback strategies in deterministic differential games, in Differential Games and Applications, P. Hagedorn, H. W. Knobloch, and G. J. Olsder, eds., Springer-Verlag, Berlin, 1977, pp. 136-142. Lecture Notes in Control and Information Sciences, vol. 3. , Viscosity Solutions and Optimal Control, Longman



Scientific and Technical, Wiley, New York, NY, 1987. R. J. Elliott and N. J. Kalton, The existence of value in differential games of pursuit and evasion, *Journal of Differential Equations*, 12 (1972), pp. 504-523. R. J. Elliott, N. J. Kalton, and L. Markus, Saddle-points for linear differential games, *SIAM Journal on Control and Optimization*, 11 (1973), pp. 100-112. J. C. ENGWERDA, On the open-loop Nash equilibrium in LQ-games, *Journal of Economic Dynamics and Control*, 22 (1998), pp. 729-762. R. A. EPSTEIN, *The Theory of Gambling and Statistical Logic*, Academic Press, New York and London, 1967. L. EVANS AND P. SOUGANIDIS, Differential games and representation formulas for solutions of Hamilton- Jacobi-Isaacs equations, *Indiana Univ. Math. J.*, 33 (1984), pp. 773-797. K. Fan, Minimax theorems, *Proceedings of the Academy of Sciences*, 39 (1953), pp. 42-47. W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, John Wiley and Sons, New York, NY, 1971. W. H. Fleming, The convergence problem for differential games, *Journal of Mathematical Analysis and Applications*, 3 (1961), pp. 102-116. , The convergence problem for differential games, in *Advances in Game Theory*, M. Dresher, L. S. Shapley, and A. W. Tucker, eds., Princeton University Press, Princeton, NJ, 1964, pp. 195-210. *Annals of Mathematics Studies*, no. 52.

494 BIBLIOGRAPHY , Optimal continuous-parameter stochastic control, *SIAM Review*, 11 (1969), pp. 470-509. W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, Berlin, 1975. W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, *Applications of Mathematics*, vol. 25, Springer-Verlag, 1993. J. G. Foreman, The princess and the monster on the circle, *SIAM Journal on Control and Optimization*, 15 (1977), pp. 841-856. M. Fox AND G. S. KIMELDORF, Noisy duels, *SIAM Journal of Applied Mathematics*, 17 (1969), pp. 353-361. B. A. FRANCIS, *A Course in Hx Control Theory*, vol. 88 of *Lecture Notes in Control and Information Sciences*, Springer-Verlag, New York, NY, 1987. J. FRANKLIN, *Methods of Mathematical Economics*, Springer-Verlag, Berlin, 1980. A. Friedman, *Differential Games*, John Wiley and Sons, New York, NY, 1971. , Stochastic differential games, *Journal of Differential Equations*, 11 (1972), pp. 79-108. J. W. FRIEDMAN, *Oligopoly and the Theory of Games*, North-Holland, Amsterdam, The Netherlands, 1977. D. FUDENBERG AND J. TIROLE, *Game theory*, MIT Press, Cambridge, MA, 1991. Z. GAJIC, D. PETKOVSKI, and X. Shen, *Singularly Perturbed and Weakly Coupled Linear*

Control Systems: A Recursive Approach, Springer-Verlag, New York, NY, 1990. S. Gal, Search Games, Academic Press, New York, 1980. W. M. Getz and M. Pachter, Two-target pursuit-evasion differential games in the plane, Journal of Optimization Theory and Applications, 34 (1981), pp. 383-403. I. I. GIKHMAN AND A. V. SKOROHOD, Stochastic Differential Equations, Springer-Verlag, Berlin, 1972. I. L. GLICKSBERG, Minimax theorem for upper and lower semicontinuous payoffs. Rand Corporation Research Memorandum RM-478, Santa Monica, CA, 1950. C. Gonzaga, Path-following methods for linear programming, SIAM Review, 34 (1992), pp. 167-224. M. Green and D. Limebeer, Linear Robust Control, Prentice-Hall, Englewood Cliffs, NJ, 1995. P. Hagedorn, H. W. Knobloch, and G. J. Olsder, eds., Differential Games and Applications, vol. 3 of Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 1977.

BIBLIOGRAPHY 495 O. Hajek, Pursuit Games, Academic Press, New York and London, 1975. , Discontinuous differential equations, I and II, Journal of Differential Equations, 32 (1979), pp. 149-185. A. Halanay, Differential games with delay, SIAM Journal on Control and Optimization, 6 (1968), pp. 579-593. J. Hale, Theory of Functional Differential Equations, Springer-Verlag, Berlin, 1977. R. P. Hamalainen, Nash and Stackelberg solutions to general linear-quadratic two-player difference games, Systems Theory Laboratory B-29, Helsinki University of Technology publication, Espoo, Finland, 1976. R. P. Hamalainen and H. K. Ehtamo, eds., Differential Games: Developments in Modelling and Computation, vol. 156, Springer-Verlag, New York, August 1991. H. Hermes and J. LaSalle, Functional analysis and time optimal control, Academic Press, New York, 1969. Y. C. Ho, Differential games, dynamic optimization and generalized control theory, survey paper, Journal of Optimization Theory and Applications, 6 (1970), pp. 179-209. Y. C. Ho, A. E. Bryson, Jr., and S. Baron, Differential games and optimal pursuit-evasion strategies, IEEE Transactions on Automatic Control, AC-10 (1965), pp. 385-389. Y. C. Ho, P. B. LUH, AND G. J. OLSDER, A control-theoretic view on incentives, in Proceedings of the Fourth International Conference on Analysis and Optimization of Systems, A. Bensoussan and J. L. Lions, eds., Springer-Verlag, Berlin, 1980, pp. 359-383. Lecture Notes in Control and Information Sciences, vol. 28. , A control-theoretic view on incentives, Automatica, 18 (1982), pp. 167- 180. Y. C. Ho AND S. K. Mitter, Directions in Large Scale Systems, Plenum Press, New York, NY, 1976. Y.-C. Ho AND

G. OLSDER, Aspects of the Stackelberg game problem,— incentive, bluff and hierarchy, in Proceedings of the 8th IFAC world congress, Kyoto, Japan, 1981, IFAC, Laxenburg, 1982, pp. 1359-1363. Y.-C. Ho AND G. J. OLSDER, Differential games: Concepts and applications, in Mathematics of Conflict, M. Shubik, ed., North-Holland, Amsterdam, The Netherlands, 1983, pp. 127-186. N. HOWARD, Paradoxes of Rationality: Theory of Metagames and Political Behavior, The MIT Press, Cambridge, MA, 1971. V. HutSON AND J. S. Pym, Applications of Functional Analysis and Operator Theory, Operator Theory, Academic Press, London, 1980.

496 BIBLIOGRAPHY M. D. Intriligator, Mathematical Optimization and Economic Theory, Prentice-Hall, Englewood Cliffs, NJ, 1971. R. ISAACS, Differential games I, II, III, IV. Rand Cooperation Research Memorandum RM-1391, 1399, 1411, 1468, Santa Monica, CA, 1954-1956. , Differential games: Their scope, nature and future, Journal of Optimization Theory and Applications, 3 (1969), pp. 283-295. , Differential Games, Kruger Publishing Company, Huntington, NY, 2nd ed., 1975. (First edition: Wiley, NY, 1965.) J. ISBELL, Finitary games, in Annals of Mathematical Studies, vol. 39, Princeton University Press, Princeton, NJ, 1957, pp. 79-95. H. ISHII, Perron's method for Hamilton-Jacobi equations, Duke Math. J, 55 (1987), pp. 369-384. , A boundary value problem of the dirichlet type for Hamilton-Jacobi equations, Ann. Sc. Norm. Sup. Pisa, XVI (1989), pp. 105-135. V. I. ISTRATESCU, Fixed Point Theory, D. Reidel Publishing Company, Dordrecht, Holland, 1981. D. H. JACOBSON, On values and strategies for infinite-time linear quadratic games, IEEE Transactions on Automatic Control, AC-22 (1977), pp. 490- 491. S. Jorgensen, Optimal dynamic pricing in an oligopolistic market: A survey, in Dynamic Games and Applications in Economics, T. Başar, ed., Lecture Notes in Economics and Mathematical Systems, volume 265, Springer- Verlag, Berlin, 1986, pp. 179-237. M. C. JOSHI AND R. K. BOSE, Some Topics in Nonlinear Functional Analysis, John Wiley and Sons, New York, NY, 1985. T. Kailath, Linear Systems, Prentice-Hall, Englewood Cliffs, NJ, 1980. S. Kakutani, Generalization of Brouwer's fixed point theorem, Duke Journal of Mathematics, 8 (1941), pp. 457-459. M. I. Kamien and N. L. Schwartz, Conjectural variations, Canadian J. Economics, 16 (1983), pp. 191-211. S. Karlin, Matrix Games, Programming and Mathematical Economics, vols. I and II, Addison-Wesley, Reading, MA, 1959. N. KARMARKAR, A new polynomial-time algorithm for linear programming, Combinatorica, 4 (1984),

pp. 373-395. P. P. Khargonekar, State-space  $H^\infty$  control theory and the LQG control problem, in *Mathematical System Theory: The Influence of R. E. Kalman*, A. C. Antoulas, ed., Springer-Verlag, Berlin, 1991, pp. 159-176. P. P. Khargonekar, K. N. Nagpal, and K. R. Poolla,  $H^\infty$  control with transients, *SIAM Journal on Control and Optimization*, 29 (1991), pp. 1373-1393.

BIBLIOGRAPHY 497 E. Kohlberg AND J.-F. Mertens, On the strategic stability of equilibria, *Econometrica*, 54 (1986), pp. 1003-1037. N. KRASOVSKII and A. Subbotin, *Game-theoretical Control Problems*, Springer-Verlag, New York, 1988. D. KREPS, *Game Theory and Economic Modelling*, Oxford University Press, Oxford, England, 1990. D. M. KREPS AND R. WILSON, Sequential equilibria, *Econometrica*, 50 (1982), pp. 863-894. K. Kuga, Brouwer's fixed point theorem: An alternative proof, *SIAM Journal of Mathematical Analysis*, 5 (1974), pp. 893-897. H. W. KUHN, Extensive games and the problems of information, in vol. 29 of *Contributions to the Theory of Games*, Princeton University Press, Princeton, NJ, 1953, pp. 193-216. H. W. Kuhn and A. W. Tucker, eds. H. W. KUHN and G. P. Szego, eds., *Differential Games and Related Topics*, North-Holland, Amsterdam, The Netherlands, 1971. H. W. Kuhn and A. W. Tucker, eds., *Contributions to the Theory of Games*, vol. I, no. 24, *Annals of Mathematical Studies*, Princeton University Press, Princeton, NJ, 1950. , eds., *Contributions to the Theory of Games*, vol. I, no. 28, *Annals of Mathematical Studies*, Princeton University Press, Princeton, NJ, 1953. P. R. KUMAR AND T. H. Shiau, Existence of value and randomized strategies in zero-sum discrete-time stochastic dynamic games, *SIAM Journal on Control and Optimization*, 19 (1981), pp. 617-634. F. Kydland, Noncooperative and dominant player solutions in discrete dynamic games, *International Journal of Economic Review*, 16 (1975), pp. 321-335. F. Kydland and E. Prescott, Rules rather than discretion: The inconsistency of optimal plans, *Journal of Political Economy*, 85 (1977), pp. 473- 491. S. LAKSHMIVARAHAN AND K. S. NARENDRA, Learning algorithms for two-person zero-sum stochastic games with incomplete information, *Mathematics of Operations Research*, 6 (1981), pp. 379-286. G. Leitmann, *Cooperative and Noncooperative Many Player Differential Games*, vol. 190 of *CISM Monograph*, Springer-Verlag, Vienna, 1974. , Many player differential games, in *Differential Games and Applications*, P. Hagedorn, H. W. Knobloch, and G. J. Olsder, eds., Springer-Verlag, Berlin, 1977, pp. 153-171. *Lecture Notes in Control and Information Sciences*, vol. 3. , On generalized Stackelberg

strategies, *Journal of Optimization Theory and Applications*, 26 (1978).

498 BIBLIOGRAPHY G. Leitmann and H. Y. Wan, Jr., Macro-economic stabilization policy for an uncertain dynamic economy, in *New Trends in Dynamic System Theory and Economics*, M. Aoki and A. Marzollo, eds., Academic Press, London and New York, 1979, pp. 105-136. C. E. Lemke AND J. F. HOWSON, Equilibrium points of bi-matrix games, *SIAM Journal on Applied Mathematics*, 12 (1964), pp. 413-423. J. Levine, Two-person zero-sum differential games with incomplete information-a Bayesian model, in *Differential Games and Control Theory III*, P. T. Liu and E. O. Roxin, eds., Plenum Press, New York, NY, 1979, pp. 119-151. Proceedings of the 3rd Kingston Conference 1978, Part A. *Lecture Notes in Pure and Applied Mathematics*, vol. 44. R. E. Levitan and M. Shubik, Noncooperative equilibrium and strategy spaces in an oligopolistic market, in *Differential Games and Related Topics*, H. W. Kuhn and G. P. Szego, eds., North-Holland, Amsterdam, The Netherlands, 1971, pp. 429-448. J. Lewin, The bang-bang-bang problem revisited, *Journal of Optimization Theory and Applications*, 18 (1976), pp. 429-432. , *Differential Games: Theory and Methods for Solving Game Problems with Singular Surfaces*. Springer-Verlag, New York, 1994. J. Lewin and G. J. Olsder, Conic surveillance evasion, *Journal of Optimization Theory and Applications*, 19 (1979), pp. 107-125. S. Li AND T. BaşAR, Distributed algorithms for the computation of noncooperative equilibria, *Automatica*, 23 (1987), pp. 523-533. D. Limebeer, B. Anderson, P. Khargonekar, and M. Green, A game theoretic approach to  $H^\infty$  control for time-varying systems, *SIAM J. Control Optim.*, 30 (1992), pp. 262-283. P. LIONS, *Generalized Solutions of Hamilton-Jacobi Equations*, Pitman Advanced Publishing Program, 1982. M. LOEVE, *Probability Theory*, Van Nostrand, Princeton, NJ, 3rd ed., 1963. R. D. LUCE AND H. RAIFFA, *Games and Decisions*, John Wiley and Sons, New York, NY, 1957. D. G. LUENBERGER, *Optimization by Vector Space Methods*, John Wiley and Sons, New York, NY, 1969. , *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA, 1973. D. L. LUKES, Equilibrium feedback control in linear games with quadratic costs, *SIAM Journal on Control and Optimization*, 9 (1971), pp. 234-252. D. L. Lukes and D. L. Russell, A global theory for linear quadratic differential games, *Journal of Mathematical Analysis and Applications*, 33 (1971), pp. 96-123.

BIBLIOGRAPHY 499 E. F. Mageirou, Values and strategies for infinite duration linear quadratic games, *IEEE Transactions on Automatic Control*, AC-21 (1976), pp. 547- 550. A. Maitra AND T. Parthasarathy, On stochastic games, *Journal of Optimization Theory and Applications*, 5 (1970), pp. 289-300. M. MARCUS AND H. MINC, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Inc., Boston, MA, 1964. J. McKinsey, *Introduction to the Theory of Games*, McGraw-Hill, New York, NY, 1952. J. MEDANIC, Closed-loop Stackelberg strategies in linear-quadratic problems, in *Proceedings of the 1977 JACC*, San Francisco, CA, 1977, pp. 1324-1329. J. F. Mertens, Formulation of Bayesian analysis for games with incomplete information, *International Journal of Game Theory*, 14 (1985), pp. 1-29. , Stabk equilibria - a reformulation. Part II Discussion of the definition and further results, *Mathematics of Operations Research*, 16 (1991), pp. 694-753. A. W. Merz, *The homicidal chauffeur-a differential game*, Tech report, Guidance and Control Laboratory 418, Stanford University, Stanford, CA, 1971. , Optimal evasive maneuvers in maritime collision avoidance, *Navigation*, 20 (1973), pp. 144-152. , A differential game solution to the coplanar tail-chase aerial combat problem. Tech report NASA-Cr-137809, NASA Langley Research Center, Hampton, VA, 1976. T. Miloh and S. D. Sharma, Maritime collision avoidance as a differential game, Berich 329, Institut für Schiffbau der Universität Hamburg, 1976. R. B. Myerson, Refinement of the Nash equilibrium concept, *International Journal of Game Theory*, 7 (1978), pp. 73-80. , *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, MA, 1991. J. NASH, Noncooperative games, *Annals of Mathematics*, 54 (1951), pp. 286- 295. H. NIKAIDO AND K. ISODA, Note on non-cooperative convex games, *Pacific Journal of Mathematics*, 5 (1955), pp. 807-815. K. OKUGUCHI, *Expectations and Stability in Oligopoly Models*, *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin, 1976, ch. 138. G. Olsder, A critical analysis of a new equilibrium concept, in *Axiomatics and Pragmatics of Conflict Analysis*, J. Paeling and P. H. Vossen, eds., Gower Pub. Co., Aldershot, England, 1987, pp. 80-100.

500 BIBLIOGRAPHY G. J. OLSDER, Some thoughts about simple advertising models as differential games and the structure of coalitions, in *Directions in Large Scale Systems*, Y. C. Ho and S. K. Mitter, eds., Plenum Press, New York, 1976, pp. 187- 206. , Information structures in differential games, in *Differential Games and Control Theory II*, E. O. Roxin, P. T. Liu, and R. L.

Sternberg, eds., Marcel Dekker, New York, NY, 1977a, pp. 99-136. , On observation costs and information structures in stochastic differential games, in *Differential Games and Applications*, P. Hagedorn, H. W. Knobloch, and G. J. Olsder, eds., Springer-Verlag, Berlin, 1977b, ch. 3, pp. 172-185. G. J. Olsder and J. V. Breakwell, Role determination in an aerial dogfight, *International Journal of Game Theory*, 3 (1974), pp. 47-66. G. J. Olsder and G. P. Papavassilopoulos, About when to use the searchlight, *Journal of Mathematical Analysis and Applications*, 136 (1988a), pp. 466-478. , A Markov chain with dynamic information, *Journal of Optimization Theory and Applications*, 59 (1988b), pp. 467-486. G. J. Olsder and J. Walter, Collision avoidance of ships, in *Proceedings 8th IFIP Conference on Optimization Techniques*, J. Stoer, ed., Springer-Verlag, Berlin, 1978, ch. 6, pp. 264-271. G. OWEN, *Game Theory*, Saunders, Philadelphia, PA, 1968. , Existence of equilibrium pairs in continuous games, *International Journal of Game Theory*, 5 (1974), pp. 97-105. , *Game Theory*, Academic Press, New York, NY, 1982. M. Pachter and Y. Yavin, A stochastic homicidal chauffeur pursuit-evasion differential game, Technical Report TWISK 95, WNNR, CSIR, Pretoria, South Africa, 1979. Z. PAN AND T. BaŞAR,  $H^\infty$ -optimal control for singularly perturbed systems. Part I: Perfect state measurements, *Automatica*, 29 (1993), pp. 401-423. ,  $H^\infty$ -optimal control for singularly perturbed systems. Part II: Imperfect state measurements, *IEEE Transactions on Automatic Control*, AC-39 (1994a), pp. 280-299. ,  $H^\infty$ -optimal control of singularly perturbed systems with sampled-state measurements, in *Advances in Dynamic Games and Applications*, T. Bas,ar and A. Haurie, eds., Birkhauser, Boston, MA, 1994b, pp. 23-55. G. P. Papavassilopoulos and J. B. Cruz, Jr., Nonclassical control problems and Stackelberg games, *IEEE Transactions on Automatic Control*, AC-24 (1979a), pp. 155-166.

BIBLIOGRAPHY 501 , On the existence of solutions to coupled matrix Riccati differential equations in linear quadratic Nash games, *IEEE Transactions on Automatic Control*, AC-24 (1979b), pp. 127-129. , Sufficient conditions for Stackelberg and Nash strategies with memory, *Journal of Optimization Theory and Applications*, 31 (1980), pp. 233-260. G. P. Papavassilopoulos, J. V. Medanic, and J. B. Cruz, Jr., On the existence of Nash strategies and solutions to coupled Riccati equations in linear-quadratic games, *Journal of Optimization Theory and Applications*, 28 (1979), pp. 49-76. G. P. Papavassilopoulos and G. J. Olsder, On the linear-quadratic, closed-loop, no-

memory Nash game, *Journal of Optimization Theory and Applications*, 42 (1984), pp. 551-560. A. PAPOULIS, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, NY, 1965. T. Parthasarathy AND T. Raghavan, *Some Topics in Two-person Games*, Elsevier, New York, NY, 1971. T. Parthasarathy and M. Stern, Markov games - a survey, in *Differential Games and Control Theory II*, E. O. Roxin, P. T. Liu, and R. L. Sternberg, eds., Marcel Dekker, New York, NY, 1977, pp. 1-46. W. Y. Peng and T. L. Vincent, Some aspects of aerial combat, *AIAA Journal*, 13 (1975), pp. 7-11. R. S. PINDYCK, Optimal economic stabilization policies under decentralized control and conflicting objectives, *IEEE Transactions on Automatic Control*, AC-22 (1977), pp. 517-530. M. Pohjola, Applications of dynamic game theory to macroeconomics, in *Dynamic Games and Applications in Economics*, T. Başar, ed., *Lecture Notes in Economics and Mathematical Systems*, volume 265, Springer-Verlag, Berlin, 1986, pp. 103-133. L. S. PONTRYAGIN, Linear differential games, I, *Soviet Math. Doklady*, 8 (1967), pp. 769-771. L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE, AND E. F. MISHCHENKO, *The Mathematical Theory of Optimal Processes*, Interscience Publishers, New York, NY, 1962. W. Poundstone, *Prisoner's Dilemma*, Doubleday, New York, NY, 1992. O. Pourtallier AND B. TOLWINSKI, Discretization of Isaacs' equation: A convergence result, Tech. report, INRIA, Sophia-Antipolis, France, 1992. R. Radner, Monitoring cooperative agreements in a repeated principal-agent relationship, *Econometrica*, 49 (1981), pp. 1127-1148. , Repeated principal-agent games with discounting, *Econometrica*, 53 (1985), pp. 1173-1198.

502 BIBLIOGRAPHY T. Raghavan and J. Filar, Algorithms for Stochastic Games - A Survey, *ZOR - Methods and Models of Operations Research*, 35 (1991), pp. 437- 472. W. T. Reid, *Riccati Differential Equations*, Academic Press, New York and London, 1972. R. ReStrepo, Tactical problems involving several, actions, in *Contributions to the Theory of Games*, M. Dresher, A. W. Tucker, and P. Wolfe, eds., Princeton University Press, Princeton, NJ, 1957, ch. III, pp. 313-335. J. ROBINSON, An iterative method of solving a game, *Annals of Mathematics*, 54 (1951), pp. 296-301. J. B. ROSEN, Existence and uniqueness of equilibrium points for concave n- person games, *Econometrica*, 33 (1965), pp. 520-534. E. O. ROXIN, The axiomatic approach in differential games, *Journal of Optimization Theory and Applications*, 3 (1969), pp. 153-163. , Differential games with partial differential equations, in *Differential*



Games and Applications, P. Hagedorn, K. W. Knobloch, and G. J. Ols- der, eds., Springer-Verlag, Berlin, 1977, vol. 3, pp. 186-204. Lecture Notes in Control and Information Sciences. W. RUPP, e-Gleichgewichtspunkte in n-Personenspiekn, in Mathematical Economics and Game Theory; Essays in Honor of Oskar Morgenstern, R. Henn and O. Moeschlin, eds., Springer-Verlag, Berlin, 1977, pp. 128-138. H. Sagan, Introduction to the Calculus of Variations, McGraw-Hill, New York, NY, 1969. R. C. SCALZO, n-person linear quadratic differential, games with constraints, SIAM Journal on Control, 12 (1974), pp. 419-425. H. SCARF, The approximation of fixed points of a continuous mapping, SIAM Journal on Applied Mathematics, 15 (1967), pp. 1328-1343. J. SCHAUDER, Der Fixpunktsatz in Funktionalraumen, Studia Mathematica, 2 (1930), pp. 171-180. W. E. SCHMITENDORF, Existence of optimal open-loop strategies for a class of differential games, Journal of Optimization Theory and Applications, 5 (1970), pp. 363-375. R. Selten, Reexamination of the perfectness concept for equilibrium points in extensive games, International Journal of Game Theory, 4 (1975), pp. 25- 55. L. S. Shapley, Stochastic games, in Proceedings of the National Academy of Sciences, vol. 39, 1953, pp. 1095-1100. , A note on the Lemke-Howson algorithm. Pivoting and extensions, Mathematical Programming Study, 1 (1974), pp. 175-189. L. D. Sharma, On ship maneuverability and collision avoidance. International report, Institut fur Schiffabu der Universitat Hamburg, 1976.

BIBLIOGRAPHY 503 M. Shubik, Game Theory in the Social Sciences, MIT Press, Cambridge, MA, 1983. M. Simaan AND J. B. CRUZ, Jr., Additional aspects of the Stackelberg strategy in nonzero sum games, Journal of Optimization Theory and Applications, 11 (1973a), pp. 613-626. , On the Stackelberg strategy in nonzero sum games, Journal of Optimization Theory and Applications, 11 (1973b), pp. 533-555. R. R. Singleton and W. F. Tyndal, Games and Programs, Freeman, San Francisco, CA, 1974. M. SION, On general minimax theorems, Pacific Journal of Mathematics, 8 (1958), pp. 171-176. M. SION AND P. Wolfe, On a game without a value, in Contributions to the Theory of Games, M. Dresher, A. W. Tucker, and P. Wolfe, eds., Princeton University Press, Princeton, NJ, 1957, ch. III, pp. 209-306. D. R. Smart, Fixed Point Theorems, Cambridge University Press, London, 1974. M. J. SOBEL, Noncooperative stochastic games, Ann. of Math. Statist., 42 (1971), pp. 1930-1935. R. Srikant and T. Başar, Iterative computation of noncooperative

equilibria in nonzero-sum differential games with weakly coupled players, *Journal of Optimization Theory and Applications*, 71 (1991), pp. 137-168. , Sequential decomposition and policy iteration schemes for M-player games with partial weak coupling, *Automatica*, 28 (1992), pp. 95-106. V. Staicu, Minimal time function and viscosity solutions, *Journal of Optimization Theory and Applications*, 60 (1989), pp. 81-91. A. W. Starr and Y. C. Ho, Further properties of nonzero-sum differential games, *Journal of Optimization Theory and Applications*, 3 (1969a), pp. 207-219. ———, Nonzero-sum differential games, *Journal of Optimization Theory and Applications*, 3 (1969b), pp. 184-206. A. A. STOORVOGEL, *The Hao Control Problem: A State Space Approach*, Prentice-Hall, New York, 1992. J. SzEP AND F. FORGO, *Introduction to the Theory of Games*, D. Reidel Publishing Co, Boston, MA, 1985. S. H. Tus, *Semi-infinite and infinite matrix games and bimatrix games*, PhD thesis, University of Nijmegen, The Netherlands, 1975. B. TOLWINSKI, Numerical-solution of N-person non-zero-sum differential games, *Control and Cybernetics*, 7 (1978a), pp. 37-50. , On the existence of Nash equilibrium points for differential games with linear and nonlinear dynamics, *Control and Cybernetics*, 7 (1978b), pp. 57- 69.

504 BIBLIOGRAPHY , Stackelberg solution of dynamic game with constraints, in *Proceedings of the IEEE 19th Conference on Decision and Control*, Albuquerque, NM, 1980, IEEE Computer Society Press, Los Alamitos, CA. , Closed-loop Stackelberg solution to multi-stage linear-quadratic games, *Journal of Optimization Theory and Applications*, 35 (1981a), pp. 485-502. , Equilibrium solutions for a class of hierarchical games, in *Applications of Systems Theory to Economics, Management and Technology*, J. Outen- baum and M. Niergodka, eds., PWN, Warsaw, Poland, 1981b. J. N. TSITSIKLIS, On the stability of asynchronous iterative processes, *Math. Systems Theory*, 20 (1987), pp. 137-153. , A comparison of Jacobi and Gauss-Seidel parallel iterations, *Applied Mathematical Letters*, 2 (1989), pp. 167-170. S. J. TURNOVSKY, T. BaşAR, AND V. D'OREY, Dynamic strategic monetary policies and coordination in interdependent economics, *The American Economic Review*, 78 (1988), pp. 341-361. K. UCHIDA, On the existence of Nash equilibrium point in n-person nonzero- sum stochastic differential games, *SIAM Journal on Control and Optimization*, 16 (1978), pp. 142-149. , A note on the existence of a Nash equilibrium point in stochastic differential games, *SIAM Journal on Control and Optimization*, 17 (1979), pp.

1-4. K. UCHIDA AND M. Fujita, On the central controller: Characterizations via differential games and LEQG control problems, *Systems & Control Letters*, 13 (1989), pp. 9-13. E. VAN Damme, A relation between perfect equilibria in extensive games and proper equilibria in normal form, *International Journal of Game Theory*, 13 (1984), pp. 1-13. , *Stability and Perfection of Equilibria*, Springer-Verlag, New York, 1987. , *Stable equilibria and forward induction*, *Journal of Economic Theory*, 48 (1989), pp. 476-496. P. P. Varaiya, The existence of solutions to a differential game, *SIAM Journal on Control*, 5 (1967), pp. 153-162. , N-person nonzero-sum differential games with linear dynamics, *SIAM Journal on Control*, 8 (1970), pp. 441-449. , N-player stochastic differential games, *SIAM Journal on Control and Optimization*, 14 (1976), pp. 538-545. P. P. Varaiya and J. G. Lin, Existence of saddle-point in differential games, *SIAM Journal on Control*, 7 (1969), pp. 141-157. M. VIDYASAGAR, A new approach to N-person, nonzero-sum, linear differential games, *Journal of Optimization Theory and Applications*, 18 (1976), pp. 171-175.

BIBLIOGRAPHY 505 J. A. Ville, Sur la theorie generale desjeux ou intervient l'habilitedes joueurs. In E. Borel, *Traite du calcul des probabilites et de ses applications IV*, 1938, pp. 105-113. T. L. Vincent and W. Y. Peng, Ship collision avoidance. Workshop on differential games, Naval Academy, Annapolis, MO, 1973. T. L. Vincent, D. L. Stricht, and W. Y. Peng, Aircraft missile avoidance, *Journal of Operations Research*, 24 (1976). J. VON Neumann, Zur theorie der geselkchaftspiele, *Mathematische Annalen*, 100 (1928), pp. 295-320. , *Uber ein okonomisches gleichungssystem und eine verallgemeinerung des Brouwerschen fixpunktsatzes*, *Ergebnisse eines Mathematik Kolloquiums*, 8 (1937), pp. 73-83. , A numerical method to determine optimum strategy, *Naval Research Logistics Quarterly*, 1 (1954), pp. 109-115. J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 2nd ed., 1947. H. VON Stackelberg, *Marktform und Gleichgewicht*. Springer-Verlag, Vienna. (An English translation appeared in 1952 entitled *The Theory of the Market Economy*, published by Oxford University Press, Oxford, England, 1934. N. H. VOROB'EV, *Game Theory*, Springer-Verlag, Berlin, 1977. A. Wald, Generalization of a theorem by von Neumann concerning zero-sum two-person games, *Annals of Mathematics*, 46 (1945), pp. 281-286. J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York

and London, 1972. A. WaSHBURN, Deterministic graphical games, *Journal of Mathematical Analysis and Applications*, 153 (1990), pp. 84-96. R. L. Weil, Game theory and eigensystems, *SIAM Review*, 10 (1968), pp. 360- 367. J. C. WILLEMS, Least squares stationary optimal control and the algebraic Riccati equation, *IEEE Transactions on Automatic Control*, AC-16 (1971), pp. 621-634. J. D. Williams, *The Compleat Strategyst*, McGraw-Hill, New York, NY, 1954. J. L. Willman, Formal solutions for a class of stochastic pursuit-evasion games, *IEEE Transactions on Automatic Control*, AC-14 (1969), pp. 504- 509. D. J. WILSON, Differential games with no information, *SIAM Journal on Control and Optimization*, 15 (1977), pp. 233-246. D. WISHART AND G. J. Olsder, Discontinuous Stackelberg solutions, *International Journal of Systems Science*, 10 (1979), pp. 1359-1368.

506 BIBLIOGRAPHY H.J. WITSENHAUSEN, On the relations between the values of a games and its information structure, *Information and Control*, 19 (1971). H. S. WITSENHAUSEN, On information structures, feedback and causality, *SIAM Journal on Control*, 9 (1971), pp. 149-160. , Alternatives to the tree model, for extensive games, in *The Theory and Applications of Differential Games*, J. D. Grote, ed., Reidel Publishing Company, The Netherlands, 1975, pp. 77-84. E. Wong AND B. E. Hajek, *Stochastic Processes in Engineering Systems*, Springer-Verlag, New York, NY, 1985. L. A. Zadeh, The concepts of system, aggregate and state in system theory, in *System Theory*, L. A. Zadeh and E. Polak, eds., McGraw-Hill, New York, NY, 1969, pp. 3-42. G. Zames, Feedback and optimal sensitivity: Model reference transformation, multiplicative seminorms and approximate inverses, *IEEE Transactions on Automatic Control*, AC-26 (1981), pp. 301-320. Y. P. Zheng and T. Başar, Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach, *International Journal of Control*, 35 (1982), pp. 997-1012. Y. P. Zheng, T. Başar, and J. B. Cruz, Jr., Incentive Stackelberg strategies for deterministic multi-stage decision processes, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-14 (1984), pp. 10-24.

Corollaries, Definitions, Examples, Lemmas, Propositions, Remarks and Theorems Chapter 1 Examples No. 1.1 1.2 1.3 1.4 p. 4 6 10 10 Chapter 2 Corollaries No. 2.1 2.2 2.3 2.4 2.5 p. 22 26 28 45 53 Definitions No. 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 2.10 2.11 2.12 2.13 2.14 2.15 Examples No. 2.1 2.2





equilibrium solution, 79 Advertising game, 6 Aeronautics application, 456 AfEne-quadratic continuous-time optimal control, 238 AfEne-quadratic differential game, 312 AfEne-quadratic discrete-time optimal control, 234 Average Stackelberg cost, 182 Average value, 480 Banach space, 472 Bang-bang-bang surface, 468 Barrier, 245, 435, 447 Battle of the sexes, 80, 87 Behavioral Stackelberg equilibrium, 143 Behavioral strategies, 44, 143 Behavioral strategies in infinite dynamic game, 216 Behavioral strategies in multi-act games, 50 Better pair of strategies, 79 Bimatrix games, 78 Borel set, 478 Braess paradox, 203 Brouwer's fixed point theorem, 92 Canonical equations, 243 Capturability, 434 Cauchy problem, 255 Cauchy sequence, 472 Cauchy-Schwarz inequality, 481 Causality, 217 Closed set, 472 Closed-loop, 247 Closed-loop imperfect state (CLIS) information, 221 Closed-loop no-memory (CLNM) Nash equilibrium solution, 267 Closed-loop perfect state (CLPS) information, 221, 225 Collision avoidance, 451 Compact set, 472 Complete vector space, 472 Completely mixed Nash equilibrium, 86 Concavity, 473 Concurrent Stackelberg solution, 186 Conflict, 1 Conic surveillance evasion, 468 Constant strategy, 7, 41 Constant-sum game, 3 Continuity, 473 Control function, 224 Cooperative game, 5 Cost function(al), 98, 221, 223 Costate vector (variable), 242 Cournot duopoly situation, 199 Cumulative (joint) probability, 479 Cusp, 441 Decision maker, 1 Decision point, 445, 451 Decision rule, 5 Delayed commitment, 108, 112, 121 Differential game, 2, 224 Dimension of a set, 472 Dimensional random vector, 478 Dirichlet problem, 255 515

516 INDEX Discount factor, 259 Discrete-time infinite game, 223 Dispersal line, 443 Dispersal surface (line), 442 Divide and rule, 396 Duel, 66 Duopoly, 199 Dynamic game, 2 Dynamic programming, 3 Dynamic single-act game, 112 Eigenvalue, 473 Eigenvector, 473 Epsilon-delayed closed-loop perfect state (eDCLPS), 226 Epsilon-neighborhood, 472 Epsilon-perfect equilibrium, 129 Equicontinuous, 484 Equivocal line, 447 Evader, 424 Event, 477 Expected (average) cost functional, 223 Extensive form, 36, 97, 119, 149, 220 Feedback, 10 Feedback imperfect state information, 221 Feedback Nash equilibrium, 277 Feedback Nash equilibrium solution (FNES), 120, 267, 278, 321 Feedback perfect state information (FB), 221, 226 Feedback saddle-point strategy, 47 Feedback Stackelberg cost, 137 Feedback Stackelberg solution, 136 Feedback strategies, 247 Finite game, 3 Finite set, 471 Finite-dimensional set, 472 Focal line, 447 Follower, 11, 131, 146 Followers' group, 146

Function, 472 Functional, 472 Game of degree, 220, 435 Game of kind, 220, 434 Global Stackelberg, 136 Globally minimizing solution, 474 Gradient, 474 Graphical solution, 30 Hamilton-Jacobi-Bellman (HJB) equation, 236 Hamilton-Jacobi-Isaacs equation, 326 Hessian matrix, 474 Hierarchical equilibrium, 146 Hierarchical equilibrium strategy, 146 Homicidal chauffeur game, 436, 437 Identity matrix, 473 Incentives, 392 Independent, 479 Infinite dimensional set, 472 Infinite game, 10 Infinite set, 471 Infinite-horizon, 262 Information, 219 Information field, 225 Information set, 37, 98, 149 Information space, 220 Information structure, 225 Informational inferiority, 103, 122 Informational nonunique Nash equilibria, 265 Informational nonuniqueness, 105 Informationally inferiority, 296 Initial state, 219 Interior, 472 Isaacs condition, 326, 353, 428 Isaacs equation, 326, 423 Iso-cost curve, 169 Isochrone, 245 Kruzkov-transform, 259 Ladder-nested form, 105 Lady in the lake, 10, 429 Leader, 11, 131 Leader's Stackelberg cost, 180 Leaking corner, 440

INDEX 517 Lebesgue measure, 479 Lemke-Howson algorithm, 96 Limit point, 472 Linear independence, 471 Linear programming, 31 Linear-quadratic continuous-time optimal control, 238 Linear-quadratic discrete-time optimal control, 234 Linearly independent vectors, 471 Lion and man, 468 Lipschitz-continuity, 227 Locally stable Nash equilibrium, 171 Loop model, 216 Loss-ceiling, 18 Lower value, 19, 164, 432 Marginal distribution function, 479 Markov games, 262 Matrix elements, 473 Matrix entries, 473 Mean (expected) value, 480 Measurable function, 478 Measurable space, 478 Measurable transformation, 478 Memoryless perfect state information (MPS), 221, 226 Mental poker, 154 Minimal cost, 233 Minimax solution, 11 Minimax theorem, 26, 211 Minimax values of a bimatrix game, 84 Minimizing element, 474 Minimum cost, 233 Minimum principle, 246 Miss distance, 452 Mixed epsilon Nash equilibrium solution, 166 Mixed saddle-point strategy, 163 Mixed security strategy, 25, 163 Mixed Stackelberg equilibrium strategy, 141, 182 Mixed strategies, 11, 23, 91, 182, 232 Mixed-strategy noncooperative (Nash) equilibrium solution, 91 Multi-act game, 39 Multi-act zero-sum games, 45  $n$ -dimensional Euclidean space, 471  $N$ -person deterministic multi-stage game, 219  $N$ -person nonzero-sum feedback game in extensive form, 119  $N$ -person nonzero-sum feedback game in ladder-nested form, 119 Nash condition, 352 Nash equilibrium, 4 Nash equilibrium solution, 163 Nature, 2, 223 Near miss, 458 Near miss maneuver, 462 Negative definite, 473 Nested, 105 Nested form,



119 Nested information, 405 Noisy observation, 362 Nonconcurrent Stackelberg solution, 186 Noncooperative (Nash) equilibrium outcome, 79, 85, 89, 91 Noncooperative (Nash) equilibrium solution, 78, 85, 88, 99 Noncooperative game, 1 Nondenumerable set, 471 Nonnegative definite, 473 Nonnested information, 405 Nonpositive definite, 473 Norm, 472 Normal form, 36, 221 Normed linear vector space, 472 Observation, 219 Obstacle tag, 468 Oligopoly, 198 One-step delayed observation sharing pattern (IDOS), 222 One-step memory, 248 Open set, 472 Open-loop (OL), 10, 221, 225

518 INDEX Open-loop games, 53 Open-loop Nash equilibrium solution, 267 Open-loop realization, 247 Open-loop strategies, 247 Open-loop value, 247 Optimal control theory, 2 Optimal response, 169 Optimal response set, 133, 143, 146 Optimal solution, 474 Optimal strategy, 134, 146, 180 Optimal value, 475 Optimality, 11 Ordered interchangeability of saddle points, 177 Pareto-optimal solution, 5 Partial differential equation (PDE), 236 Pay-off function, 60 Penetration, 436 Perfect equilibrium, 129 Person-by-person (pbp) optimality, 196 Payability, 229 Playable strategies, 229 Player, 1 Player sets, 98 Players' set, 219, 224 Positive definite, 473 Positive semidefinite, 473 Precedence of players, 105 Price game, 201 Princess and the monster, 10 Prior commitment, 54, 108 Prisoners' dilemma, 83 Probability density function, 479 Probability distribution, 149, 478, 479 Probability measure, 478 Probability space, 478 Proper equilibrium, 130 Pure strategy, 21, 38 Pursuer, 424 Quadratic games, 190 Qualitative game, 220 Quantitative game, 220 Quantity game, 201 Quasi-perfect equilibrium, 131 Random variable, 478 Randomized strategies, 56 Rational reaction set, 169 Reaction curve, 169 Reaction function, 169 Relative coordinate system, 448 Repeated games, 159 Representation of a strategy, 248, 297 Robust Nash solution, 171 Rope-pulling game, 4 Saddle point in behavioral strategies, 44 Saddle point in mixed strategies, 25 Saddle-point equilibrium, 21 Searchlight game, 66 Security level, 18 Security strategy, 18 Semi-infinite bimatrix game, 166 Semipermeable, 427 Semipermeable surface, 435, 460 Separating hyperplane theorem, 474 Sequential equilibrium, 131 Ship collision avoidance, 458 Sigma algebra, 478 Simplex, 23 Simultaneous confrontation, 458 Single-act game, nonzero-sum, 40 Single-act game, zero-sum, 41 Singleton set, 374 Singular lines, 442 Singular surfaces (line), 442 Stable Nash equilibrium, 171 Stackelberg cost, 141 Stackelberg equilibrium, 133, 134, 179 Stackelberg game, 11 Stage of the game, 219 Stage-additive cost function, 221

INDEX 519 Stagewise (Nash) equilibrium solution, 298 State equation, 219 State space, 219 State trajectory, 224 Static (open-loop) version, 122 Static normal form, 103, 122 Static sub-extensive form, 106 Static version of a dynamic game, 103 Stationary strategies, 262 Stochastic differential equation, 230 Stochastic differential game, 230, 263 Stochastic games, 222 Stochastic incentive problems, 405 Stochastic systems, 252 Strategic equilibrium, 131 Strategic equivalence, 34, 81, 90 Strategic form, 3 Strategy, 5, 18, 220 Strategy set, 39, 99 Strategy space, 39, 91, 99 Strict dominance, 35 Strong positive definiteness, 190 Strong time consistency, 250 Switching envelope, 447 Switching function, 244 Switching surface, 442 Symmetric game, 71 Target set, 228, 424 Team problem, 81, 196, 299, 378 Team-optimal solution, 196 Terminal cost functional, 222 Terminal time, 228 Termination, 220, 229, 431 Termination dilemma, 432 Threat, 392 Threat strategy, 380 Time consistency, 249 Time inconsistency, 250 Trajectory space, 224 Transition surface (line), 442 Tree structure, 36 Trembling hand equilibrium, 129 Two-cars model, 437 Two-target games, 456 Undecomposable extensive form, 112 Universal surface (line), 444 Unstable Nash equilibrium, 171 Upper value, 19, 164, 432 Usable part (UP), 245, 435 Value, 21, 164 Value function, 234, 236, 424 Value in mixed strategies, 26 Vector space, 471 Viscosity solution, 257, 432 Viscosity subsolution, 257, 260 Viscosity supersolution, 257, 260 Weak time consistency, 250 Weierstrass theorem, 475 Wiener process, 230 Zero-sum game, 3 Zero-sum games with chance moves, 57