BIRKHÄUSER

Annals of the International Society of Dynamic Games

Advances in Dynamic Game Theory

Numerical Methods, Algorithms, and Applications to Ecology and Economics

Steffen Jørgensen
Marc Quincampoix
Thomas L. Vincent
Editors

BIRKHÄUSER

Annals of the International Society of Dynamic Games
Volume 9

Series Editor
Tamer Başar

Annals of the International Society of Dynamic Games

Advances in Dynamic Game Theory

Numerical Methods, Algorithms, and Applications to Ecology and Economics

Steffen Jørgensen
Marc Quincampoix
Thomas L. Vincent
Editors

Birkhäuser
Boston • Basel • Berlin

Steffen Jørgensen
University of Southern Denmark
Department of Economics and Business
Campusvej 55
DK-5230 Odense
Denmark

Marc Quincampoix
Université de Bretagne Occidentale
Laboratoire de Mathématiques
6 Avenue Le Georgeu
29200 Brest
France

Thomas L. Vincent
University of Arizona
Department of Aerospace and
Mechanical Engineering
Tucson, AZ 85721
USA

Mathematics Subject Classification: 91Axx, 91A05, 91A06, 91A10, 91A12, 91A15, 91A20, 91A22, 91A23, 01A24, 91A40, 91A50, 91A80

Library of Congress Control Number: 2006934736

ISBN-10: 0-8176-4399-0 e-ISBN-10: 0-8176-4553-5
ISBN-13: 978-0-8176-4399-7 e-ISBN-13: 978-0-8176-4553-3

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

www.birkhauser.com (Lap/SB)

Contents

Part VI Numerical Methods and Algorithms in Dynamic Games

Preface

The theory of dynamic games continues to evolve, and one purpose of this volume is to report a number of recent theoretical advances in the field, which are covered in Parts I, II and IV. Another aim of this work is to present some new applications of dynamic games in various areas, including pursuit-evasion games (Part III), ecology (Part IV), and economics (Part V). The volume concludes with a number of contributions in the field of numerical methods and algorithms in dynamic games (Part VI).

With a single exception, the contributions of this volume are outgrowths of talks that were presented at the Eleventh International Symposium on Dynamic Games and Applications, held in Tucson, Arizona, USA, in December 2004, and organized by the International Society of Dynamic Games. The symposium was co-sponsored by the University of Arizona, College of Engineering and Aerospace and Mechanical Engineering, as well as GERAD, Montréal, Canada, and the ISDG Organizing Society.

The volume contains thirty-five chapters that have been peer-reviewed according to the standards of international journals in game theory and applications.

Part I deals with the *theory of dynamic games* and contains six chapters.

Cardaliaguet, Quincampoix, and Saint-Pierre provide a survey of the state-of-the-art of the use of viability theory in the formulation and analysis of differential games, in particular zero-sum games. An important result of viability theory is that many zero-sum differential games can be formulated as viability problems. The main achievements of viability theory are assessed and a number of recent developments are explained, for instance, the formulation of viability problems for hybrid differential games. The chapter contains a substantial list of references.

Chikrii, Matychyn, and Chikrii are concerned with differential games of pursuit and evasion in which one or more players can use impulse controls. The state dynamics are ordinary differential equations that are affected by jumps in the state at discrete instants of time. The method of "resolving functions" provides a general framework for the analysis of such problems and essentially employs the theory of set-valued mappings.

Cardaliaguet investigates a nonzero-sum differential game played by tow players on a line. The dynamics are very simple and players wish to maximize their respective terminal payoffs. In the zero-sum case, the situation is well understood. However, the situation in the nonzero-sum game is completely different. The feedback equilibrium payoffs (FEP) are extremely unstable and small perturbations of the terminal payoffs lead, in general, to large changes in the FEPs.

Ganebny, Kumkov, Patsko, and Pyatko suggest a method for constructing robust feedback controllers for differential games that have linear dynamics with disturbances. The method is based on results from the theory of differential games with geometric constraints on players' controls. The authors also provide an algorithm for constructing a robust control and present simulation results for the practical case of lateral motion control of an aircraft during landing under wind disturbances.

Beck and Filar question a standard assumption of game theory that payoffs in noncooperative matrix games are determined directly by the players' choice of strategies. In real life, palyers may—for several reasons—be unable to execute their chosen strategies. Such inability is referred to as "incompetence." A method for analyzing incompetence in matrix games is suggested, assessed, and demonstrated. The method is motivated by applications where investments in efforts that will decrease incompetence can be made.

Margiocco and Pusillo study the classical Stackelberg game in which the first player is the leader and the second player is the follower. By well-posedness it is meant that the solution exists, is unique, and the approach of "maximizing sequences" is valid. Various general characterizations of Stackelberg well-posedness are provided. Furthermore, hierarchical potential games are considered, and it is proved that some properties of well-posedness are equivalent to the Tikhonov well-posedness of a maximum problem of the potential function.

Part II contains two chapters dealing with the theory of *stochastic differential games.*

Alvaréz and Bardi propose and study a notion of ergodicity for deterministic, zero-sum differential games that extends the one in classical ergodic control theory of systems with two opponent controllers. The connections to the existence of a constant and uniform long-time limit of the value of finite horizon games are established. Moreover, a series of conditions for ergodicity are stated and some extensions to stochastic differential games are provided.

Yeung, Petrosyan, and Yeung address the problem of designing mechanisms that guarantee subgame consistency in the framework of cooperative stochastic differential games with white noise. Recent results of the authors for the case of transferable payoffs (utility) are extended to the—highly intractable—case where payoff cannot be transferred among players.

Part III is devoted to *pursuit-evasion games* and contains four chapters.

Melikyan studies the geometry of pursuit-evasion games on 2-D manifolds. The analysis is done for a variety of game spaces. Because of their simple motion, optimal trajectories are, in general, geodesic lines of the game space manifolds. In some cases there is a singular surface consisting trajectories that are envelopes of family of geodesics. Necessary and sufficient conditions for such singularities are stated. The analysis is based upon viscosity solutions to the Isaacs equations, variational calculus, and geometrical methods.

Shinar, Glizer, and Turetsky propose a class of pursuit-evasion differential games in which two finite sets of possible dynamics of the pursuer and the evader, respectively, are give. These sets are known by both players. The evader chooses her dynamics once before the game starts, and this choice is unobserved by the pursuer. The latter can change his dynamics a finite number of times during the course of the game. Optimal strategies of the players are characterized, and the existence of a saddle point is established.

Crück, Quincampoix, and Saint-Pierre are concerned with pursuit-evasion games with impulsive dynamics (see also Chikrii, Matychyn, and Chikrii in Part I). The system controlled by a player consists of an ordinary differential equation, describing continuous evolution, and a discrete equation that accounts for jumps in the state. For qualitative games, a geometric characterization of the victory (capture) domains are given. For quantitative games, value functions are determined using the Isaacs partial differential inequalities.

Shevchenko studies a game with simple motion in which a pursuer and coalition of evaders move with constant speeds in a plane. The pursuer wishes to minimize the distance to the coalition (defined in a particular way) and terminates the game when distance reduction no longer is guaranteed. The game with two evaders,which can be called a game of alternative pursuit, is studies in detail.

Part IV is devoted to *evolutionary game theory* and *applications*. It contains seven chapters.

Garay examines a situation in which a resident population of interacting individuals is described by a logistic model in which interaction parameters depend on the phenotypes of the individuals. A new mutant clone arises in the population. Among the questions addressed are: what kind of mutuant can (cannot) invade the population, and, if invasion occurs, when does stable co-existence arise? The work establishes a connection between adaptive dynamics and dynamic evolutionary stability.

Mitchell presents an analysis of a resource-consumer model in which individuals are allowed to adaptively vary their resource use as a function of competitor density and strategy. It is demonstrated that habitat specialization, stable minima, community invasibility, and sympatric speciation are more likely when individuals are more efficient at converting resources into viable offspring. The work suggests possible links between species diversity and factors influencing resource conversion efficiency (climate, habitat fragmentation, environmental toxins).

Vincent and Vincent suggest a series of modifications of a well-known model for coexistence. Their starting point is a classical version of the Lotka-Volterra competition equation, which subsequently is made frequency dependent in three different ways and allows the modeling of relative abundance. The purpose is to examine the conditions that determine the relative abundance of species which are in an evolutionary stable state. It is assumed that the ecosystem is at or near an evolutionary equilibrium, and the authors seek evolutionary stable strategies to identify a coalition of individual species.

Hamelin, Bernhard, Nain, and *Wajnberg*, and *Hamelin, Bernhard, Shaiju, and Wajnberg* are concerned with the optimal behavior of foragers that reach a patch at random arrival times. In the first chapter, competition is limited to sharing a common resource. In this case, optimal behavior can be characterized by using a Charnov rule with "carefully chosen" parameters. The second chapter deals with the case of interference competition. Here, an earlier result in the literature is extended to asynchronous arrivals. The resulting problem requires the solution of a war of attrition game with random terminal time. In both chapters, the analysis is valid no matter the arrival law, provided that it is Markovian.

Rael, Vincent, Costantino, and Cushing explore the persistence of corn oil sensitivity in a population of one particular flour beetle. The authors use evolutionary game theory to model and analyze population dynamics and changes in the mean strategy of a population over time. Corn oil sensitivity is a strategy of the flour beetle and is a trait in an evolutionary game that affects the fitness of the organisms. Equilibrium allele frequencies resulting from the game are evolutionary stable strategies and compare favorably with those obtained from experimental data.

Whelan, Brown, and Moll propose a game of resource competition (see also the chapters by Hamelin, et al.). To a forager, food items have three properties that relate to the value of a particular strategy: profitability, richness, and ease of digestion. When foraging on foods that differ with respect to these properties, adjustment (modulation) of gut size and throughput rate leads to a specialization of the digestive system. Modulation of digestive physiology to a particular food type causes different food types to be antagonistic resources. Adjustment of gut volume and processing may promote niche diversification and, in turn, sympatric speciation.

Part V contains ten chapters dealing with the *application of dynamic game theory* to various branches of *economics*.

Beard and McDonald examine the issue of improving the efficiency of water usage. One particular instrument here is the trading of water rights. An important problem is how to design an allocation system for a long period of time such that the desirable properties of the system are sustained at each point of time.

This involves the question of time consistency of the water trading contract (see also the chapter by Yeung, et al. in Part II). A model of dynamic recontracting of water rights is developed and its time consistency properties are assessed.

Alpcan and Basşar investigate a (hybrid) noncooperative game that is motivated by the practical problem of joint power control and base station assignment in code division, multiple access wireless a data networks. Each mobile user's feasible actions include not only the transmission power level, but also the choice of base station. Existence and uniqueness of pure strategy Nash equilibria of the hybrid game are studied. Because of a lack of analytical tractability of the game, numerical analyses and simulations are conducted.

Shen and Başar study pricing issues in communication networks, in particular, the Internet. Pricing is an instrument that can be used to control congestion and prompt efficient network utilization, as well as recover costs of the network service provider and generate revenue for future network development and expansion. A Stackelberg game model with nonlinear pricing strategies is analyzed in which the Internet service provider's pricing policy is designed as an incentive strategy. The problem is studied under complete and incomplete information. In both cases, the game is not, in general, incentive controllable.

Behrens, Caulkins, Feichtinger, and Tragler present a differential game model of international terrorism. A Stackelberg game model is set up (see also the chapter by Shen and Başar). Quantative as well as qualitative analyses are carried out of the incentive strategies of the the two players, called "The West" and an "International terrorist organization." A series of implications of equilibrium behavior are identified, for instance, that outcomes can be considerably improved for "The West" if the game is played for shorter periods of time, or if the terrorist organization's ability to recruit new terrorists can be diminished.

Cellini and Lambertini use a differential game model that represents a variation on the classical Ramsey capital accumulation problem. The game is noncooperative and played by oligopolistic firms. An important aim of the authors is to assess the effects of market power on the long-run equilibrium performance of the economic system. Nonlinear market demand functions are employed, and it is shown that the game admits multiple steady-state equilibria. Also the steady-state relationship between demand curvature and the capital commitments of the players is identified.

Luckraz studies the effects of process imitation on economic growth under stochastic intra-industry spillovers. A main result is that in an economy where the representative industry is a duopoly, R&D spillovers positively affect the economic growth. The game is played in two stages: in the first, the duopolists use Markovian

R&D investment strategies. In the second stage, the duopolists compete in the product market. It turns out that in steady state there is a positive and monotonic relationship between the spillover rate and economic growth.

Kim and El Ouardighi develop a differential game model of collaboration between a manufacturer and a supplier. For each player, the issue is how to allocate her resources among two activities, viz., improving the existing product and developing a new one. In a noncooperative setting, equilibrium time paths of individual player's strategies are identified and confronted with those that would apply in a cooperative game. Numerical simulations suggest that if players cooperate, they prefer to increase the quality of the existing product.

Rubel and Zaccour consider a supply chain (see also the chapter by Kim and El Ourdighi) with one manufacturer and one retailer. The former can choose to sell directly to the consumers (e.g., by e-trade). A noncooperative differential game model is suggested in which players control their respective marketing efforts that are directed at keeping/attracting consumers to a player's preferred channel (i.e., the conventional one for the retailer, the online one for the manufacturer). The authors identify a feedback Nash equilibrium in an infinite-horizon,linear-quadratic game.

Jørgensen and Di Liddo suggest a two-stage game of fashion imitation. A fashion firm sells its product in an exclusive, high-price market for a limited period of time. Afterwards, a less expensive version of the product is introduced in a market with many more consumers. In this market, the fashion firm faces competition from a fringe of firms that sell an illegal imitation of the fashion firm's product. One problem for the fashion firm is that its brand image is diluted by excessive amounts of goods sold in the lower-price market (including its own).

Friesz, Mookherjee, and Rigdon wish to advance the use of differential variational inequalities (DVI) in noncooperative Cournot–Nash dynamic games. With this purpose in mind, the chapter provides a guided tour of how to use DVI in formulating, analyzing, and computing solutions in such games, As a practical example, the author consider a problem of competitive revenue management that is related to pricing and resource allocation in service networks.

Part VI contains six chapters that deal with the development and use of *numerical methods and algorithms* in dynamics games.

Kushner considers a class of stochastic zero-sum differential games with a reflect diffusion system model and ergodic cost criterion. The controls of the players are separated in the dynamics as well as in the cost function. Such ergodic and "Separated" models sometimes occur in risk-sensitive and robust control problems. The numerical method proposed by the author solves a stochastic game for a finite-state Markov chain and an ergodic cost criterion. Essential conditions are nondegeneracy of the diffusion and a weak local consistency condition.

Grantham develops and analyzes a family of gradient transformation differential equation algorithms to find a stationary min-max point of a scalar-valued function. The family includes as special cases the Min-Max Ascent, Newton's method, and a Gradient Enhanced Min-Max algorithm (GEMM) suggested by the author. Using the Stingray function, the stiffness of the gradient transformation family is studied in terms of Lyapunov exponent time histories. GEMM is globally convergent, it is not stiff, and faster than Newton's method and Min-Max Ascent.

McDonald and Grantham examine trajectory following algorithms for differential games. Closed-loop strategies are generated directly by solving ordinary differential equations. The numerical method eliminates the need to solve (i) a min-max optimization problem at each point along the state trajectory; and (ii) nonlinear two-point boundary-value problems. A new algorithm, Efficient Cost Descent, has desirable characteristics that are unique to the trajectory following method. In addition, several issues in the design and implementation of a trajectory following algorithm in a differential game setting are identified and resolved.

Le Ménec addresses a particular problem in the design of optimal guidance laws for radar-guided missiles. The problem is, in general, cast as a zero-sum differential game with noisy measurements of the target location. The sensitivity of several guidance laws to so-called "radome aberrations" is studied. The laws considered are the classical Proportional Navigation, Augmented Proportional Navigation, and two optimal linear-quadratic guidance laws. As to the later, one law is based upon one-sided optimization and the other on a differential game.

Szilagyi reports on a series of computer simulation experiments, based on the author's agent-based simulation tool, for the multi-person Chicken Dilemma game. The agents in the game may cooperate for their collective interest, or they may defect to pursue their individual interests only. The behavior of agents is investigated under a wide range of payoff functions. Results suggest that is quite possible to have a situation in which a vast majority of agents prefer cooperation to defection. A practical application of the game is the problem of using private or public transportation in large cities.

Mikaelyan and Khatchaturyan investigate a partially cooperative, extended-form game. Partial cooperative behavior means that a player can either be cooperating or she can play individually. An algorithmic method for finding optimal behaviors of players and the value of the game is offered. The construction is by backward induction. Some numerical examples are given.

We are grateful to all those researchers who served as reviewers of the manuscripts submitted for possible publication in this volume. They provided excellent reviews and worked very efficiently. We wish to thank the contributing authors for choosing to publish their research work in the Annals of the International Society of Dynamic Games. We also thank Walter Grantham for resolving some technical problems concerning the layout style files. Finally, our thanks go

to the editorial staff of Birkhäuser Boston for their efforts in turning a somewhat heterogenous set of papers into such a nice, cohesive volume. It has been an honorable, but also demanding, task to serve as editors of the volume.

Steffen Jørgensen
Marc Quincampoix
Thomas L. Vincent

Contributors

Tansu Alpcan, Deutsche Telekom Laboratories, Technische Universität Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany

Olivier Alvarez, UMR 60-85, Université de Rouen, 76821 Mont-Saint Aignan Cedex, France

Martino Bardi, Dipartimento di Matematica Pura ed Applicata, Università di Padova, via Belzoni 7, 35131 Padova, Italy

Tamer Başar, Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801, USA

Rodney Beard, School of Economics and the ARC Centre for Complex Systems, The University of Queensland, St. Lucia 4072, Queensland, Australia

Justin Beck, Defence Systems Analysis Division, Defence Science and Technology Organisation, P.O. Box 1500, Edinburgh, South Australia 5111, Australia

Doris A. Behrens, Institute for Mathematical Methods in Economics, ORDYS, Vienna University of Technology, 1040 Vienna, Austria and Department of Economics, Alps Adriatic University of Klagenfurt, 9020 Klagenfurt, Austria

Pierre Bernhard, CNRS and University of Nice Sophia Antipolis–I3S, École Polytechnique de l'Université de Nice Sophia Antipolis, 930 Route des Colles, B.P. 145, 06903 Sophia Antipolis, France

Joel S. Brown, Department of Biological Sciences, University of Illinois at Chicago, M/C 066, 845 West Taylor Street, Chicago, IL 60607, USA

Pierre Cardaliaguet, Laboratoire de Mathématiques, Unité CNRS UMR 6205, Université de Bretagne Occidentale, 6 avenue Le Gorgeu, 29200 Brest, France

Jonathan P. Caulkins, H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA 15213, USA and Qatar Campus, P.O. Box 24866, Doha, Qatar

Roberto Cellini, Department of Economics, University of Catania, Corso Italia 55, 95129 Catania, Italy

Arkadii A. Chikrii, Glushkov Institute of Cybernetics NAS Ukraine, 40 Glushkov Prsp., Kiev 03187, Ukraine

Kirill A. Chikrii, Glushkov Institute of Cybernetics NAS Ukraine, 40 Glushkov Prsp., Kiev 03187, Ukraine

R. F. Costantino, Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

Eva Crück, Laboratoire de Mathématiques, Unité CNRS UMR6205, Université de Bretagne Occidentale, 6 avenue Le Gorgeu, 29200 Brest, France

J. M. Cushing, Department of Mathematics, Program in Applied Mathematics, University of Arizona, Tucson, AZ 85721, USA

Gustav Feichtinger, Institute for Mathematical Methods in Economics, ORDYS, Vienna University of Technology, 1040 Vienna, Austria

Jerzy A. Filar, Centre for Industrial and Applied Mathematics, University of South Australia, G.P.O. Box 2471, Adelaide, South Australia 5001, Australia

T. L. Friesz, Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA 16802, USA

S. A. Ganebny, Institute of Mathematics and Mechanics, Russian Academy of Sciences, S. Kovalevskaya Street 16, Ekaterinburg 620219, Russia

József Garay, Research Group of Ecology and Theoretical Biology, Department of Plant Taxonomy and Ecology, Hungarian Academy of Science and Eötvös University, Pázmány P. sétány 1/C, H-1117 Budapest, Hungary

Valery Y. Glizer, Faculty of Aerospace Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel

Walter J. Grantham, School of Mechanical and Materials Engineering, Washington State University, Pullman, WA 99164-2920, USA

Frédéric Hamelin, CNRS and University of Nice Sophia Antipolis–I3S, École Polytechnique de l'Université de Nice Sophia Antipolis, 930 Route des Colles, B.P. 145 06903 Sophia Antipolis, France

Steffen Jørgensen, Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230, Denmark

Rafik Khachaturyan, Department of Mechanics, Yerevan State University, 1 Alex Manoukyan Street, 375025 Yerevan, Republic of Armenia

Bowon Kim, KAIST Graduate School of Management, 207-43 Cheongryangri, Dongdaemoon-Ku, Seoul 130-722, Korea

S. S. Kumkov, Institute of Mathematics and Mechanics, Russian Academy of Sciences, S. Kovalevskaya Street 16, Ekaterinburg 620219, Russia

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

Luca Lambertini, Department of Economics, University of Bologna, Strada Maggiore 45, 40125 Bologna, Italy and ENCORE, University of Amsterdam, Roetersstraat 11, WB1018 Amsterdam, The Netherlands

Andrea Di Liddo, Department of Economics, Mathematics and Statistics, University of Foggia, Via Gramsci 89-91, 71100 Foggia, Italy

Shravan Luckraz, Faculty of Business, Bond University, Queensland 4229, Australia

Stéphane Le Ménec, EADS/MBDA-F, 78141 Vélizy-Villacoublay, France

Marco Margiocco, Department of Mathematics, University of Genoa, via Dodecaneso 35, 16146 Genoa, Italy

Ivan I. Matychyn, Glushkov Institute of Cybernetics NAS Ukraine, 40 Glushkov Prsp., Kiev 03187, Ukraine

Dale B. McDonald, School of Mechanical and Materials Engineering, Washington State University, Pullman, WA 99164-2920, USA

Stuart McDonald, School of Economics and the ARC Centre for Complex Systems, The University of Queensland, St. Lucia 4072, Queensland, Australia

Arik Melikyan, Institute for Problems in Mechanics, Russian Academy of Sciences, 101-1 Vernadskogo pr., Moscow 117526, Russia

Onik Mikaelyan, Department of Mechanics, Yerevan State University, 1 Alex Manoukyan Street, 375025 Yerevan, Republic of Armenia

William A. Mitchell, Department of Ecology and Organismal Biology, Indiana State University, Terre Haute, IN 47809, USA

Jason Moll, Department of Biological Sciences, University of Illinois at Chicago, M/C 066, 845 West Taylor Street, Chicago, IL 60607, USA

R. Mookherjee, Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA 16802, USA

Philippe Nain, INRIA, 2004 Route des Lucioles, 06902 Sophia Antipolis, France

Fouad El Ouardighi, ESSEC Business School, Avenue Bernard Hirsch, B.P. 105, 95021 Cergy Pontoise, France

V. S. Patsko, Institute of Mathematics and Mechanics, Russian Academy of Sciences, S. Kovalevskaya Street 16, Ekaterinburg 620219, Russia

Leon Petrosyan, Faculty of Applied Mathematics-Control Processes, St. Petersburg State University, 198904 Petrodvorets, St. Petersburg, Russia

Lucia Pusillo, Department of Mathematics, University of Genoa, via Dodecaneso 35, 16146 Genoa, Italy

S. G. Pyatko, Academy of Civil Aviation, Pilotov Street 38, St. Petersburg 196210, Russia

Marc Quincampoix, Laboratoire de Mathématiques, Unité CNRS UMR 6205, Université de Bretagne Occidentale, 6 avenue Le Gorgeu, 29200 Brest, France

R. C. Rael, Program in Applied Mathematics, University of Arizona, Tucson, AZ 85721, USA

M. A. Rigdon, Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA 16802, USA

Olivier Rubel, HEC Montréal and GERAD, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, PQ, H3T 2A7, Canada

Patrick Saint-Pierre, Centre de Recherche Viabilité, Jeux, Contrôle, Université Paris IX-Dauphine, Place du Maréchal de Lattre de Tassigny, F-75775 Paris Cedex 16, France

A. J. Shaiju, CNRS and University of Nice Sophia Antipolis–I3S, École Polytechnique de l'Université de Nice Sophia Antipolis, 930 Route des Colles, B. P. 145, 06903 Sophia Antipolis, France

Hongxia Shen, Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801, USA

Igor Shevchenko, Pacific Fisheries Research Center, 4 Shevchenko Alley, Vladivostok, 690950, Russia and Informatics Department, Far Eastern State University, 8 Sukhanova Street, Vladivostok, 690950, Russia

Josef Shinar, Faculty of Aerospace Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel

Miklos N. Szilagyi, Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721, USA

Gernot Tragler, Institute for Mathematical Methods in Economics, ORDYS, Vienna University of Technology, 1040 Vienna, Austria

Vladimir Turetsky, Faculty of Aerospace Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel

Tania L. S. Vincent, Alaska Pacific University, 4101 University Drive, Anchorage, AK 99508, USA

T. L. Vincent, Department of Aerospace and Mechanical Engineering, University of Arizona, Tucson, AZ 85721, USA

Éric Wajnberg, INRA, 400 Route des Chappes, B.P. 167, 06903 Sophia Antipolis Cedex, France

Christopher J. Whelan, Illinois Natural History Survey, Midewin National Tallgrass Prairie, 30239 South State Highway 53, Wilmington, IL 60481, USA

David W. K. Yeung, Department of Decision Sciences, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong and Centre for Game Theory, St. Petersburg State University, 198904 Petrodvorets, St. Petersburg, Russia

Patricia Melody Yeung, Faculty of Arts and Science, University of Toronto, Toronto, ON, M5S 2E4, Canada

Georges Zaccour, HEC Montréal and GERAD, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, PQ, H3T 2A7, Canada

PART I
Dynamic Game Theory

Differential Games Through Viability Theory: Old and Recent Results

Pierre Cardaliaguet
Laboratoire de Mathématiques
Unité CNRS UMR 6205
Université de Bretagne Occidentale
6 avenue Le Gorgeu, 29200 Brest, France
pierre.cardaliaguet@univ-brest.fr

Marc Quincampoix
Laboratoire de Mathématiques
Unité CNRS UMR 6205
Université de Bretagne Occidentale
6 avenue Le Gorgeu, 29200 Brest, France
marc.quincampoix@univ-brest.fr

Patrick Saint-Pierre
Centre de Recherche Viabilité, Jeux, Contrôle
Université Paris IX-Dauphine
Place du Maréchal de Lattre de Tassigny
F-75775 Paris Cedex 16, France
patrick.saint-pierre@dauphine.fr

Abstract

This article is devoted to a survey of results for differential games obtained through Viability Theory. We recall the basic theory for differential games (obtained in the 1990s), but we also give an overview of recent advances in the following areas: games with hard constraints, stochastic differential games, and hybrid differential games. We also discuss several applications.

1 Introduction

Viability theory is concerned with the study of state constrained problems for controlled systems or differential games. The terminology has been introduced by J.-P. Aubin [6] for controlled systems, but the story of the problem goes back to Poincaré, who introduced the notion of the "shadow of a set" for differential equations. The investigation of such questions for differential games started with the pioneering works of Aubin [1], [2], [5], and has attracted an increasing interest since then. We have tried here to explain the main achievements of the theory, and to give some ideas of its new developments during the last years. We also provide a large—although not exhaustive—list of references.

For a two player differential game, the viability problem in a set K, with target C, consists, for the evader, to ensure that the state of the system reaches the target C before leaving the set of constraints K. On the contrary, the pursuer tries to make the state of the system leave the set of constraints before it reaches the target. Such a problem is a typical "game of kind" in Isaacs' terminology. It is also closely connected with Krasovskii and Subbotin's stable bridges [45], [68].

A first achievement of viability theory is the complete characterization of the solution of the viability game, as well as the numerical approximation of the solution ([23], [22]). It can be shown, in particular, that there is a partition of the set $K\backslash C$ starting from which one, and only one, of the players wins. The common boundary of the partition turns out to be a "semipermeable barrier" in some Isaacs sense [50], [24].

The second important result of viability theory is the fact that many zero-sum differential games can be expressed in terms of viability problems. These ideas are strongly inspired by Frankowska's works (see [37] for instance). As a consequence, such games have a value, and this value can be characterized in terms of a minimal solution of some Hamilton–Jacobi equation (in the viscosity sense [16]). Furthermore, this value can be numerically approximated. For instance, the minimal time problem, in which the pursuer tries to capture the evader in a minimal time, can be treated in that way, without any controllability assumption on the dynamics, even in problems with hard state constraints for the pursuer [25].

Applications of these techniques are manifold. Without trying to be exhaustive, let us first cite the applications in the economic domains, in particular for dynamical economic models [9] and in finance [12], [13]. In engineering science, such a modelization has been used for the study of worst case design [33], [65], [66].

Recent extensions of the viability techniques for differential games have been obtained in several directions. The case of games with separate dynamics and hard state constraints for both players has been solved in [27] and [19] (see also [28] for the numerical approximation). This concerns typically the famous "Lion and Man" game, for which one can prove that there is a value, discontinuous in general.

Viability for stochastic control problems has attracted a lot of attention lately (see [8], [10], [11], [21], [32], [39], [43], [44], [70]). The generalization to differential games, which is not completely understood, has been investigated in [17].

Another very active research field is the domain of uncertain systems that one wants to optimize against the unknown disturbance. This leads to a Min-Max problem, which can naturally be interpreted as a differential game. The viability techniques for solving this problem have been investigated in [35], [47], [57] and [55].

Finally, one of the most challenging problems in the domain is the study of viability problems for hybrid differential games. Substantial contributions have appeared very recently (see [7], [15], [29], [30], [31], [61], [64]). In particular, the worst-case design for these games is now well understood. Note however that the question of the existence of a value is only partially solved. This very new theory has already appeared to be an extremely useful tool for modelization (see for instance [7], [20], [59], [63]).

The paper is organized in two parts. In the first part, we recall the basic theory of viability theory for differential games: the definitions of the discriminating domains and kernels, the alternative theorem, the barrier phenomenon, and the application to the minimal time problem. In the second part, we give an overview of the recent advances in the theory: games with hard state constraints, stochastic viability problems, worst case design and hybrid or impulsive games. We complete the paper by presenting several applications.

2 Basic Viability Theory for Differential Games

In this section, we present the main basic achievements of viability theory. The starting point is the viability game, in which one player wants the state of the system to leave a given set, while the other player wants the state of the system to remain in that set forever. The solution of this game is clearly a set, or, more precisely, a partition of the set of constraints: in one part of this set one of the players wins, while his opponent obtains the victory in the other part.

The geometric characterization of this partition is the first goal of viability theory. It was developed in [23], [24]. Its numerical approximation has been the aim of intensive work: since this was described and explained in detail in the paper [25], we do not recall this subject here. Instead, we briefly show how to apply the main results on the viability game to treat a typical game of kind: the minimal time problem.

2.1 Statement of the Viability Problem

We investigate a differential game with dynamic described by the differential equation

$$\begin{cases} x'(t) = f(x(t), u(t), v(t)), \\ u(t) \in U, \ \ v(t) \in V \end{cases} \tag{1}$$

where $f : \mathbb{R}^N \times U \times V \to \mathbb{R}^N$, U and V being the control sets of the players. Throughout this section, we study the *viability game*. Beside the dynamics, this game has two data: $K \subset \mathbb{R}^N$ which is a closed set of constraints, and $\mathcal{E} \subset \mathbb{R}^N$ which is a closed evasion set. The first player—Ursula, playing with u—wants the state of the system to leave K in finite time while avoiding $\mathcal{E}$. The goal of the second player—Victor, playing with v—is for the state of the system to remain in K until reaching the evasion set $\mathcal{E}$.

In this game the main objects of investigation are the *victory domains* of the players, i.e., the set of initial positions starting from which a player can find a strategy which leads him or her to victory.

We work here in the framework of *nonanticipative strategies* (also called Varaiya–Roxin–Elliot–Kalton strategies). Let

$$\begin{cases} \mathcal{U} = \{u(\cdot) : [0, +\infty[\to U, \text{ measurable function}\} \\ \mathcal{V} = \{v(\cdot) : [0, +\infty[\to V, \text{ measurable function}\} \end{cases} \tag{2}$$

be the sets of time-measurable controls of respectively the first (Ursula) and the second (Victor) player.

Definition 1 (Nonanticipative strategies). A map $\alpha : \mathcal{V} \to \mathcal{U}$ is a nonanticipative strategy (for Ursula) if it satisfies the following condition: for any $s \geq 0$, for any $v_1(\cdot)$ and $v_2(\cdot)$ belonging to $\mathcal{V}$ such that $v_1(\cdot)$ and $v_2(\cdot)$ coincide almost everywhere on $[0, s]$, the images $\alpha(v_1(\cdot))$ and $\alpha(v_2(\cdot))$ coincide almost everywhere on $[0, s]$. Nonanticipative strategies $\beta : \mathcal{U} \to \mathcal{V}$ (for Victor) are defined in a symmetric way.

Assume now that f is continuous and Lipschitz with respect to x. Then, for any $u(\cdot) \in \mathcal{U}$ and $v(\cdot) \in \mathcal{V}$, for any initial position x_0, there exists only one solution to (1). We denote this solution by $x[x_0, u(\cdot), v(\cdot)]$. In order to define the victory domains of the players, we introduce the following notation: if $C \subset \mathbb{R}^N$ is closed and $\varepsilon > 0$, we denote by $C + \varepsilon B$ the set

$$C + \varepsilon B = \{x \in \mathbb{R}^N \mid d_C(x) \leq \varepsilon\},$$

where $d_C(x)$ denotes the distance from x to C.

Definition 2 (Victory domains).
– Victor's victory domain is the set of initial positions $x_0 \in K$ for which Victor can find a nonanticipative strategy $\beta : \mathcal{U} \to \mathcal{V}$ such that for any time-measurable control $u(\cdot) \in \mathcal{U}$ played by Ursula, the solution $x[x_0, u(\cdot), \beta(u(\cdot))]$ remains in K until it reaches $\mathcal{E}$, or remains in K forever if it does not reach $\mathcal{E}$. Namely,

$$\begin{array}{l} \exists \tau \in [0, +\infty],\ \forall t \in [0, \tau),\ x[x_0, u(\cdot), \beta(u(\cdot))](t) \in K \\ \text{and, if } \tau < +\infty,\ \text{ then } x[x_0, u(\cdot), \beta(u(\cdot))](\tau) \in \mathcal{E}. \end{array}$$

– Ursula's victory domain is the set of initial positions $x_0 \in K$ for which Ursula can find a nonanticipative strategy $\alpha : \mathcal{V} \to \mathcal{U}$, positive ε and T, such that, for any $v(\cdot) \in \mathcal{V}$ played by Victor, the solution $x[x_0, \alpha(v(\cdot)), v(\cdot)]$ leaves $K + \varepsilon B$ before reaching the set $\mathcal{E} + \varepsilon B$ and before T.
Namely,

$$\begin{array}{l} \exists \tau \leq T,\ d_K(x[x_0, \alpha(v(\cdot)), v(\cdot)](\tau)) \geq \varepsilon \\ \text{and } \forall t \in [0, \tau],\ x[x_0, \alpha(v(\cdot)), v(\cdot)](t) \notin \mathcal{E} + \varepsilon B. \end{array}$$

In the definition of Victor's victory domain, the solution has to remain in the constraint until reaching the evasion set.

In the definition of Ursula's victory domain, the solution not only has to leave the constraint while avoiding the evasion set, but also has to leave it "sufficiently" (say to leave $K + \varepsilon B$) while remaining "sufficiently far" from the evasion set (say at a distance not smaller than ε), and in a finite time (say not larger than T). Both ε and T have to be independent of Victor's response $v(\cdot)$.

Assumptions on f: In the sequel, we need the following assumptions:

$$\begin{cases} (i) & U \text{ and } V \text{ are compact} \\ (ii) & f : \mathbb{R}^N \times U \times V \to \mathbb{R}^N \text{ is continuous,} \\ (iii) & f \text{ is } \ell\text{-Lipschitz w.r.t. } x, \\ (iv) & \forall x \in \mathbb{R}^N,\ \forall u \in U,\ \text{the set } \bigcup_{v \in V} \{f(x,u,v)\} \text{ is convex.} \end{cases} \tag{3}$$

A sufficient condition for assumption (3-iv) to be satisfied is that V is convex and f is affine with respect to v. We also assume that Isaacs' condition holds:

$$\forall (x,p) \in \mathbb{R}^{2N}, \sup_{u \in U} \inf_{v \in V} \langle f(x,u,v), p \rangle = \inf_{v \in V} \sup_{u \in U} \langle f(x,u,v), p \rangle. \tag{4}$$

2.2 Discriminating Domains

We introduce in this part a class of sets defined by geometric conditions and justify this definition by characterizing it in terms of trajectories.

Definition 3 (Proximal normal). Let D be a closed subset of $\mathbb{R}^N$ and $x \in D$. A vector $p \in \mathbb{R}^N$ is a proximal normal to D at x if $d_D(x+p) = |p|$.

We denote by $\mathcal{NP}_D(x)$ the set of proximal normals to D at x.

This definition means that the ball centered at $x + p$ and of radius $|p|$ is tangent to D at x. Note also that $\mathcal{NP}_D(x)$ is nonempty because $0 \in \mathcal{NP}_D(x)$.

Definition 4 (Discriminating domains). Let $H : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$. A closed set $D \subset X$ is a discriminating domain for H if

$$\forall x \in D,\ \forall p \in \mathcal{NP}_D(x),\ H(x,p) \le 0.$$

We are mainly interested here in the following H:

$$H(x,p) := \begin{cases} \sup_u \inf_v \langle f(x,u,v), p \rangle & \text{if } x \notin \mathcal{E} \\ \min\{\sup_u \inf_v \langle f(x,u,v), p \rangle; 0\} & \text{otherwise} \end{cases}. \tag{5}$$

Theorem 5 (Interpretation of discriminating domains). *We assume that f satisfies (3) and that Isaacs' condition (4) holds. Let $D \subset \mathbb{R}^N$ be a closed set. The following statements are equivalent:*

- *D is a discriminating domain for H defined by (5),*
- *for any initial position $x_0 \in D$, there is a nonanticipative strategy $\beta : \mathcal{U} \to \mathcal{V}$ such that the solution $x[x_0, u, \beta(u)]$ remains in D as long as it has not reached $\mathcal{E}$,*
- *for any initial position $x_0 \in D$, for any positive ε and T, for any nonanticipative strategy $\alpha : \mathcal{V} \to \mathcal{U}$, there is some control $v \in \mathcal{V}$ such that the solution $x[x_0, u, v]$ remains in $K + \varepsilon B$ on the time interval $[0, T]$ until it reaches $\mathcal{E} + \varepsilon B$.*

Remark. If we go back to our viability game, with constraints K and target $\mathcal{E}$, and if a discriminating domain D is contained in K, then D lies in Victor's victory domain, and has an empty intersection with Ursula's victory domain. Indeed, starting from D, Victor can ensure the state of the system to remain in D, hence in K.

2.3 The Alternative Theorem and the Characterization of the Victory Domains

We now characterize the victory domains of the game. For this we have to introduce the notion of discriminating kernel of the constraint set K.

Theorem 6 (Discriminating kernel). *Let $H : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ be a lower semicontinuous map. Any closed subset K of $\mathbb{R}^N$ contains a largest (closed) discriminating domain for H. This set is called the discriminating kernel of K for H and is denoted by $Disc_H(K)$.*

Remarks

- Any discriminating domain for H contained in a closed K is contained in $Disc_H(K)$. Moreover, $Disc_H(K)$ is itself a discriminating domains for H, and $Disc_H(K)$ may be empty if K does not contain any discriminating domains for H.
- One of the most interesting aspects of discriminating kernels consists in the fact that there exist algorithms for approximating these sets. We refer the reader to [25] for a survey of numerical methods.

The main achievement of viability theory for differential games is the following characterization of the victory domains.

Theorem 7 (Characterization of the victory domains). *Assume that f satisfies (3) and that Isaacs' condition (4) holds. Let the constraint set K be closed and let $\mathcal{E}$ be the evasion set. Recall that the Hamiltonian H of the system is defined by (5). Then*

- *Victor's victory domain is equal to $Disc_H(K)$,*
- *Ursula's victory domain is equal to $K \backslash Disc_H(K)$.*

In particular, the victory domains of the two players form a partition of K. A similar Alternative Theorem has been obtained by Krasovskii and Subbotin in the framework of the positional strategies [45]. In fact, the discriminating domains are very close to Krasovskii and Subbotin's stable bridges, while the discriminating kernel is related to the maximal stable bridges. What singularizes the viability approach is the place taken by the geometric characterization of the victory domains, which plays an important role in the sequel.

2.3.1 Isaacs' Semipermeable Barriers

Under the assumption that the common boundary of the victory domains is smooth, Isaacs proved that it satisfies the geometric equation, now known as *Isaacs' equation*:

$$H(x, \nu_x) = 0 \qquad \text{for any } x \text{ on the boundary of the victory domain,} \tag{6}$$

where ν_x stands for the outward normal at x to Victor's victory domain. As a consequence, each player can prevent the state of the system from crossing the boundary in one direction. Such a boundary is called *a semipermeable barrier*.

The boundary $\partial Disc_H(K)$ of the discriminating domain $Disc_H(K)$ actually enjoys this property in a weak sense. This phenomenon was first noticed in [50] for control problems and then extended to differential games in [24].

For simplicity we assume here that there is no evasion set: $\mathcal{E} = \emptyset$.

Proposition 8 (Geometric point of view). *Let x belong to $\partial Disc_H(K)\backslash\partial K$. Then,*

$$(i)\ H(x, p) \leq 0,\ \forall p \in \mathcal{NP}_{Disc_H(K)}(x)$$

and

$$(ii)\ H(x, -p) \geq 0,\ \forall p \in \mathcal{NP}_{\overline{K\backslash Disc_H(K)}}(x).$$

Remark: Since proximal normals are in fact *outward* normals, a proximal normal to $\overline{K\backslash Disc_H(K)}$ is in fact an inward normal to $Disc_H(K)$. Hence putting (i) and (ii) together is a weak formulation of Isaacs' equation (6).

Proposition 9 (Dynamic point of view). *Let x_0 belong to $\partial Disc_H(K)$ but not to ∂K. Then*

- *there is a nonanticipative strategy $\beta : \mathcal{U} \to \mathcal{V}$ for Victor such that*

$$x[x_0, u, \beta(u)](t) \in Disc_H(K) \qquad \forall t \geq 0,\ \forall u \in \mathcal{U},$$

- *there is a nonanticipative strategy $\alpha : \mathcal{V} \to \mathcal{U}$ for Ursula and a time $T > 0$ such that*

$$x[x_0, \alpha(v), v](t) \in \overline{K\backslash Disc_H(K)} \qquad \forall t \in [0, T],\ \forall v \in \mathcal{V}.$$

Remark 10. Roughly speaking, on $[0, T]$, the "solution" $x[x_0, \alpha, \beta]$ remains on $\partial Disc_H(K)$.

Although viability for differential games is now quite well understood, there remain several important and interesting open questions on the subject:

- What kind of regularity can be expected for the boundary of $Disc_H(K)$?
- There is a natural way to define the solution $x[x_0, \alpha, \beta]$ in Remark 10. Is there a maximum principle satisfied by this solution, even in a weak sense?
- In general, $Disc_H(K)$ is not the unique set satisfying Isaacs equation. When is there uniqueness?

2.4 Application to the Minimal Time Problem

Most zero-sum differential games can be expressed in terms of a viability game. We just show how this principle applies for the minimal time problem.

In the minimal time problem, the dynamics of the game is still given by (1). Victor—playing with v—is now the pursuer: he aims at reaching a given target C as fast as possible. Ursula—playing with u—is now the evader: she wants to avoid the target C as long as possible. We assume here that the target $C \subset \mathbb{R}^N$ is closed. If $x(\cdot) : [0, +\infty) \to \mathbb{R}^N$ is a continuous trajectory, the first time $x(\cdot)$ reaches C is

$$\vartheta_C(x(\cdot)) = \inf\{t \geq 0 \mid x(t) \in C\},$$

with convention $\inf(\emptyset) = +\infty$.

The value functions of the game are defined as follows: the *lower value function* is given by

$$\vartheta^\flat(x_0) = \inf_\beta \sup_{u \in \mathcal{U}} \vartheta_C(x[x_0, u, \beta(u)]) \qquad \forall x_0 \notin C,$$

while the *upper value function* is

$$\vartheta^\sharp(x_0) = \lim_{\epsilon \to 0^+} \left(\sup_\alpha \inf_{v \in \mathcal{V}} \vartheta_{C+\epsilon B}(x[x_0, \alpha(v), v]) \right) \qquad \forall x_0 \notin C.$$

Remarks

(1) In general the value functions are discontinuous.
(2) The definition of the upper value is slightly more complicated than that of the lower one. Unfortunately, this formulation is generally necessary for the game to have a value—unless some controllability condition at the boundary of the target holds.

We now explain how to transform the minimal time problem into a viability game in $\mathbb{R}^{N+1}$. For this we define the *extended dynamics* $\widetilde{f} : \mathbb{R}^{N+1} \times U \times V \to \mathbb{R}^{N+1}$ by

$$\tilde{f}(\rho, x, u, v) = \{-1\} \times \{f(x, u, v)\}$$

and

$$\widetilde{K} = \mathbb{R}^+ \times \mathbb{R}^N \qquad \text{and} \qquad \widetilde{C} = \{0\} \times C.$$

For an initial data $\widetilde{x}_0 = (\rho_0, x_0)$, we denote by $\widetilde{x}[\widetilde{x}_0, u, v]$ the solution to

$$\begin{cases} \widetilde{x}' = \widetilde{f}(\widetilde{x}, u, v) \\ \widetilde{x}(0) = \widetilde{x}_0 \end{cases}$$

Remark: $\rho(t) = \rho_0 - t$, the first component of $\widetilde{x}(t)$, is the running cost of the problem.

The viability game associated to the minimal time problem is the following:

- Victor becomes *the evader*: he wants $\widetilde{x}(t)$ to reach the target $\widetilde{C}$ before leaving the constraints $\widetilde{K} := \mathbb{R}^+ \times \mathbb{R}^N$.
- Ursula becomes *the pursuer*: she wants $\widetilde{x}(t)$ to leave $\widetilde{K}$ before reaching the target $\widetilde{C}$.

As seen previously, there is an alternative theorem for this game, and we shall use it to prove the existence of a value for the minimal time problem. The link between the minimal time and the viability game is the following.

Proposition 11. *Let $\widetilde{H} : \mathbb{R}^{N+1} \times \mathbb{R}^{N+1} \to \mathbb{R}$ be defined by*

$$\widetilde{H}(\widetilde{x}, p_x, p_\rho) = \begin{cases} \inf_{v\in V} \sup_{u\in U} \langle f(x,u,v), p_x\rangle - p_\rho & \text{if } \widetilde{x} \notin \widetilde{C} \\ \min\{0, \inf_{v\in V} \sup_{u\in U} \langle f(x,u,v), p_x\rangle - p_\rho\} & \text{otherwise} \end{cases}.$$

Then we have

$$\mathcal{E}pi(\vartheta^\sharp) = Disc_{\widetilde{H}}(\widetilde{K}) = \mathcal{E}pi(\vartheta^\flat),$$

where $\mathcal{E}pi(\phi) = \{(x,\rho) \in \mathbb{R}^{N+1} \to \mathbb{R} \mid \rho \geq \phi(x)\}$ denotes the epigraph of a function ϕ.

Idea of the proof: Let $\widetilde{x} = \widetilde{x}[\widetilde{x}_0, u, v] = (\rho(t), x(t))$ be some solution. We note that $\rho(t) = \rho_0 - t$ because $\widetilde{f} = \{-1\} \times \{f\}$. Hence $\widetilde{x}(t)$ belongs to $\widetilde{K} := \mathbb{R}^+ \times \mathbb{R}^N$ if and only if $\rho(t) = \rho_0 - t \geq 0$.

If (ρ_0, x_0) belongs to Victor's victory domain, then Victor can ensure that $\widetilde{x}(\cdot)$ reaches $\widetilde{C}$ before leaving $\widetilde{K}$. This means that $x(\cdot)$ reaches C before ρ_0. Hence

$$\rho_0 \geq \vartheta_C(x(\cdot)) \geq \vartheta^\sharp(x_0).$$

This shows that Victor's victory domain is contained in the epigraph of $\vartheta^\sharp$. The other inclusions can be proved similarly, by using the different characterizations of Victor's and Ursula's victory domains. □

As a consequence, we have the following result of existence of a value for the game.

Theorem 12. *If assumptions (3) and Isaacs' condition (4) hold, we have*

$$\vartheta^\flat(x_0) = \vartheta^\sharp(x_0) \qquad \forall x_0 \notin C.$$

Furthermore, both functions are the minimal nonnegative viscosity supersolutions to

$$\sup_u \inf_v \langle f(x,u,v), D\vartheta\rangle = 1 \qquad \text{in } \mathbb{R}^N \backslash C.$$

Remark. The last statement is just a way to rewrite the fact that the epigraph of $\vartheta^\flat = \vartheta^\sharp$ is the largest discriminating domain contained in $\mathbb{R}^N \times \mathbb{R}_+$. An introduction to viscosity solutions with a link to differential games can be found in [16].

3 Recent Results on Viability Theory for Differential Games

3.1 Differential Games with State Constraints

Viability theory for differential games can be extended to differential games with separate dynamics and separate state constraints for both players. A typical example of such a game is the famous "Lion and Man" game, in which both lion and man have to remain in the arena. The results for this can be found in [19], [27]. The numerical approximations of state constrained problems is the aim of the work [28].

We assume throughout this section that the game has separate dynamics: the first player, Ursula, playing with u, controls a first system

$$\begin{cases} y'(t) = g(y(t), u(t)), \quad u(t) \in U \\ y(0) = y_0 \in K_U \end{cases} \tag{7}$$

and has to ensure the state constraint $y(t) \in K_U$ to be fulfilled for any $t \geq 0$, where K_U is a closed subset of $\mathbb{R}^l$. On the other hand, the second player, playing with v, controls a second system

$$\begin{cases} z'(t) = h(z(t), v(t)), \quad v(t) \in V \\ z(0) = z_0 \in K_V \end{cases} \tag{8}$$

and has to ensure the state constraint $z(t) \in K_V$ for any $t \geq 0$, where K_V is a closed subset of $\mathbb{R}^m$. We denote by $y[y_0, u]$ the solution to (7), and by $z[z_0, v]$ the solution to (8). We set $N = l + m$.

We study again the viability game: Let $K \subset K_U \times K_V$ be a closed set of constraint and $\mathcal{E} \subset K_U \times K_V$ be a closed evasion set. Ursula wants the state of the system (y, z) to leave K before reaching $\mathcal{E}$, while, on the contrary, Victor wants (y, z) to remain in K as long as it has not reached $\mathcal{E}$. Our aim is again to characterize the victory domains of the players.

To do so, let us set $x = (y, z)$, $f = (g, h)$. Throughout this part, we assume that f, K_U and K_V satisfy the following regularity conditions:

$$\begin{cases} (i) & U \text{ and } V \text{ are compact subsets of some finite-} \\ & \text{dimensional spaces,} \\ (ii) & f : \mathbb{R}^N \times U \times V \to \mathbb{R}^N \text{ is continuous and} \\ & \text{Lipschitz continuous (with Lipschitz constant } M) \\ & \text{with respect to } x \in \mathbb{R}^N, \\ (iii) & \bigcup_u f(x, u, v) \text{ and } \bigcup_v f(x, u, v) \text{ are convex for any } x, \\ (iv) & K_U = \{y \in \mathbb{R}^l,\ \phi_U(y) \leq 0\} \text{ with } \phi_U \in \mathcal{C}^2(\mathbb{R}^l; \mathbb{R}), \\ & \nabla\phi_U(y) \neq 0 \text{ if } \phi_U(y) = 0, \\ (v) & K_V = \{z \in \mathbb{R}^m,\ \phi_V(z) \leq 0\} \text{ with } \phi_V \in \mathcal{C}^2(\mathbb{R}^m; \mathbb{R}), \\ & \nabla\phi_V(z) \neq 0 \text{ if } \phi_V(z) = 0, \\ (vi) & \forall y \in \partial K_U,\ \exists u \in U \text{ such that } \langle \nabla\phi_U(y), g(y, u)\rangle < 0 \\ (vii) & \forall z \in \partial K_V,\ \exists v \in V \text{ such that } \langle \nabla\phi_V(z), h(z, v)\rangle < 0 \end{cases} \tag{9}$$

We note that Isaacs' condition (4) always holds.

We now introduce the notion of *admissible controls and strategies*. For an initial position $(y_0, z_0) \in K_U \times K_V$,

$$\mathcal{U}(y_0) = \{u(\cdot) : [0, +\infty) \to U \text{ measurable} \mid y[y_0, u(\cdot)](t) \in K_U \ \ \forall t \geq 0\}$$

and

$$\mathcal{V}(z_0) = \{v(\cdot) : [0, +\infty) \to V \text{ measurable} \mid z[z_0, v(\cdot)](t) \in K_V \ \forall t \geq 0\}.$$

Under the assumptions (9), it is well known that there are admissible controls for any initial position: namely,

$$\mathcal{U}(y_0) \neq \emptyset \quad \text{and} \quad \mathcal{V}(z_0) \neq \emptyset \qquad \forall x_0 = (y_0, z_0) \in K_U \times K_V.$$

For any $y \in K_U$, we set

$$U(y) = U \text{ if } y \in Int(K_U), \qquad U(y) = \{u \in U \mid g(y, u) \in T_{K_U}(y)\} \text{ if } y \in \partial K_U,$$

where $T_{K_U}(y)$ is the tangent half-space to the set K_U at $y \in \partial K_U$.

Remark 13. We note for later use that, under assumptions (9), the set-valued map $y \rightsquigarrow g(y, U(y))$ is lower semicontinuous with convex compact values (for definitions and properties, see [3]).

Both players now play *admissible nonanticipative strategies*. A nonanticipative strategy α for Ursula is admissible at the point $x_0 = (y_0, z_0) \in K_U \times K_V$ if $\alpha : \mathcal{V}(z_0) \to \mathcal{U}(y_0)$. Admissible nonanticipative strategies β for the second player, Victor, are symmetrically defined. For any point $x_0 \in K_U \times K_V$ we denote by $S_U(x_0)$ and by $S_V(x_0)$ the sets of the nonanticipative strategies for Ursula and Victor respectively.

The victory domains of the players are defined as in Definition 2, but the players now play *admissible* controls and nonanticipative strategies.

As before, the characterization of the victory domains uses the notion of discriminating domains and kernels. However, on account of the state constraints, the definition of the Hamiltonian has to be adapted: for $x = (y, z) \in K_U \times K_V$ and $p \in \mathbb{R}^N$,

$$H(x, p) := \begin{cases} \sup_{u \in U(y)} \inf_{v \in V} \langle f(x, u, v), p\rangle & \text{if } x \notin \mathcal{E} \\ \min\{\sup_{u \in U(y)} \inf_{v \in V} \langle f(x, u, v), p\rangle, 0\} & \text{otherwise} \end{cases}. \tag{10}$$

We note that H is lower semicontinuous, because of assumptions (9) and Remark 13. Therefore, the notion of discriminating kernel of K is well defined.

As before, our main result is the characterization of the victory domains.

Theorem 14 (Characterization of the victory domains). *Assume that $f = (g, h)$, K_U and K_V satisfy (9). Let $K \subset K_U \times K_V$ be the closed constraint and*

$\mathcal{E} \subset K_U \times K_V$ be the closed evasion set. Let H be defined by (10). Then

- *Victor's victory domain is equal to $Disc_H(K)$,*
- *Ursula's victory domain is equal to $K \backslash Disc_H(K)$.*

The main difficulty in proving this theorem is the presence of state constraints for both players. To overcome this difficulty, we use repetitively a technical lemma that we state here because of its usefulness for games with state constraints.

Lemma 15 ([19]). Assume that conditions (9) are satisfied. For any $R > 0$ there exists $C_0 = C_0(R) > 0$ such that, for any $y_0, y_1 \in K_U$ with $|y_0|, |y_1| \le R$, there is a nonanticipative strategy $\sigma : \mathcal{U}(y_0) \longrightarrow \mathcal{U}(y_1)$ with the following property: for any $u_0(\cdot) \in \mathcal{U}(y_0)$

$$|y_0(t) - y_1(t)| + \int_{t_0}^{t} |u_0(s) - \sigma(u_0(\cdot))(s)|\, ds \le C_0 |y_0 - y_1| e^{C_0(t-t_0)},$$

where for simplicity we have set $y_0 = y[y_0, u_0(\cdot)]$ and $y_1 = y[y_1, \sigma(u_0(\cdot))](t)$.

In other words, any admissible control u_0 for y_0 can be approximated by an admissible control $u_1 = \sigma(u_0(\cdot))$ for y_1, in a nonanticipative way.

As in the standard viability for differential games described in the first part, the alternative result for games with state constraints has several counterparts: we have applied it in [28] to prove that the minimal time problem with state constraints has a value, while [19] deals with the Bolza problem. The numerical approximation of such games is studied in [27].

3.2 Viability for Stochastic Differential Games

There are some extensions of viability theory to stochastic controlled systems. For differential games, the main problem is the following: let (X_t) be a controlled stochastic process driven by a stochastic differential equation

$$dX_t = f(X_t, u_t, v_t)dt + \sigma(X_t, u_t, v_t)dW_t \tag{11}$$

and let K be a closed set of constraints. The viability game consists in finding the set of initial position $x_0 \in K$ from which Victor can ensure that $X_t \in K$ *almost surely* for any $t \ge 0$.

For controlled problems, the pioneering works are Gautier and Thibault [39], Aubin and Da Prato [4], Buckdahn, Peng, Quincampoix and Rainer [21] and Aubin, Da Prato and Frankowska [11], to cite only a few. Viability has not been thoroughly investigated for differential games. Some partial characterization results have been obtained by Bardi and Jensen in [17]. This is what we describe now.

Let $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ be a complete stochastic basis endowed with a k-dimensional standard $(\mathcal{F}_t)$-Brownian motion (W_t). An admissible control for Ursula is an $(\mathcal{F}_t)$-progressively measurable process $u : [0, +\infty) \to U$. We denote by $\mathcal{U}$ the set of admissible controls. In the same way, we denote by $\mathcal{V}$ the set of admissible controls for Victor.

Throughout this part, we assume that $f : \mathbb{R}^N \times U \times V \to \mathbb{R}^N$ and $\sigma : \mathbb{R}^N \times U \times V \to \mathbb{R}^{N\times k}$ are continuous, Lipschitz continuous w.r.t. the x-variable and that U and V are compact subsets of some finite-dimensional space. Under this assumption, for any initial position $x_0 \in \mathbb{R}^N$ and any admissible controls $u \in \mathcal{U}$ and $v \in \mathcal{V}$, equation (11) has a unique solution denoted by $X_t^{x_0,u,v}$.

Following [36], we define admissible strategies for Victor as follows: an admissible strategy for Victor is a mapping $\beta : \mathcal{U} \to \mathcal{V}$ such that, if $u_1 \equiv u_2$ a.s. on some time interval $[0, t]$, then $\beta(u_1) \equiv \beta(u_2)$ a.s. on $[0, t]$. We denote by S_V the strategies for Victor.

We now introduce the notion of a second-order proximal set: if $x \in \partial D$, a pair $(p, X) \in \mathbb{R}^N \times \mathbb{R}^{N\times N}$ belongs to the second-order proximal set $\mathcal{NP}_D^2(x)$ to D at x if X is symmetric and if, for any $y \in D$, $p.(y-x)+\frac{1}{2}(y-x).Y(y-x) \leq o(|y-x|^2)$. We note that, if $(p, X) \in \mathcal{NP}_D^2(x)$, then p is a proximal normal to D at x.

Recall that $BUC(\mathbb{R}^N)$ is the set of bounded uniformly continuous maps on $\mathbb{R}^N$.

Theorem 16. *Let $D \subset \mathbb{R}^N$ be a closed set. The following statements are equivalent:*

- *for all $\ell \in BUC(\mathbb{R}^N)$, $\ell \geq 0$, $\ell \equiv 0$ in D, and for all $\lambda > 0$,*

$$\inf_{\beta\in S_V} \sup_{u\in\mathcal{U}} E \int_0^{+\infty} e^{-\lambda t} \ell(X_t^{x_0,u,\beta(u)})dt = 0 \qquad \forall x_0 \in D$$

- $\forall x \in \partial D$, $\forall (p, X) \in \mathcal{NP}_D^2(x)$,

$$\inf_{v\in V} \sup_{u\in U} \left\{ \langle f(x, u, v), p\rangle + \frac{1}{2} Tr\left(\sigma(x, u, v)\sigma^T(x, u, v)X\right) \right\} \leq 0. \qquad (12)$$

The first condition can be understood as a condition of approximate viability for the differential game. The second condition is purely of geometric nature.

Using the ideas of [17] it is not difficult to define the notion of discriminating kernel of a closed set K as the largest closed subset of K satisfying the geometric condition (12), and to characterize it as Victor's victory domain *for an approximate* viability problem.

One should also expect that the two above equivalent conditions in the theorem are equivalent to the following third one (under suitable convexity assumptions):

- for all $x_0 \in D$, there exists some $\beta \in S_V$ such that, for any admissible control u, we have $X_t^{x_0,u,\beta(u)} \in D$ a.s. for any $t \geq 0$.

However this has not been proved up to now.

4 Worst-Case Design

Viability theory can also be used to control systems with imperfectly known uncertainty [55].

We consider the system

$$x'(t) = f(x(t), u(t), y(t), v(t)), \quad x(0) = e \in E_0, \tag{13}$$

where $x \in \mathbb{R}^n$ is the state, $u \in U$ is the control, and $y \in Y$ and $v \in V(y)$ are disturbances (U, Y and $V(y)$ are given subsets of finite-dimensional spaces, $E_0 \subset \mathbb{R}^n$). The main concern of this section is the optimal control problem where the controller wants to minimize (by choosing u) the cost

$$g(T, x(T)) \tag{14}$$

against the worst case of disturbances y and v and initial state $e \in E_0$. We distinguish two types of disturbances:

– *observable uncertainty* y, for which the current realization $y(t) \in Y$ becomes known to the controller;

– *unobservable uncertainty* $v \in V(y)$, for which the realization of $v(t) \in V(y(t))$ remains unknown.

Thus we consider a Min-Max problem or a differential game where the second player wants to maximize (by choosing e, y and v) the cost (14) while the first player—the controller—wants to minimize it (by choosing u). This specific aspect of the information available to the controller implies that the control u should be considered in a feedback form which may depend on the current and the past values of y, but not on v.

For every given control $u(\cdot)$ and observable uncertainty $y(\cdot)$, the unobservable uncertainty $v(\cdot)$ gives rise to a differential inclusion

$$x'(t) \in f(x, u(t), y(t), V(y(t))), \quad x(0) \in E_0, \tag{15}$$

whose solution is a time-dependent tube providing the deterministic estimation of the trajectory. Notice that a (set-valued) tube starting from E_0 is involved even in the case of precisely known initial state $x(0)$, if unobservable uncertainty is present.

Let us formulate the problem in a more precise way. Let $\mathcal{U}_{[t,\theta]}$ be the set of admissible control functions on the interval $[t, \theta]$, that is, the measurable functions with values in U. Similarly, $\mathcal{Y}_{[t,\theta]}$ denotes the set of all measurable selections of Y on $[t, \theta]$, and $\mathcal{V}_{[t,\theta]}(y(\cdot))$ denotes the set of all measurable selections of the mapping $V(y(\cdot))$ (for a given $y(\cdot)$) on the same interval. The following suppositions hold true.

Condition C

C.1. U, Y and $\bar{V}$ are compact subsets of finite-dimensional vector spaces, U is convex, the mapping $y \longrightarrow V(y) \subset \bar{V}$ is compact valued and Lipschitz continuous.

C.2. The function $f : \mathbb{R}^n \times U \times Y \times \bar{V} \mapsto \mathbb{R}^n$ has the form

$$f(x, u, y, v) = f_0(x, y, v) + B(x, y)u,$$

where f_0 and B are continuous, locally Lipschitz in x uniformly with respect to the other variables. The sets $f_0(x, y, V(y))$ are convex. f has linear growth with respect to x, uniformly in u, y, v.

These assumptions will imply that for every $t \in [0, T)$, $e \in \mathbb{R}^n$, $u \in \mathcal{U}_{[t,T]}$, $y \in \mathcal{Y}_{[t,T]}$, and $v \in \mathcal{V}_{[t,T]}(y)$ system (13) has a unique solution on $[t, T]$ starting from e, denoted by $x[t, e; u, y, v](\cdot)$.

The optimal control for initial time t and initial compact set E is sought as a *nonanticipative strategy*, $\alpha : \mathcal{Y}_{[t,T]} \mapsto \mathcal{U}_{[t,T]}$. The guaranteed result obtained by using the strategy α for initial data (t, E) is

$$I(t, E; \alpha) := \sup\{g(T, x[t, e; \alpha(y), y, v](T)),\ e \in E,\ y \in \mathcal{Y}_{[t,T]},\ v \in \mathcal{V}_{[t,T]}(y)\}.$$

Then

$$I(t, E) := \inf_{\alpha} I(t, E; \alpha) \tag{16}$$

is the minimal guaranteed (lower) value that can be achieved starting from the set E at time t.

We denote by $X_{u,y}[t, E](\cdot)$ the reachable set, namely the set of all solutions to (15) evaluated at time t for all possible initial conditions $x_0 \in E$. This is called the *solution tube* of (15) in the set $\mathrm{comp}(\mathbb{R}^n)$ of all compact subsets of R^n. Moreover, for a compact set $Z \subset \mathbb{R}^n$ we define $G : \mathrm{comp}(\mathbb{R}^n) \mapsto \mathbb{R}$ as

$$G(T, Z) = \sup_{z \in Z} g(T, z).$$

Then obviously definition (16) is equivalent to

$$I(t, E) = \inf_{\alpha \in \mathcal{A}_{[t,T]}} \sup_{y \in \mathcal{Y}_{[t,T]}} G(T, X_{\alpha(y),y}[t, E](T)). \tag{17}$$

In this formulation of the original problem there is no unobservable uncertainty. We passed to a problem with complete information (here y is an observable disturbance), but over the solution tubes to differential inclusion (15).

Because of the complexity of the so-obtained problem we can restrict the consideration to the solution tubes of (15) in a given collection of sets $\mathcal{E}$ instead of the whole $\mathrm{comp}(\mathbb{R}^n)$. Namely we are interested in tubes which give an outer estimation of the unknown set of evolutions in a specific class of subsets (boxes ellipsoids, polyhedral sets ...). By doing so we avoid dealing with the geometry of the reachable sets, which could be rather complicated. Thus we come up with the more general problem, formulated below in the case of a target that determines the termination time. In addition we suppose the following.

C.3. The collection $\mathcal{E}$ satisfies conditions:

Condition B.1. The collection $\mathcal{E}$ consists of nonempty compact sets and is closed in the Hausdorff metric. For every compact Z there is some $E \in \mathcal{E}$ containing Z.

Condition B.2. There exists a constant $L_{\mathcal{E}}$ such that for each $\varepsilon > 0$ and each $E \in \mathcal{E}$ there exists $E' \in \mathcal{E}$ for which $E + \varepsilon B \subset E' \subset E + \varepsilon L_{\mathcal{E}} B$.

Condition B.3. For every $Z \in \mathrm{comp}(\mathbb{R}^n)$ there is a unique minimal element of $\mathcal{E}$ containing Z.

The following dynamic programming principle adapts standard arguments.

Proposition 17. *Under conditions* **C.1–C.3***, for every $E \in \mathcal{E}$, $t \in [0, T)$ and $s \in (t, T]$*

$$I_{\mathcal{E}}(t, E) = \inf_{\alpha \in \mathcal{A}_{[t,s]}} \sup_{y \in \mathcal{Y}_{[t,s]}} I_{\mathcal{E}}(s, E(s)),$$
$$\textit{where } E(\cdot) := X_{\alpha(y),y}[t, E](\cdot).$$

One can characterize the value function using Hamilton–Jacobi equations based on Dini epiderivatives which are dual versions of the proximal normals defined in Section 2.2.

Definition 18. Let $\mathcal{E}$ be a closed collection of nonempty compact sets in $\mathbb{R}^n$ and $J : \mathcal{E} \longrightarrow \mathbb{R}$ will be a lower semicontinuous function. Let $E \in \mathcal{E}$ be fixed and let $F : E \mapsto \mathrm{comp}(\mathbb{R}^n)$ be a set-valued field on E.

We define the lower Dini derivative of J at E in the direction of the field F as

$$D^-_{\mathcal{E}} J(E; F) := \liminf_{h,\delta \to 0+} \inf \left\{ \frac{J(E') - J(E)}{h},\ E' \in \mathcal{E},\ (\mathcal{I} + \langle \mathcal{F})(\mathcal{E}) \subset \mathcal{E}' + h\delta\mathcal{B} \right\}.$$

For $\mathcal{E}$ and J as above we define as usual

$$\mathrm{epi}\, J := \{(E, \nu),\ E \in \mathcal{E},\ \nu \in \mathbb{R} :\ \mathcal{J}(\mathcal{E}) \le \nu\} \subset \mathcal{E} \times \mathbb{R},$$

which is also a closed collection of sets.

Define the extended closed collection of compact sets $\hat{} E$ in $\mathbb{R} \times \mathbb{R}^n$ as

$$\hat{} E := \{(t, E),\ t \ge 0,\ E \in \mathcal{E}\}.$$

Theorem 19. *Under conditions* **C.1**–**C.3***, the value function $I_{\mathcal{E}}$ is the uni-que minimal l.s.c. solution of the differential inequality*

$$\sup_{y \in Y} \min_{u \in U} D^-_{\hat{E}} J(t, E; (1, f(\cdot, u, y, V(y)))) \le 0 \qquad \forall\, (t, E) \in \hat{} E, \tag{18}$$

with the condition

$$J(t, E) \ge G(t, E), \quad \forall\, (t, E) \in \hat{E}. \tag{19}$$

That is,

$$I_{\mathcal{E}}(t, E) = \min\{J(t, E),\ J - \textit{l.s.c. solution of } (18), (19)\}.$$

The proof of this theorem is based on a generalization of Theorem 5 but in the context of tubes with values in collections of sets [53] and with results on the regularity of such tubes [54]. In [57], the above worst case design method is applied to the case of a pursuit game with uncomplete information of the pursuer: the pursuer knows only the state of the evader at finitely many times given in advance.

5 Viability for Games with Impulsive Dynamics

We consider a two-player differential game with separated impulsive dynamics where jumps are allowed. The first controller acts on the system

$$\begin{cases} y' &= g(y,u) \\ y^+ &= p(y^-,\mu) \end{cases}, \tag{20}$$

where u and μ are respectively a continuous and a discrete control chosen by Ursula. The solution of (20) is a discontinuous trajectory $t \mapsto y(t)$.

Similarly, the second player, using controls v and ν, controls a system

$$\begin{cases} z' &= h(z,v) \\ z^+ &= q(z^-,\nu) \end{cases}. \tag{21}$$

The outcome of the game is defined in the following way:

- The first player's goal consists in driving the state $x := (y,z)$ into an open set Ω, while keeping it outside a closed set $\mathcal{T}$.
- The second player's aim consists in driving the state in $\mathcal{T}$ while keeping it outside Ω.

We associate with a trajectory $(y(\cdot), z(\cdot))$ of (20)–(21) a payoff

$$\theta^K_{\mathcal{T}}(y(\cdot),z(\cdot)) := \inf\{t : (y(t),z(t)) \in \mathcal{T} \quad \text{and} \quad \forall s \le t, \quad (y(s),z(s)) \in K\},$$

where $K := \mathbb{R}^n \setminus \Omega$.

This kind of game, in the control case has been studied in [15], [29].

We assume that

Assumption 1.

$$\begin{cases} (i) & U \text{ and } M \text{ are compact convex subsets of some finite-dimensional spaces;} \\ (ii) & g : \mathbb{R}^l \times U \mapsto \mathbb{R}^l, \text{ and } p : A_U \times M \mapsto \mathbb{R}^l \text{ are Lipschitz continuous with} \\ & \text{respect to their first variable, and continuous with respect to} \\ & \text{their second variable;} \\ (iii) & g \text{ has linear growth with respect to the first variable;} \\ (iv) & A_U \text{ is compact and for all } y \in A_U, \quad p(y,M) \cap A_U = \emptyset. \end{cases}$$

The last assumption ensures that, after a jump, the trajectory is continuous for some time.

Ursula's system can be characterized by a pair of set-valued functions (G, P), by changing (20) into

$$\begin{cases} y' \in G(y) := \{g(y,u) : u \in U\} \\ y^+ \in P(y^-) := \{p(y^-, \mu) : \mu \in M\} \end{cases}. \tag{22}$$

The reset map P is defined only on A_U, so that the set from which jumps are allowed coincides with the domain of P, denoted by Dom (P).

Similarly, Victor's system can be characterized by a pair of set-valued functions (H, Q), by changing (21) into

$$\begin{cases} z' \in H(y) := \{h(z,v) : v \in V\} \\ z^+ \in Q(y^-) := \{q(z^-, \nu) : \nu \in N\} \end{cases}. \tag{23}$$

The assumptions on g and on h ensure the existence of absolutely continuous solutions defined on $[0, +\infty)$ to the differential inclusions $y' \in G(y)$ and $z' \in H(z)$. Let us denote by $S_G(y_0)$ and respectively $S_H(z_0)$ the set of such solutions.

Definition 20 (Runs and trajectories). We call a *run* of impulse system (G, P) (resp. (H, Q)) with initial condition y_0 (resp. z_0) a finite or infinite sequence $\{\tau_i, y_i, y_i(\cdot)\}_{i\in I}$ (resp. $\{\tau_i, z_i, z_i(\cdot)\}_{i\in I}$) of $(\mathbb{R}^+ \times \mathbb{R}^l \times S_G(\mathbb{R}^l))$ such that for all $i \in I$

$$\begin{cases} y_i'(t) \in G(y_i(t)) \\ y_i(0) = y_i \end{cases} \quad \text{and} \quad \begin{cases} y_i(\tau_i) \in \text{Dom}\,(P) \\ y_{i+1} \in P(y_i(\tau_i)) \end{cases}$$

(resp.

$$\begin{cases} z_i'(t) \in H(z_i(t)) \\ z_i(0) = z_i \end{cases} \quad \text{and} \quad \begin{cases} z_i(\tau_i) \in \text{Dom}\,(Q) \\ z_{i+1} \in Q(z_i(\tau_i)) \end{cases}).$$

A *trajectory* of impulse system (G, P) is a function $y : \mathbb{R} \mapsto \mathbb{R}^l$ associated with a run in the following way:

$$y(t) = \begin{cases} y_0 & \text{if } t < 0 \\ y_i(t - t_i) & \text{if } t \in [t_i, t_i + \tau_i) \end{cases}, \tag{24}$$

where $t_i = \sum_{j<i} \tau_i$. We denote by $S_{G,P}(y_0)$ the set of trajectories with initial condition y_0.

We also assume that the impulse system controlled by Victor satisfies Assumption 1 and can be described similarly by a pair of set-valued maps (H, Q), and we denote by $S_{H,Q}(z_0)$ the set of trajectories with initial condition z_0.

We are now in a position to define strategies.

Definition 21. We call a Varaiya–Roxin strategy (VR-strategy) for Ursula at initial condition $x_0 = (y_0, z_0)$ a map

$$\mathbf{A} : S_{H,Q}(z_0) \longrightarrow S_{G,P}(y_0)$$

such that for any $\theta > 0$, and for any trajectories $z(\cdot)$ and $\tilde{z}(\cdot)$ of $S_{H,Q}(z_0)$ which coincide on $[0, \theta]$, the trajectories $y(\cdot) = \mathbf{A}(z(\cdot))$ and $\tilde{y}(\cdot) = \mathbf{A}(\tilde{z}(\cdot))$ coincide on $[0, \theta]$.
We denote by $\mathcal{A}(x_0)$ the set of VR-strategies for Ursula at x_0.

A VR-strategy for Victor at initial condition $x_0 = (y_0, z_0)$ is defined symmetrically as a map

$$\mathbf{B} : S_{G,P}(y_0) \longrightarrow S_{H,Q}(z_0)$$

such that for any $\theta > 0$, and for any trajectories $y(\cdot)$ and $\tilde{y}(\cdot)$ of $S_{G,P}(y_0)$ which coincide on $[0, \theta]$, the trajectories $z(\cdot) = \mathbf{B}(y(\cdot))$ and $\tilde{z}(\cdot) = \mathbf{B}(\tilde{y}(\cdot))$ coincide on $[0, \theta]$.
We denote by $\mathcal{B}(x_0)$ the set of VR-strategies for Victor at x_0.

We define, for all $x = (y, z) \in \mathbb{R}^n$ and all $D \subset \mathbb{R}^n$ closed, the functions

$$\mathcal{H}(x, D) := \sup_{\pi \in \mathcal{NP}_D(x)} \left\{ \sup_{u \in U} \inf_{v \in V} \langle f(x, u, v), \pi \rangle \right\}, \tag{25}$$

$$\mathcal{L}_V(x, D) := \inf_{\nu \in N} \chi_D(y, q(z, \nu)), \tag{26}$$

$$\mathcal{L}_{UV}(x, D) := \sup_{\mu \in M} \inf \left\{ \chi_D(p(y, \mu), z),\ \inf_{\nu \in N} \chi_D(p(y, \mu), q(z, \nu)) \right\}, \tag{27}$$

where $f(x, u, v) = f((y, z), u, v) = (g(y, u), h(z, v))$, and $\chi_D(\cdot)$ denotes the characteristic function of the set D:

$$\chi_D(x) = \begin{cases} 0 & \text{if } x \in D \\ +\infty & \text{otherwise.} \end{cases}$$

Definition 22 (Impulse discriminating domains). A closed subset D of $\mathbb{R}^n$ is an impulse discriminating domain with target $\mathcal{T}$ if

$$\forall x \in (D \setminus \mathcal{T}), \quad \max\{\min\{\mathcal{H}(x, D),\ \mathcal{L}_V(x, D \cup \mathcal{T})\},\ \mathcal{L}_{UV}(x, D \cup \mathcal{T})\} \leq 0. \tag{28}$$

We obtain the following generalization of Theorem 5 in the impulsive case.

Theorem 23. *Assume the following.*

Assumption 2. *For all time $\theta > 0$, there exist a neighborhood $\mathcal{N}_U$ of* Dom (P) *and a neighborhood $\mathcal{N}_V$ of* Dom (Q) *such that*

$$\begin{array}{ll} (i) & \forall y_0 \in \mathcal{N}_U, \quad \exists y(\cdot) \in S_G(y_0) \quad \textit{such that} \; \exists\, t_0 \le \theta, \quad y(t_0) \in \mathrm{Dom}\,(P) \\ (ii) & \forall z_0 \in \mathcal{N}_V, \quad \exists z(\cdot) \in S_H(z_0) \quad \textit{such that} \; \exists\, t_0 \le \theta, \quad z(t_0) \in \mathrm{Dom}\,(Q)\,. \end{array}$$

Let D and $\mathcal{T}$ be closed subsets of $\mathbb{R}^n$. Under Assumptions 1 and 2, D is an impulse discriminating domain with target $\mathcal{T}$ if and only if for all $x_0 = (y_0, z_0) \in D$, there exists a VR-strategy $\mathbf{B}$ such that for any $y(\cdot) \in S_{G,P}(y_0)$, the trajectory $(y(\cdot), \mathbf{B}(y(\cdot)))$ stays in D as long as $\mathcal{T}$ has not been reached, namely:

$$\forall t \le \inf\{s : (y(s), \mathbf{B}(y(\cdot))(s)) \in \mathcal{T}\}, \qquad (y(t), \mathbf{B}(y(\cdot))(t)) \in D.$$

As in Section 2.4, this theorem can be applied to the characterization and the existence of a value associated with the payoff $\theta_{\mathcal{T}}^{K}(y(\cdot), z(\cdot))$ (cf. [30]).

Remark 24. Differential games with impulsive dynamics is a very new field, and there are several important and interesting open questions on the subject:

- What are the conditions for existence of the value for hybrid differential games?: the players can not only jump but they can also "switch" between several dynamics.
- What are the natural hypotheses allowing one to consider games with many instantaneous jumps (which are forbidden here by assumption 1)?

6 Examples

In this section, we illustrate the different results presented in the previous section through several examples taken from recent studies in ecology, in economics and in finance. The first one deals with the management of renewable resources in the case when we don't really know the precise dynamic which governs the evolution of the resource. The second one is related to the evaluation of a call in the presence of barriers. Such problems can be formalized in the frame of hybrid differential game theory.

6.1 Application I: Management of Renewable Resources

A full discussion of this application can be found in [14].

6.1.1 Verhulst and Graham

The simplest ecological model for managing resources was originated by Graham and taken up by Schaeffer; in this model it is assumed that the exploitation rate of the resource is proportional to the biomass and the economic activity.

Let $x \in \mathbb{R}_+$ denote the biomass of the renewable resource and $v \in \mathbb{R}_+$ the economic effort for exploiting it.

The constraints are of different types and written as follows:

1. *Ecological constraints*: $\forall\ t \geq 0,\ \ 0 \leq x(t) \leq b$ where b is the carrying capacity of the resource.
2. *Economic constraints*: $\forall\ t \geq 0,\ \ cv(t) + C \leq \gamma v(t)\ x(t)$ where $C \geq 0$ is a fixed cost, $c \geq 0$ the unit cost of economic activity and $\gamma \geq 0$ the price of the resource with $\gamma b > c$.
3. *Production constraints*: $\forall\ t \geq 0,\ \ 0 \leq v(t) \leq \overline{v}$, where $\overline{v}$ is the maximal exploitation effort satisfying $\frac{C}{\gamma b - c} \leq \overline{v}$.

Setting $a := \frac{C + c\overline{v}}{\gamma \overline{v}}$, the economic constraints imply that $\forall t \geq 0,\ \ x(t) \in [a, b]$ and $v \in V(x) := \left[\frac{C}{\gamma x - c}, \overline{v}\right]$.

The Verhulst logistic dynamics and the Schaeffer proposal are summed up as follows:

$$\begin{cases} x'(t) = rx(t)\left(1 - \dfrac{x(t)}{b}\right) - v(t)x(t) \\ v(t) \in V(x(t)) := \left[\dfrac{C}{\gamma x(t) - c}, \overline{v}\right] \end{cases}, \tag{29}$$

where the admissible economic effort depends on the very level of the resource. The viability kernel of the interval $[a, b]$ is clearly an interval of the form $[x_-, x_+]$.

The *rigidity* of the economic behavior is expressed through a constraint on the velocity of the economic effort v':

$$\forall\ t \geq 0,\ \ -d \leq v'(t) \leq +d.$$

Taking into account the rigidity of the economic efforts leads us to study what is called the *metasystem* governing the state x and the regulon v which comes with the differential inclusion system

$$\begin{cases} (i) \quad x'(t) = rx(t)\left(1 - \dfrac{x(t)}{b}\right) - v(t)x(t) \\ (ii) \quad v'(t) \in [-d, +d] \end{cases} \tag{30}$$

and the meta-constrained set which is the graph of the set-valued map V:

$$K := (\mathrm{Graph}(V)) = \{(x, v) | v \in V(x)\}. \tag{31}$$

One can prove that the viability kernel of K, Viab(Graph(V)), is of the form $\{(x, v) \in \mathrm{Graph}(V) | x \geq \rho^\sharp(v)\}$, where $\rho^\sharp$ is the solution to the differential equation

$$-d\frac{d\rho^\sharp}{dv} = r\left(1 - \frac{\rho^\sharp(v)}{b}\right) - u\rho^\sharp(v)$$

satisfying the initial condition $\rho^\sharp(v_-) = x_-$.

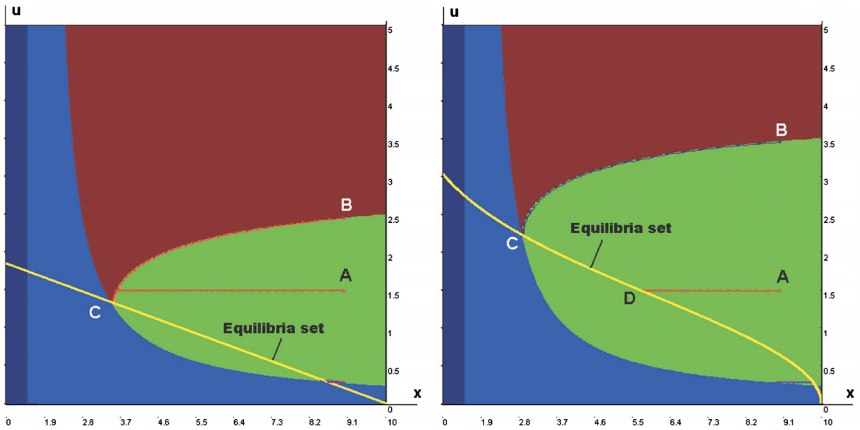


Figure 1: Regulation maps and heavy solutions under Verhulst–Schaeffer metasystems: $x'(t) = rx(t)\left(1 - \frac{x(t)}{b}\right) - v(t)x(t), \quad v'(t) \in [-d, +d]$ on left side and Verhulst-inert meta-systems:
$x'(t) = \sqrt{\alpha}\, x(t) \sqrt{2\log\left(\frac{b}{x(t)}\right)} - v(t)x(t), \ v'(t) \in [-d, +d]$ on right side. The equilibrium sets are respectively the graphs $r\left(1 - \frac{x}{b}\right)$ and $\sqrt{\alpha}\sqrt{2\log\left(\frac{b}{x}\right)}$. Starting from A, heavy solutions minimize $|v'|$ as long as the evolution remains in the viability kernel. The velocity of the economic effort first is null then changes when it reaches the boundary of the viability kernel, at a suitable time for maintaining viability. Here this occurs as soon as it reaches the equilibrium set.

6.2 Towards Dynamical Meta-Games

In fact, we do not really know what dynamical equations govern the evolution of the resource. We could fix Malthusian feedbacks $\widetilde{u}$ in a given class $\widetilde{\mathcal{U}}$ of continuous feedbacks as parameters and study the viability kernel $\mathrm{Viab}_{\widetilde{u}}([a, b])$ of the interval $[a, b]$ under the system

$$\begin{cases} (i) \quad x'(t) = (\widetilde{u}(x(t)) - v(t))x(t) \\ (ii) \quad v(t) \in V(x(t)) := \left[\dfrac{C}{\gamma x(t) - c}, \overline{v}\right] \end{cases}, \tag{32}$$

where the control parameter is v, but, instead of fixing feedbacks, we can study "meta-games" by setting bounds c and d on the velocities of the growth rate $u(t)$ and the exploitation effort $v(t)$, regarded as meta-controls, whereas the meta-states of the meta-game are the triples (x, u, v):

$$\begin{cases} (i) \quad x'(t) = (u(t) - v(t))x(t) \\ (ii) \quad u'(t) \in B(0, c) \\ (iii) \quad v'(t) \in B(0, d) \end{cases} \tag{33}$$

subjected to the viability constraints $u(t) \in \mathbb{R}$ and $\frac{C}{\gamma x(t) - c} \leq v(t) \leq \overline{v}$.

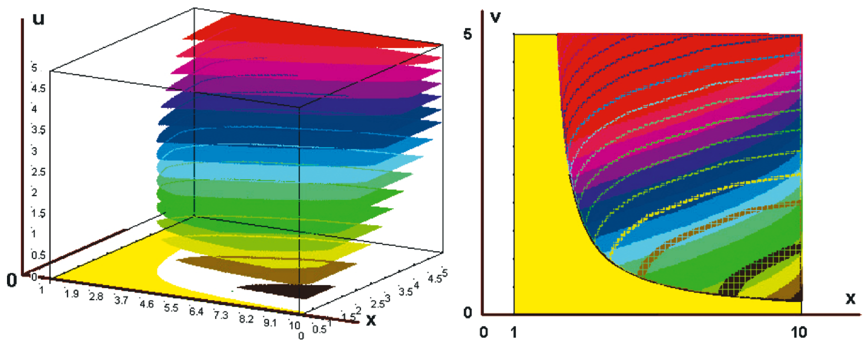


Figure 2: Discriminating Kernel of meta-game (35). The constrained set $\{(x, v, u) \in [a, b] \times [0, \overline{v}] \times [-c, +c] \mid v \geq \frac{C}{\gamma x - c}\}$ translates economic constraints.

Then computing the discriminating kernel of the meta game we get the following result as shown in Figure 2 where the meta-controls are the velocities $|v'(t)| \leq d$ of economic activity bounded by a constant d and the "meta-tyches" corresponding to uncertainty are the velocities of the growth rates of the renewable resources.

6.2.1 The Palikinesia Function

Considering that the velocities of the growth rate $u(t)$ are unknown, we now aim at measuring the maximal norm of this velocity for which a given position (x, u, v) belongs to the discriminating kernel. This can be considered as an inverse problem: If the state initial position is (x_0, v_0), what is the maximal level of the velocities of the growth rate in any future time that the system might face if the velocity of the economic activity is "well chosen?" The palikinesia value is connected to the *"range of environmental tolerance over which an organism can persist"* introduced by Joel Brown when defining the Hutchinsonian Niche in his T2.

The palikinesia function which indicates the maximal level of risk under which viability can be maintained is defined by

$$\psi(x, u, v) := \inf_{\beta} \sup_{u(\cdot)\in\mathcal{U}} \mathbf{esssup}_{t>0} \|v'(t)\|, \tag{34}$$

where the infimum is over the strategies β ensuring the viability.

Proposition 25. *The hypograph of the palikinesia function is the discriminating kernel associated with the following system:*

$$\begin{cases} (i) & x'(t) = (u(t) - v(t))x(t) \\ (ii) & u'(t) \in B(0, |z|) \\ (iii) & v'(t) \in B(0, d) \\ (iv) & z'(t) = 0 \end{cases} \tag{35}$$

subjected to the viability constraints $u(t) \in \mathbb{R}$ *and* $\frac{C}{\gamma x(t) - |z|} \leq v(t) \leq \overline{v}$.

6.3 Application II: Evaluation of Barrier Options

This application is an extension to hybrid games of the guaranteed capture basin method developed in [49], [62], [63].

A put or a call is a contract giving the right to buy (call) or to sell (put) a quantity of an asset of price S at a given date or at any date t before a fixed date T (American put or call). The point is to determine the value W of the contract at the start. Facing the inherent risks, the seller builds up a theoretical portfolio in investing into the underlying asset by self-financing yielding the same losses and profits as the put or call.

6.3.1 Constraints, Payoff and "Target"

1) $\forall (t, S) \in \mathbb{R}_+ \times \mathbb{R}^2_+$, $W \geq \mathbf{b}(t, S) \geq 0$ describes the "floor" constraints where $\mathbf{b} : (t, S) \mapsto \mathbf{b}(t, S) = U(S_0, S_1)$ for American puts and calls, and $\mathbf{b} = 0$ for European puts and calls.
2) $t = 0,\ \ W \geq \mathbf{c}(0, S)$ describes the targets at maturity.
3) The payoff function is $U(S_0, S_1) = (S_1 - K)^+$.

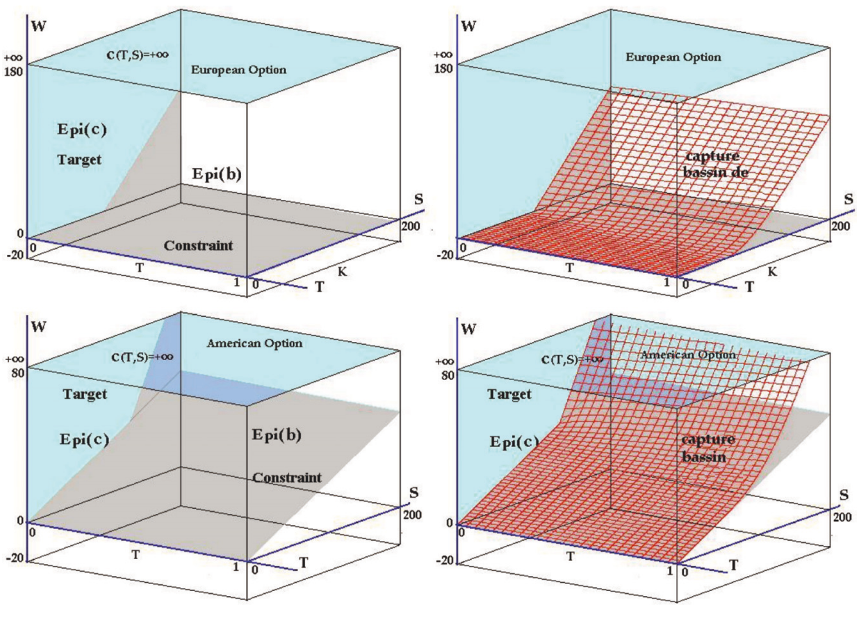


Figure 3: Evaluating a call can be expressed as a guaranteed capture basin problem where the target is the epigraph of the payoff function at $T = 0$ and the constraint is the epigraph of the "floor" function $\mathbf{b}$ (which appears when evaluating American calls).

6.3.2 Objective

We want to determine the set of portfolio strategies $S \to \widetilde{\pi}(S)$ such that, whatever the variations of capital $S(\cdot)$ are, the following conditions hold true:

$$\begin{array}{ll} (i) & \forall t \in [0, T],\ W_{\widetilde{\pi}(\cdot)}(t) \geq \mathbf{b}(T - t, S(t)) \\ (ii) & W_{\widetilde{\pi}(\cdot)}(T) \geq \mathbf{c}(0, S(T)) \end{array}$$

and, amongst them, we want to select the portfolio strategy such that, for all predictable variations, the initial value of the portfolio corresponds to the cheapest capital $V(T, S(0))$, which we identify as the *evaluation function* of the put or call.

6.3.3 The Dynamical Game Describing the Replicating Process

Consider a riskiness asset and an underlying risky asset of respective prices S_0 and S_1. Let $S = (S_0, S_1) \in \mathbb{R}^2$ and $\pi = (\pi_0, \pi_1)\mathbb{R}^2$ be the array of which each component is the total number of assets in a portfolio of value: $W_\pi = \pi_0 S_0 + \pi_1 S_1$. The riskiness and the risky assets are governed by a deterministic and a nondeterministic differential equation

$$\begin{cases} S_0'(t) = S_0(t)\gamma_0(S_0(t)) \\ S_1'(t) = S_1(t)\gamma_1(S_1(t), v(t)). \end{cases}$$

The variations of price $S(t)$ of assets at date t help find the variations $W_{\pi(\cdot)}(t)$ of capital as a function of a strategy $\pi(\cdot)$ of the replicating portfolio. Indeed, the value of the replicating portfolio is given by $W_\pi(t) := \pi_0(t)S_0(t) + \pi_1(t)S_1(t)$.

The self-financing principle of the portfolio reads

$$\forall\ t \geq 0,\ \langle \pi'(t), S(t) \rangle = \pi_0'(t)S_0(t) + \pi_1'(t)S_1(t) = 0$$

so that the value of the portfolio satisfies

$$W'(t) = \langle \pi(t), S'(t) \rangle = \pi_0(t)S_0(t)\gamma_0(S(t)) + \pi_1(t)S_1(t)\gamma_1(S_1(t), v(t)),$$

which is

$$W'(t) = W(t)\gamma_0(S(t)) - \pi_1(t)S_1(t)(\gamma_0(S_0(t)) - \gamma_1(S_1(t), v(t))).$$

Let $\tau(t) := T - t$, then $S(t)$ and $W(t)$ change according to

$$\begin{cases} \tau'(t) & = -1 \\ S_0'(t) & = S_0(t)\gamma_0(S_0(t)) \\ S_1'(t) & = S_1(t)\gamma_1(S_1(t), v(t)) \\ W'(t) & = W(t)\gamma_0(S(t)) - \pi_1(t)S_1(t)(\gamma_0(S_0(t)) - \gamma_1(S_1(t), v(t))), \end{cases}$$

where $\pi(t) \in \Pi(S(t))$ and $v(t) \in Q(S(t))$.

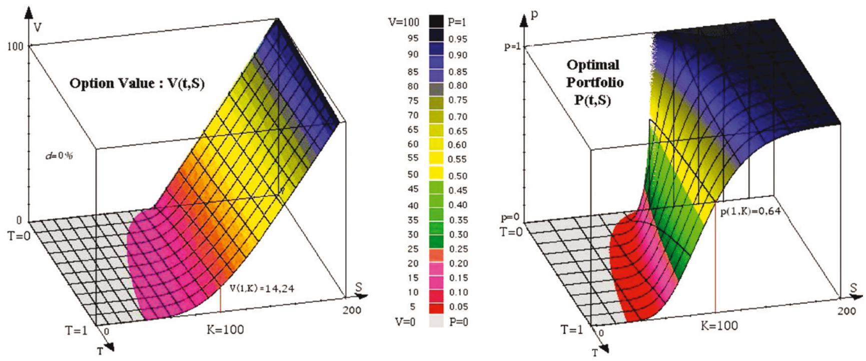


Figure 4: Evaluation function of a European call with uncertainty of Cox, Ross and Rubinstein type and the optimal portfolio: π_1.

It is outside of the scope of this section to develop the numerical point of view allowing us to evaluate calls or puts, as shown in Figure 4. Let us just point out, that the discretization process allows to recover the values obtained by Cox, Ross and Rubinstein's binomial approach as well as the evaluation of contract in more general cases. Let us also quote some recent results linking the stochastic approach and that using differential games that can be found in a forthcoming paper of Aubin and Doss.

Value of S_1	Value of the Call	π_0	π_1
80	4.53	−22.10	0.3332
90	8.61	−34.94	0.4844
100	14.24	−48.17	0.6240
110	21.04	−59.94	0.7363

6.3.4 The Barrier Mechanism

Barriers complicate the evaluation of the replicating portfolio. A barrier is a particular value S^* for the price $S(t)$ beyond which the contract changes. There are four types of options with barriers:

"up and in": the contract becomes effective at the first time t^* when $S(t) < S^*$, $\forall t < t^*$, $S(t^*) = S^*$
"down and in": the contract becomes effective at the first time t^* when $S(t) > S^*$, $\forall t < t^*$, $S(t^*) = S^*$
"up and out": the contract ceases at the first time t^* when $S(t) < S^*$, $\forall t < t^*$, $S(t^*) = S^*$
"down and out": the contract ceases at the first time t^* when $S(t) > S^*$, $\forall t < t^*$, $S(t^*) = S^*$.

The challenge is to evaluate today a contract which will vanish at some unknown future date.

Consider an "up and in" call and introduce a discrete variable $L \in \{0, 1\}$ which "labels" the state of the contract: effective for $L = 1$ or noneffective for $L = 0$. The label L increases because we study "in" options. We consider the hybrid dynamical system

$$\begin{cases} \tau'(t) & = -1 \\ S_0'(t) & = S_0(t)\gamma_0(S_0(t)) \\ S_1'(t) & = S_1(t)\gamma_1(S_1(t), v(t)) \\ L'(t) & = 0 \\ W'(t) = & \begin{cases} W(t)\gamma_0(S(t)) - \pi_1(t)S_1(t)(\gamma_0(S_0(t)) - \gamma_1(S_1(t), v(t))) \\ \qquad \text{if } L(t) = 1 \\ 0 \quad \text{if } L(t) = 0 \end{cases} \end{cases}$$

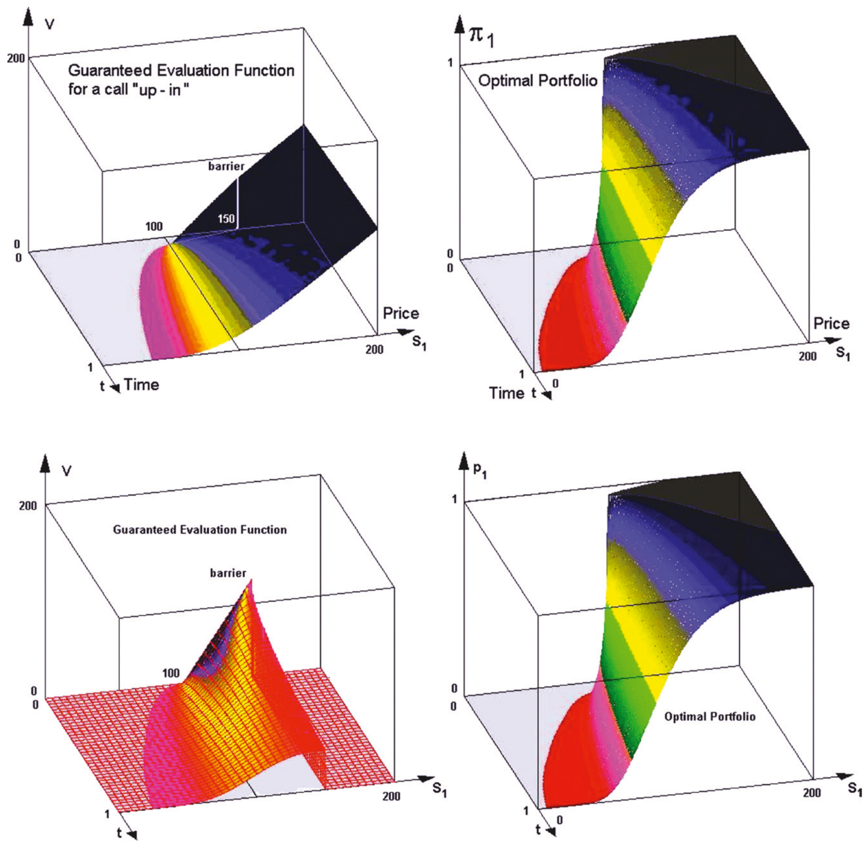


Figure 5: Evaluation function of a European call with barrier "up in" and "up out" with the corresponding optimal strategy: π_1.

$$\begin{cases} \tau^+ &= \tau^- \\ S_0^+ &= S_0^- \\ S_1^+ &= S_1^- \\ L^+ &= \begin{cases} 1 & \text{if } S_1^- \geq S^* \\ L^- & \text{if } S_1^- < S^* \end{cases} \\ W^+ &= W^- \end{cases}$$

Then applying the discriminating kernel algorithm extended to hybrid systems allows us to compute the evaluation function of contracts embedding impulse events as presented in Figure 5.

References

[1] Aubin J.-P. *Contingent Isaacs' Equations of a Differential Game*, in Differential Games and Applications, Eds. T. Başar & P. Bernhard, Lecture Notes in Control and Information Sciences, Springer-Verlag (1989).

[2] Aubin J.-P. *Victory and Defeat in Differential Games*, in Modeling and Control of Systems, Proceedings of the Bellman Continuum, June 1988, Ed. A. Blaquière, Lecture Notes in Control and Information Sciences, Springer-Verlag (1989).

[3] Aubin J.-P. & Frankowska H. SET-VALUED ANALYSIS. Birkhäuser, Boston (1990).

[4] Aubin J.-P. & Da Prato G. *Stochastic viability and invariance.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 17, no. 4, 595–613 (1990).

[5] Aubin J.-P. *Differential games: a viability approach.* SIAM J. Control Optim. 28, no. 6, 1294–1320 (1991).

[6] Aubin J.-P. VIABILITY THEORY. Birkhäuser, Boston (1991).

[7] Aubin J.-P. IMPULSE DIFFERENTIAL INCLUSIONS AND HYBRID SYSTEMS: A VIABILITY APPROACH, Lecture Notes, University of California at Berkeley (1999).

[8] Aubin J.-P. & Da Prato G. *Stochastic Nagumo's Viability Theorem*, Stochastic Analysis and Applications, 13, 1–11 (1995).

[9] Aubin J.-P. DYNAMIC ECONOMIC THEORY: A VIABILITY APPROACH, Springer-Verlag, Berlin and New York (1997).

[10] Aubin J.-P. & Da Prato G. *The Viability Theorem for Stochastic Differential Inclusions*, Stochastic Analysis and Applications, 16, 1–15 (1998).

[11] Aubin J.-P., Da Prato G. & Frankowska H. *Stochastic invariance for differential inclusions.* Set-valued analysis in control theory. Set-Valued Anal. 8, no. 1–2, 181–201 (2000).

[12] Aubin J.-P., Pujal D. & Saint-Pierre P. *Dynamic Management of Portfolios with Transaction Costs under Tychastic Uncertainty*, Preprint (2001).

[13] Aubin J.-P. & Shi Shuzhong, Eds STOCHASTIC AND VIABILITY APPROACHES TO DYNAMICAL UNCERTAINTY IN ECONOMICS AND finance, Presses de l'Université de Beijing (2001).

[14] Aubin J.-P. & Saint-Pierre P. *An introduction to Viability Theory and management of renewable resources, in ch.2, Progress of Artificial Intelligence in Sustainability Science*, Kropp J. & Scheffran J. (Eds.) Nova Science Publ. Inc., New York (2004).

[15] Aubin J.-P., Lygeros J., Quincampoix. M., Sastry S. & Seube N. *Impulse differential inclusions: a viability approach to hybrid systems.* IEEE Trans. Automat. Control 47, no. 1, 2–20 (2002).

[16] Bardi, M., Capuzzo-Dolcetta I. OPTIMAL CONTROL AND VISCOSITY SOLUTIONS OF HAMILTON-JACOBI-BELLMAN EQUATIONS. Systems and Control: Foundations and Applications. Birkhäuser, Boston (1997).

[17] Bardi, M., Jensen R. *A geometric characterization of viable sets for controlled degenerate diffusions.* Calculus of variations, nonsmooth analysis and related topics. Set-Valued Anal. 10, no. 2–3, 129–141 (2002).

[18] Barles G. SOLUTIONS DE VISCOSITÉ DES ÉQUATIONS DE HAMILTON-JACOBI. Mathématiques & Applications (Paris), 17. Springer-Verlag, Paris (1994).

[19] Bettiol P., Cardaliaguet P. & Quincampoix M. Zero-sum state constrained differential games: Existence of value for Bolza problem, Preprint (2004).

[20] Bonneuil N. & Saint-Pierre P. *The Hybrid Guaranteed Capture Basin Algorithm in Economics*, Preprint (2004).

[21] Buckdahn R., Peng S., Quincampoix M. & Rainer C. *Existence of stochastic control under state constraints.* C. R. Acad. Sci. Paris Sér. I Math. 327, no. 1, 17–22 (1998).

[22] Cardaliaguet P., Quincampoix M. & Saint-Pierre P. *Some Algorithms for Differential Games with two Players and one Target*, RAIRO Mathematical Modeling and Numerical Analysis 28, no. 4, 441–461 (1994).

[23] Cardaliaguet P. *A differential game with two players and one target*, SIAM J. Control and Optimization 34, no. 4, 1441–1460 (1996).

[24] Cardaliaguet P. *Non smooth semi-permeable barriers, Isaacs equation and application to a differential game with one target and two players*, Appl. Math. Opti. 36, 125–146 (1997).

[25] Cardaliaguet P., Quincampoix M. & Saint-Pierre P. *Numerical methods for differential games*, in "Stochastic and differential games: Theory and

numerical methods", pp. 177–247. Annals of the international Society of Dynamic Games, M. Bardi, T.E.S. Raghavan, T. Parthasarathy Eds. Birkhäuser (1999).

[26] Cardaliaguet P. & Plaskacz S. *Invariant solutions of differential games and Hamilton-Jacobi equations for time-measurable hamiltonians*, SIAM Journal on Control and Optim. 38, no. 5, 1501–1520 (2000).

[27] Cardaliaguet P., Quincampoix M. & Saint-Pierre P. *Pursuit differential games with state constraints.*, SIAM J. Control and Optimization 39, no. 5, 1615–1632 (2001).

[28] Cardaliaguet P., Quincampoix M. & Saint-Pierre P. *Differential games with state-constraints.* ISDG2002, Vol. I, II (St. Petersburg), 179–182, St. Petersburg State Univ. Inst. Chem., St. Petersburg (2002).

[29] Cruck E. *Target problems under state constraints for nonlinear controlled impulsive systems*, J. Math. Anal. Appl. 270, no. 2, 636–656 (2002).

[30] Crück E., Quincampoix, M. & Saint-Pierre P., *Existence of a value for Pursuit-Evasion Games with Impulsive Dynamics*, Preprint (2003).

[31] Crück E. & Saint-Pierre P. *Non linear Impulse Target Problems under State Constraint: A Numerical Analysis based on Viability Theory*. Set-Valued Analysis. Set-Valued Analysis. 12, no. 4, 383–416 (2004).

[32] Da Prato G. & Frankowska H. *A stochastic Filippov Theorem*, Stochastic Calculus 12, 409–426 (1994).

[33] Doyen L. & Seube N. *Control of uncertain systems under bounded chattering.* Dynam. Control 8, no. 2, 163–176 (1998).

[34] Evans L.C. & Souganidis P.E. *Differential games and representation formulas for solutions of Hamilton-Jacobi Equations* Indiana Univ. Math. J. 282, 487–502 (1984).

[35] Filippova T.F., Kurzhanski A. B., Sugimoto K. & Valyi I. *Ellipsoidal calculus, singular perturbations and the state estimation problems for uncertain systems.* J. Math. Systems Estim. Control 6, no. 3, 323–338 (1996).

[36] Fleming W. & Souganis P. *On the existence of value functions of two-player, zero-sum stochastic differential games.* Indiana Univ. Math. J. 38, no. 2, 293–314 (1989).

[37] Frankowska H. *Lower semicontinuous solutions of Hamilton-Jacobi-Bellman equation*, SIAM J. on Control and Optimization 31, no. 1, 257–272 (1993).

[38] Gao Y., Lygeros J., Quincampoix M. & Seube N. *Approximate stabilisation of uncertain hybrid systems* Proceedings of Computation and control. 6th

international workshop, HSCC 2003, Springer-Verlag, Berlin, Lect. Notes Comput. Sci. 2623, 203–215 (2003).

[39] Gautier S. & Thibault L. *Viability for constrained stochastic differential equations.* Differential Integral Equations 6, no. 6, 1395–1414 (1993).

[40] Haddad G. *Monotone trajectories of differential inclusions with memory*, Isr. J. Math. 39, 83–100 (1981).

[41] Haddad G. *Monotone viable trajectories for functional differential inclusions*, J. Diff. Eq. 42, 1–24 (1981).

[42] Isaacs R. DIFFERENTIAL GAMES Wiley, New York (1965).

[43] Jachimiak W. *Stochastic invariance in infinite dimension*, Preprint (1998).

[44] Kisielewicz M. *Viability theorem for stochastic inclusions*, Discussiones Mathematicae, Differential Inclusions 15, 61–74 (1995).

[45] Krasovskii N.N. & Subbotin A.I. GAME-THEORICAL CONTROL PROBLEMS Springer-Verlag, New York (1988).

[46] Kurzhanski A. B. & Filippova T.F. *On the theory of trajectory tubes—a mathematical formalism for uncertain dynamics, viability and control.* Advances in nonlinear dynamics and control: a report from Russia, 122–188, Progr. Systems Control Theory, 17, Birkhäuser, Boston (1993).

[47] Moitie G., Quincampoix M. & Veliov V. *Optimal control of uncertain discrete systems with imperfect measurements*, IEEE Trans. Automat. Control 47, no. 11, 1909–1914 (2002).

[48] Plaskacz S. & Quincampoix M. *Value-functions for differential games and control systems with discontinuous terminal cost.* SIAM J. Control and Optim. 39, no. 5, 1485–1498 (2001).

[49] Pujal D. & Saint-Pierre P. *L'Algorithme du bassin de capture appliqué pour évaluer des options européennes, américaines ou exotiques*, Revue de l'Association Française de Finance, Paris (2004).

[50] Quincampoix M. *Differential inclusions and target problems*, SIAM J. Control and Optimization 30, 324–335 (1992).

[51] Quincampoix M. & Saint-Pierre *An algorithm for viability kernels in Hölderian case: Approximation by discrete viability kernels*, J. Math. Syst. Est. and Control, Summary: 115–120 (1995).

[52] Quincampoix M. & Veliov V. *Viability with a target: theory and applications*, in Applications of mathematics in engineering, 47–54, Heron Press, Sofia, Bulgaria (1998).

[53] Quincampoix M. & Veliov V. *Open-loop viable control under uncertain initial state information.* Set-Valued Anal. 7, no. 1, 55–87 (1999).

[54] Quincampoix M. & Veliov V. *Solution tubes to differential equations within a collection of sets.* Control and Cybernetics 31, 3 (2002).

[55] Quincampoix M. & Veliov V. *Optimal control of uncertain systems with incomplete information for the disturbance*, Siam J. Control Opti. 43, no. 4, 1373–1399 (2005).

[56] Quincampoix M. & Seube N. *Stabilization of uncertain Control Systems through Piecewise Constant Feedback.* J. Math. Anal. Appl. no. 196, 240–255 (1998).

[57] Rigal S. *A set evolution approach to the control of uncertain systems with discrete time measurement*, Preprint (2004).

[58] Saint-Pierre P. *Approximation of the viability kernel*, Applied Mathematics & Optimisation, 29, 187–209 (1994).

[59] Saint-Pierre P. *Evaluation of crisis, reversibility, alert management for constrained dynamical systems using impulse dynamical systems*. Numerical methods and applications, 255–263, Lecture Notes in Comput. Sci., 2542, Springer, Berlin (2003).

[60] Saint-Pierre P.) *Viable Capture Basin for Studying Differential and Hybrid Games*, (to appear) International Game Theory Review, World Scientific Publishing Company, Singapore (2003).

[61] Saint-Pierre P. *Hybrid Kernels and Capture Basins for Impulse Constrained Systems*, Proceedings of Hybrid Systems (2003).

[62] Saint-Pierre P. *The Guaranteed Hybrid Kernel Algorithm applied to evaluate barrier options in Finance*, Proceeding MTNS04 (2004).

[63] Saint-Pierre P. *Viable capture basin for studying differential and hybrid games: application to finance.* Int. Game Theory Rev. 6, no. 1, 109–136 (2004).

[64] Saint-Pierre P. *Approximation of capture basins for hybrid systems*, in *Proceedings of the 2001 European Control Conference*, Porto, Portugal, September 4–7 (2001).

[65] Seube N., Moitie R. & Leitmann G. *Viability analysis of an aircraft flight domain for take-off in a windshear.* Lyapunov's methods in stability and control. Math. Comput. Modelling 36, no. 6, 633–641 (2002).

[66] Seube N., Moitie R. & Leitmann G. *Aircraft take-off in windshear: a viability approach.* Set-valued analysis in control theory. Set-Valued Anal. 8, no. 1–2, 163–180 (2000).

[67] Serea O.-S. *Discontinuous differential games and control systems with supremum cost.* J. Math. Anal. Appl. 270, no. 2, 519–542 (2002).

[68] Subbotin A. I. GENERALIZED SOLUTIONS OF first-ORDER PDES. THE DYNAMICAL OPTIMIZATION PERSPECTIVE. Translated from the Russian. Systems & Control: Foundations & Applications. Birkhäuser, Boston (1995).

[69] Veliov V. *Sufficient conditions for viability under imperfect measurement.* Set-Valued Anal. 1, no. 3, 305–317 (1993).

[70] Zabczyk J. *Stochastic invariance and consistency of financial models*, preprint Scuola Normale di Pisa (1999).

Differential Games with Impulse Control

Arkadii A. Chikrii
Glushkov Institute of Cybernetics NAS Ukraine
40 Glushkov Prsp., Kiev 03187, Ukraine
dept165@insyg.kiev.ua

Ivan I. Matychyn
Glushkov Institute of Cybernetics NAS Ukraine
40 Glushkov Prsp., Kiev 03187, Ukraine

Kirill A. Chikrii
Glushkov Institute of Cybernetics NAS Ukraine
40 Glushkov Prsp., Kiev 03187, Ukraine

Abstract

This chapter deals with the pursuit games in which players (pursuer, evader, or both) employ impulse control. We consider continuous-time dynamical systems modeled by ordinary differential equations that are affected by jumps in state at discrete instants. The moments of jump comply with the condition for a finite number of jumps in finite time. In so doing, the Dirac delta function is used to describe the impulse control. Such systems represent a special case of hybrid systems. The method of resolving functions provides a general framework for analysis of the above-mentioned problems. This method essentially uses the technique of the theory of set-valued mappings. The following cases are examined in succession: impulse control of the pursuer; impulse control of the evader; impulse control of both players. The problem of approaching a cylindrical terminal set is studied for each case, and the sufficient conditions for its solvability are derived. Obtained results are supported by a model example of the game with simple motion dynamics.

1 Introduction

In this chapter a class of linear pursuit-evasion differential games is studied, where one or both of the players employ impulse control. The impulse character of control is symbolically conveyed by using the Dirac delta function.

Quite an extensive literature covers the impulse dynamic games. Apparently the papers [1,2] were among the first. Krasovskii's methods are used in the works of the reseachers from the Ural school (see, e.g., [3]). In the papers [4,5] the notion of backward procedure is employed. The recent papers [6,7] treat impulse differential games in the framework of the viability of control systems. In these papers the methods developed for continuous dynamic systems are extended to the case of impulse systems. This chapter is no exception. The approach presented here is based on the main ideas of the method of resolving functions [8]. It is aimed at deriving sufficient conditions for hitting a terminal set, rather than obtaining optimality conditions.

The material is arranged as follows. In Section 2, the general problem of hitting a cylindrical terminal set is formulated. The following cases are examined successively in the next three sections: only the pursuer employs impulse control (Section 3); only the evader employs impulse control (Section 4); both players employ impulse control (Section 5). For each case sufficient conditions for hitting the terminal set are derived in the form of an existence theorem. In the proof of each theorem we construct a pursuer's control which ensures hitting the target for all the admissible evader's counteractions. For the case when only the evader employs impulse control (Section 4) it is shown that the pursuit control may be obtained as a solution of an appropriate Fredholm–Volterra integral equation. Section 6 features a model example of the pursuit game with simple motion dynamics. This example illustrates the theoretical results of the paper. Detailed accounts of these results may be found in [9–12].

2 Problem Statement

Let us consider a linear controlled dynamic system, whose evolution is modeled by the following differential equation:

$$\dot{x} = Ax + u - v, \quad x \in \mathbb{R}^m. \tag{1}$$

Here A is a constant square matrix of order m; $u = u(t)$, $v = v(t)$ are the controls of the pursuer and evader respectively. The exact structure of the players' controls is specified in each case treated.

Also a terminal set M^*, $M^* \subset \mathbb{R}^m$, having cylindrical form, is given:

$$M^* = M^0 + M. \tag{2}$$

Here M^0 is a linear subspace from $\mathbb{R}^m$ and the set M belongs to the subspace L being an orthogonal complement to M^0 in $\mathbb{R}^m$.

The goal of the pursuer is to drive a trajectory of the system (1) to the terminal set M^* in a finite time. The goal of the evader is the opposite one: to deviate the system's trajectory from the terminal set as long as possible. We will treat the game, standing on the pursuer's side.

3 Impulse Control of the Pursuer

Let $\{\tau_i\}_{i=0}^{\infty}$ be a sequence of time instants numbered in ascending order, which satisfies the following condition:

Condition 3.1. Any compact segment $[a, b]$ contains a finite number of instants τ_i.

Assume that the pursuer can affect the system (1) only at discrete instants τ_i and that his impact at these moments is of impulse nature, and described using the Dirac delta function [13]:

$$u(t) = \sum_{i=0}^{\infty} u_i \delta(t - \tau_i). \tag{3}$$

Here u_i are jump vectors chosen from the compact set U, $U \subset \mathbb{R}^m$. Hereafter $\delta(t)$ denotes the Dirac delta function.

Let us also assume that the evader's control $v(t)$ is a (Lebesgue) measurable function taking values in the compact set V, $V \subset \mathbb{R}^m$.

Hence the right-hand side of equation (1) additively contains generalized functions. It is known that for any initial condition

$$x_0 = x(0), \tag{4}$$

and given controls of the players there exists a unique solution to system (1), which is absolutely continuous on intervals (τ_{i-1}, τ_i), $i \in \mathbb{N}$ [14]. Here $\mathbb{N}$ is the set of natural numbers, $\tau_0 = 0$.

The function

$$\mathfrak{v}_t(\cdot) = \{v(s) : \ v(s) \in V, \ s \in [0, t], \ v(s) - \text{measurable}\}$$

describes the pre-history of the evader's control at the instant t, $t \geq 0$. Let us assume that the pursuer chooses its control as a *quasistrategy* [8]. This means that each jump vector u_i is defined as $u_i = u(\tau_i, x_0, \mathfrak{v}_{\tau_i}(\cdot))$, where $u(\tau_i, x_0, \mathfrak{v}_{\tau_i}(\cdot))$ is a function of the current instant τ_i, the initial state x_0, and the pre-history of the evader's control $\mathfrak{v}_{\tau_i}(\cdot)$, taking values in the set U. Suppose also that $u(\tau_i, x_0, \mathfrak{v}_{\tau_i}(\cdot))$ takes equal values for almost everywhere equal functions $v(\cdot)$.

Denote by π the operator of the orthogonal projection from $\mathbb{R}^m$ onto L and let e^{At} be the sum of the convergent series

$$e^{At} = \sum_{n=0}^{\infty} \frac{A^n t^n}{n!}.$$

Now we consider the sets

$$W_0(n, v(\cdot)) = W_0(n) = \pi e^{A(\tau_n - \tau_0)} U,$$

$$W_i(n, v(\cdot)) = \pi e^{A(\tau_n - \tau_i)} U - \int_{\tau_{i-1}}^{\tau_i} \pi e^{A(\tau_n - \vartheta)} v(\vartheta) d\vartheta,$$

$$W_i(n) = \bigcap_{v(\cdot) \in V[\tau_{i-1}, \tau_i]} W_i(n, v(\cdot)) = \pi e^{A(\tau_n - \tau_i)} U \overset{*}{-} \int_{\tau_{i-1}}^{\tau_i} \pi e^{A(\tau_n - \vartheta)} V d\vartheta, \tag{5}$$

where $n \in \mathbb{N} \cup \{0\}$, $i = 1, \ldots, n$, and $v(\cdot) \in V[\tau_{i-1}, \tau_i]$. Hereafter $V[a, b]$ is the set of all measurable on $[a, b]$ functions taking values in V, and $X \overset{*}{-} Y = \{x : x + Y \subset X\} = \bigcap_{y \in Y} (X - y)$, the Minkowski difference of two sets. The integral in (5) is an integral of set-valued mapping [15].

The following condition is an analog of Pontrjagin's condition [16], which is well known in differential game theory.

Condition 3.2. Sets $W_i(n)$ are nonempty for all $n, i, n \in \mathbb{N} \cup \{0\}$, $i = 0, \dots, n$.

By virtue of Condition 3.2, one can select some element $w_i(n)$ from set $W_i(n)$, $i = 0, \dots, n$. Let us fix some collection $\omega = \omega(n) = \{w_i(n)\}_{i=0}^{n}$ and set

$$\xi(n, x, \omega) = \pi e^{A(\tau_n - \tau_0)} x + \sum_{i=0}^{n} w_i(n). \tag{6}$$

Now we introduce the functions

$$\begin{aligned} \tilde{\alpha}_i(n, x, v(\cdot), \omega) = \sup\{\alpha \geq 0 :\ & \alpha[M - \xi(n, x, \omega)] \cap \\ & [W_i(n, v(\cdot)) - w_i(n)] \neq \varnothing\}, \end{aligned} \tag{7}$$

and denote

$$k = k(n, x, v(\cdot), \omega) = \min \left\{ j \in \{0, \dots, n\} :\ \sum_{i=0}^{j} \tilde{\alpha}_i(n, x, v(\cdot), \omega) \geq 1 \right\}. \tag{8}$$

If the inequality in braces fails for all j, $j \in \{0, \dots, n\}$, we set $k = n + 1$.

Now we proceed to the definition of the resolving functions [8]. If $\xi(n, x, \omega) \in M$, then

$$\alpha_i(n, x, v(\cdot), \omega) = \frac{1}{n+1}, \quad i = 0, \dots, n.$$

If, otherwise, $\xi(n, x, \omega) \notin M$, then we set

$$\alpha_i(n, x, v(\cdot), \omega) = \begin{cases} \tilde{\alpha}_i(n, x, v(\cdot), \omega), & i = 0, \dots, k-1, \\ 1 - \sum_{j=0}^{k-1} \tilde{\alpha}_j(n, x, v(\cdot), \omega), & i = k, \\ 0, & i = k+1, \dots, n, \end{cases} \tag{9}$$

where $0 < k < n$.

We will examine separately the cases $k = 0$, $k = n$, and $k = n+1$. If $\xi(n, x, \omega) \notin M$ and $k = 0$, then

$$\alpha_i(n, x, v(\cdot), \omega) = \begin{cases} 1, & i = 0, \\ 0, & i = 1, \dots, n. \end{cases} \tag{9$'$}$$

If $\xi(n, x, \omega) \notin M$ and $k = n$, then

$$\alpha_i(n, x, v(\cdot), \omega) = \begin{cases} \tilde{\alpha}_i(n, x, v(\cdot), \omega), & i = 0, \dots, n-1 \\ 1 - \sum_{j=0}^{n-1} \tilde{\alpha}_j(n, x, v(\cdot), \omega), & i = n. \end{cases} \tag{9$''$}$$

If $\xi(n, x, \omega) \notin M$ and $k = n + 1$, we set

$$\alpha_i(n, x, v(\cdot), \omega) = \tilde{\alpha}_i(n, x, v(\cdot), \omega), \quad i = 0, \dots, n.$$

By definition of the resolving function $\alpha_0(n, x, v(\cdot), \omega)$ does not depend on the evader's control, i.e.,

$$\alpha_0(n, x, v(\cdot), \omega) = \alpha_0(n, x, \omega).$$

Lemma 3.1. *For the game (1), (2) with impulse control of the pursuer (3), let the sets M, U be convex, the Condition 3.2 hold, and $w_i(n) \in W_i(n)$, $n \in \mathbb{N} \cup \{0\}$, $i = 0, \dots, n$. Then*

$$\alpha_i(n, x, v(\cdot), \omega)(M - \xi(n, x, \omega)) \cap (W_i(n, v(\cdot)) - w_i(n)) \neq \varnothing \tag{10}$$

for all $n \in \mathbb{N} \cup \{0\}$, $i = 0, \dots, n$, $x \in \mathbb{R}^m$, $v(\cdot) \in V[\tau_0, \tau_n]$.

Proof. One can see from (9) that

$$\alpha_i(n, x, v(\cdot), \omega) = \tilde{\alpha}_i(n, x, v(\cdot), \omega), \ i = 0, \dots, k - 1.$$

Then the statement of the lemma follows from the definition of functions (7). For $i = k+1, \dots, n$ $\alpha_i(n, x, v(\cdot), \omega) = 0$ and therefore relationship (10) is true, since the sets $W_i(n, v(\cdot)) - w_i(n)$ contain the null vector.

Now let $i = k$. Consider the set $\mathcal{A} = \{\alpha \geq 0 : \alpha[M - \xi(n, x, \omega)] \cap [W_i(n, v(\cdot)) - w_i(n)] \neq \varnothing\}$. By virtue of the lemma's assumptions, the sets $M - \xi(n, x, \omega)$ and $W_i(n, v(\cdot)) - w_i(n)$ are convex. Let us suppose that $x_1 \in \alpha_1[M - \xi(n, x, \omega)] \cap [W_i(n, v(\cdot)) - w_i(n)]$, $x_2 \in \alpha_2[M - \xi(n, x, \omega)] \cap [W_i(n, v(\cdot)) - w_i(n)]$, $\alpha_1, \alpha_2 \in \mathcal{A}$. Due to the convexity of the sets $M - \xi(n, x, \omega)$ and $W_i(n, v(\cdot)) - w_i(n)$, the inclusion $\lambda x_1 + (1 - \lambda)x_2 \in (\lambda\alpha_1 + (1 - \lambda)\alpha_2)[M - \xi(n, x, \omega)] \cap [W_i(n, v(\cdot)) - w_i(n)]$ holds for any $\lambda \in [0, 1]$. From this it follows that $\lambda\alpha_1 + (1 - \lambda)\alpha_2 \in \mathcal{A}$ and the set $\mathcal{A}$ is convex. Convexity of the set $\mathcal{A}$ implies that $\alpha[M - \xi(n, x, \omega)] \cap [W_i(n, v(\cdot)) - w_i(n)] \neq \varnothing$ for all α such that $0 \leq \alpha \leq \tilde{\alpha}_k(n, x, v(\cdot), \omega)$. One can see from formulas (8), (9), that $\alpha_k(n, x, v(\cdot), \omega) \leq \tilde{\alpha}_k(n, x, v(\cdot), \omega)$ and therefore Lemma 3.1 is true in the case $i = k$ too. □

Let us introduce the function

$$N(x, \omega) = \min \left\{ n \in \mathbb{N} \cup \{0\} : \sum_{i=0}^{n} \inf_{v(\cdot) \in V[\tau_{i-1}, \tau_i]} \alpha_i(n, x, v(\cdot), \omega) = 1 \right\}.$$

If the equality in braces fails for all n, then we set $N(x, \omega) = +\infty$.

Theorem 3.1. *In the game (1), (2) with impulse control of the pursuer (3), let the sets M and U be convex, Condition 3.2 hold, and also let $N(x_0, \omega) < +\infty$ for given initial state x_0 and some collection ω. Then a trajectory of the system (1) starting at x_0 can be driven to the terminal set (2) at the time instant $\tau_{N(x_0, \omega)}$.*

Proof. Denote $N = N(x_0, \omega)$ and fix some function $v(\cdot)$, $v(\cdot) \in V[\tau_0, \tau_N]$.

To begin, we consider the case when $\xi(N, x_0, \omega) \notin M$.

Let us set $K = k(N, x_0, v(\cdot), \omega)$. Then, according to (8), (9),

$$\sum_{i=0}^{K} \alpha_i(N, x_0, v(\cdot), \omega) = 1.$$

Now we select vector u_0 satisfying the inclusion

$$\pi e^{A(\tau_N - \tau_0)} u_0 - w_0(N) \in \alpha_0(N, x_0, v(\cdot), \omega)[M - \xi(N, x_0, \omega)], \tag{11}$$

and for $i = 1, \ldots, K$ we select jump vectors u_i satisfying inclusions

$$\begin{aligned} &\pi e^{A(\tau_N - \tau_i)} u_i - \int_{\tau_{i-1}}^{\tau_i} \pi e^{A(\tau_N - \vartheta)} v(\vartheta) d\vartheta - w_i(N) \\ &\quad \in \alpha_i(N, x_0, v(\cdot), \omega)[M - \xi(N, x_0, \omega)]. \end{aligned} \tag{12}$$

By virtue of Lemma 3.1 and in view of Condition 3.2 and formulas (5), (7), there exist solutions to inclusions (11), (12).

For $i = K + 1, \ldots, N$ let us select vectors u_i satisfying the equations

$$\pi e^{A(\tau_N - \tau_i)} u_i - \int_{\tau_{i-1}}^{\tau_i} \pi e^{A(\tau_N - \vartheta)} v(\vartheta) d\vartheta = w_i(N). \tag{13}$$

We can do this due to Condition 3.2.

Now let us write for the system (1) the Cauchy formula using properties of the Dirac delta function:

$$\begin{aligned} \pi x(\tau_N) &= \pi e^{A(\tau_N - \tau_0)} x_0 + \int_{\tau_0}^{\tau_N} \pi e^{A(\tau_N - \vartheta)} (u(\vartheta) - v(\vartheta)) d\vartheta \\ &= \pi e^{A(\tau_N - \tau_0)} x_0 + \sum_{i=0}^{N} \pi e^{A(\tau_N - \tau_i)} u_i - \int_{\tau_0}^{\tau_N} \pi e^{A(\tau_N - \vartheta)} v(\vartheta) d\vartheta \\ &= \pi e^{A(\tau_N - \tau_0)} x_0 + \pi e^{A(\tau_N - \tau_0)} u_0 \\ &\quad + \sum_{i=1}^{N} \left[\pi e^{A(\tau_N - \tau_i)} u_i - \int_{\tau_{i-1}}^{\tau_i} \pi e^{A(\tau_N - \vartheta)} v(\vartheta) d\vartheta \right]. \end{aligned} \tag{14}$$

Let us add and subtract the term $\sum_{i=0}^{N} w_i(N)$ from the right-hand side of (14). Then, taking into account the convexity of the set M and formulas (11)–(13), we obtain the following inclusion:

$$\begin{aligned} \pi x(\tau_N) \in \xi(N, x_0, \omega) &\left[1 - \sum_{i=0}^{N} \alpha_i(N, x_0, v(\cdot), \omega) \right] \\ &+ \sum_{i=0}^{N} \alpha_i(N, x_0, v(\cdot), \omega) M = M. \end{aligned}$$

This is equivalent to the inclusion $x(\tau_N) \in M^*$.

Now let us assume that $\xi(N, x_0, \omega) \in M$. We select vector u_0 satisfying the equation

$$\pi e^{A(\tau_N-\tau_0)}u_0 = w_0(N).$$

As the jump vectors u_i, $i = 1, \dots, N$, we take solutions of the equations (13). Then the inclusion $\pi x(\tau_N) \in M$ will immediately follow from formula (14). □

4 Impulse Control of the Evader

Here we suppose that the pursuer's control is a measurable function of time, taking values in the compact set U, $U \subset \mathbb{R}^m$. The evader affects the system (1) only at the instants $\{\tau_i\}$ and his control is of the form

$$v(t) = \sum_{i=0}^{\infty} v_i \delta(t-\tau_i), \tag{15}$$

where the jump vectors v_i belong to the compact set V, $V \subset \mathbb{R}^m$, $\tau_0 = 0$.

Analogously to the case analyzed in Section 3, a solution to the system (1) exists for any initial condition (4). Moreover, it is unique and absolutely continuous on intervals (τ_i, τ_{i+1}), $i \in \mathbb{N} \cup \{0\}$.

Let us introduce the function $n(t) = \max\{i \in \mathbb{N} \cup \{0\} : \ \tau_i \le t\}$ and consider the set

$$\mathfrak{v}_t = \{v_i : \ i = 0, \dots, n(t), v_i \in V\},$$

which describes the pre-history of the evader's control at instant t, $t \ge 0$. As before, we assume that the pursuer employs quasistrategies, i.e., there exists a map $U(t, x_0, \mathfrak{v}_t)$ such that $u(t) = U(t, x_0, \mathfrak{v}_t)$, $t \ge 0$.

Consider the sets

$$\begin{aligned} W_i(t, v) &= \int_{\tau_i}^{\tau_{i+1}} \pi e^{A(t-\vartheta)} U d\vartheta - \pi e^{A(t-\tau_i)} v, \quad i = 0, 1, \dots, n(t) - 1, \\ W_{n(t)}(t, v) &= \int_{\tau_{n(t)}}^{t} \pi e^{A(t-\vartheta)} U d\vartheta - \pi e^{A(t-\tau_{n(t)})} v, \end{aligned} \tag{16}$$

$$\begin{aligned} W_i(t) &= \int_{\tau_i}^{\tau_{i+1}} \pi e^{A(t-\vartheta)} U d\vartheta \overset{*}{-} \pi e^{A(t-\tau_i)} V, \quad i = 0, 1, \dots, n(t) - 1, \\ W_{n(t)}(t) &= \int_{\tau_{n(t)}}^{t} \pi e^{A(t-\vartheta)} U d\vartheta \overset{*}{-} \pi e^{A(t-\tau_{n(t)})} V. \end{aligned} \tag{17}$$

Let us introduce a set $\mathfrak{T} = \{t \ge 0 : \ W_i(t) \ne \varnothing, \ i = 0, \dots, n(t)\}$. Note that as a rule $\tau_i \notin \mathfrak{T}$ for all $i \in \mathbb{N} \cup \{0\}$, since the sets $W_i(\tau_i)$ are nonempty only if the set πV consists of a single element.

Condition 4.1. The set $\mathfrak{T}$ is nonempty.

For each instant t, $t \in \mathfrak{T}$, let us fix a collection $\omega = \omega(t) = \{w_i(t) : \; w_i(t) \in W_i(t), \; i = 0, \ldots, n(t)\}$ and set

$$\xi(t, x, \omega) = \pi e^{A(t-\tau_0)} x + \sum_{i=0}^{n(t)} w_i(t). \tag{18}$$

We introduce the functions

$$\tilde{\alpha}_i(t, x, v, \omega) = \sup\{\alpha \geq 0 : \; \alpha(M - \xi(t, x, \omega)) \cap (W_i(t, v) - w_i(t)) \neq \varnothing\}. \tag{19}$$

Let us set

$$k = k(t, x, v, \omega) = \min \left\{ j \in \{1, \ldots, n(t)\} : \; \sum_{i=1}^{j} \tilde{\alpha}_i(t, x, v, \omega) \geq 1 \right\}. \tag{20}$$

If the inequality in braces fails for all j, $j \in \{1, \ldots, n(t)\}$, than we set $k = n(t)+1$.

Now we proceed to the definition of the resolving function [8]. Suppose that $\xi(t, x, \omega) \in M$. Then

$$\alpha_i(t, x, v, \omega) = \frac{1}{n(t)+1}, \quad i = 0, \ldots, n(t).$$

If $\xi(t, x, \omega) \notin M$, then we set

$$\alpha_i(t, x, v, \omega) = \begin{cases} \tilde{\alpha}_i(t, x, v, \omega), & i = 0, \ldots, k-1, \\ 1 - \sum_{j=0}^{k-1} \tilde{\alpha}_j(t, x, v, \omega), & i = k, \\ 0, & i = k+1, \ldots, n(t), \end{cases} \tag{21}$$

where $0 < k < n(t)$.

We will analyze separately the cases $k = 0$, $k = n(t)$, and $k = n(t) + 1$. If $\xi(t, x, \omega) \notin M$ and $k = 0$, then

$$\alpha_i(t, x, v, \omega) = \begin{cases} 1, & i = 0, \\ 0, & i = 1, \ldots, n(t). \end{cases}$$

If $\xi(t, x, \omega) \notin M$ and $k = n(t)$, then

$$\alpha_i(t, x, v, \omega) = \begin{cases} \tilde{\alpha}_i(t, x, v, \omega), & i = 0, \ldots, n(t) - 1 \\ 1 - \sum_{j=0}^{n(t)-1} \tilde{\alpha}_j(t, x, v, \omega), & i = n(t). \end{cases}$$

In the case when $\xi(t, x, \omega) \notin M$ and $k = n(t) + 1$, let us set

$$\alpha_i(t, x, v, \omega) = \tilde{\alpha}_i(t, x, v, \omega), \quad i = 0, \ldots, n(t). \tag{22}$$

One can easily show that if set M is convex, then the following relationship holds:

$$\alpha_i(t, x, v, \omega)(M - \xi(t, x, \omega)) \cap (W_i(t, v) - w_i(t)) \neq \varnothing, \tag{23}$$

where $t \in \mathfrak{T}$, $i = 0, 1, \dots, n(t)$, $x \in \mathbb{R}^m$, $v \in V$, $w_i(t) \in W_i(t)$. This assertion can be proved similarly to Lemma 3.1. In so doing, convexity of the sets $W_i(t, v)$ is proved in view of the fact that the integral of a uniformly bounded compact-valued mapping is a convex compact set (*Aumann theorem* [17]).

Consider the set

$$T(x, \omega) = \left\{ t \in \mathfrak{T} : \sum_{i=0}^{n(t)} \inf_{v \in V} \alpha_i(t, x, v, \omega) = 1 \right\}. \tag{24}$$

If the equality in braces fails for all $t \in \mathfrak{T}$, then we set $T(x, \omega) = \varnothing$.

Theorem 4.1. *In the game (1), (2) with impulse control of the evader (15), let the set M be convex, Condition 4.1 hold, $\tau_i \notin \mathfrak{T}$ for all $i \in \mathbb{N} \cup \{0\}$, and let the set $T(x_0, \omega)$ be nonempty for given initial state x_0 and some collection ω. Then for any fixed $T \in T(x_0, \omega)$ a trajectory of the system (1) can be driven to the terminal set (2) at the time instant T.*

Proof. Let us denote $N = n(T)$ and fix some collection of jump vectors $\{v_i\}$.

To begin, we assume that $\xi(T, x_0, \omega) \notin M$. Denote $K = k(T, x_0, v, \omega)$. Then, in view of formulas (20), (21),

$$\sum_{i=0}^{K} \alpha_i(T, x_0, v_i, \omega) = 1.$$

Now we suppose that $K = N$. On the intervals $[\tau_i, \tau_{i+1})$, $i = 0, 1, \dots, K - 1$, we select the pursuer's control from the inclusions

$$\begin{aligned} &\int_{\tau_i}^{\tau_{i+1}} \pi e^{A(T-\vartheta)} u(\vartheta) d\vartheta - \pi e^{A(T-\tau_i)} v_i - w_i(T) \\ &\quad \in \alpha_i(T, x_0, v_i, \omega)(M - \xi(T, x_0, \omega)), \end{aligned} \tag{25}$$

and on the interval $[\tau_K, T]$ from the inclusion

$$\begin{aligned} &\int_{\tau_K}^{T} \pi e^{A(T-\vartheta)} u(\vartheta) d\vartheta - \pi e^{A(T-\tau_K)} v_K - w_K(T) \\ &\quad \in \alpha_K(T, x_0, v_K, \omega)(M - \xi(T, x_0, \omega)). \end{aligned} \tag{26}$$

From Condition 4.1, in view of formula (23), it follows that these inclusions have measurable solutions.

In the case $K < N$, on intervals $[\tau_i, \tau_{i+1})$, $i = 0, 1, \dots, K$, we select the pursuer's control from inclusion (25), and on the intervals $[\tau_i, \tau_{i+1})$, $i = K + 1, \dots, N - 1$, as solutions of the following Fredholm integral equations of the first kind:

$$\int_{\tau_i}^{\tau_{i+1}} \pi e^{A(T-\vartheta)} u(\vartheta) d\vartheta - \pi e^{A(T-\tau_i)} v_i = w_i(T). \tag{27}$$

On the final interval $[\tau_N, T]$ as the pursuer's control we take a solution of the following Volterra integral equation of the first kind:

$$\int_{\tau_N}^{T} \pi e^{A(T-\vartheta)} u(\vartheta) d\vartheta - \pi e^{A(T-\tau_N)} v_N = w_N(T). \tag{28}$$

By virtue of Condition 4.1 equations (27), (28) have measurable solutions.

The Cauchy formula for the system (1) in view of properties of the Dirac delta function yields the following expression:

$$\begin{aligned}\pi x(T) &= \pi e^{A(T-\tau_0)} x_0 \\ &\quad + \sum_{i=0}^{N-1} \left(\int_{\tau_i}^{\tau_{i+1}} \pi e^{A(T-\vartheta)} u(\vartheta) d\vartheta - \pi e^{A(T-\tau_i)} v_i \right) \\ &\quad + \int_{\tau_N}^{T} \pi e^{A(T-\vartheta)} u(\vartheta) d\vartheta - \pi e^{A(T-\tau_N)} v_N. \end{aligned} \tag{29}$$

Let us add and subtract the value $\sum_{i=0}^{N} w_i(T)$ from the right-hand side of (29). Then, taking into account convexity of the set M and expressions (25)–(28) for the pursuer's control, we obtain the following inclusion:

$$\pi x(T) \in \xi(T, x_0, \omega) + \sum_{i=0}^{K} \alpha_i(T, x_0, v_i, \omega)(M - \xi(T, x_0, \omega)) = M. \tag{30}$$

This is equal to the inclusion $x(T) \in M^*$.

Now consider the case $\xi(T, x_0, \omega) \in M$. On the intervals $[\tau_i, \tau_{i+1})$, $i = 0, 1, \dots, N - 1$, we take as the pursuer's control solutions of the equations (27), and on the interval $[\tau_N, T]$ we take solutions of the equation (28). As a result, taking into account (29), we have that $\pi x(T) = \xi(T, x_0, \omega) \in M$ or $x(T) \in M^*$. □

5 Impulse Control of Both Players

Let $\{\tau_i\}$, $\{\eta_j\}$ be sequences of time instants numbered in ascending order, which satisfy Condition 3.1.

Let us consider a linear controlled dynamic system, the evolution of which is modeled by the equation (1). Here we assume that controls, employed by both the pursuer and the evader, are of impulse nature and have the form

$$u(t) = \sum_{i=0}^{\infty} u_i \delta(t - \tau_i), \quad v(t) = \sum_{j=0}^{\infty} v_j \delta(t - \eta_j). \tag{31}$$

Here the pursuer's jump control vectors u_i are taken from the compact $U, U \subset \mathbb{R}^m$, and the evader's jump control vectors v_j from the compact V, $V \subset \mathbb{R}^m$.

In view of results from [14] for any initial state (4) and given controls (31) there exists a solution to system (1), which is unique and absolutely continuous on intervals between the instants of jumps $\{\tau_i\}$, $\{\eta_j\}$, i, $j \in \mathbb{N} \cup \{0\}$.

As in previous sections we assume that the pursuer uses quasistrategies, i.e., $u_i = u(\tau_i, x_0, \mathfrak{v}_{\tau_i})$, where $u(\tau_i, x_0, \mathfrak{v}_{\tau_i}(\cdot))$ is a function of the current instant τ_i, the initial state x_0, and the pre-history of the evader's control $\mathfrak{v}_{\tau_i} = \{v_j : \eta_j \leq \tau_i, v_j \in V\}$, taking values in the set U.

Denote $J_0 = \varnothing$, $J_i = \{j \in \mathbb{N} \cup \{0\} : \eta_j \in [\tau_{i-1}, \tau_i)\}$, $i \in \mathbb{N}$, and consider the sets

$$W_i(n, \{v_j\}) = \pi e^{A(\tau_n - \tau_i)} U - \sum_{j \in J_i} \pi e^{A(\tau_n - \eta_j)} v_j, \tag{32}$$

$$W_i(n) = \pi e^{A(\tau_n - \tau_i)} U \overset{*}{-} \sum_{j \in J_i} \pi e^{A(\tau_n - \eta_j)} V, \tag{33}$$

where $i = 0, \ldots, n$.

Condition 5.1. The sets $W_i(n)$ are nonempty for all n, i, $n \in \mathbb{N} \cup \{0\}$, $i = 0, \ldots, n$.

By virtue of Condition 5.1, one can select some element $w_i(n)$ from set $W_i(n)$, $i = 0, \ldots, n$. Let us fix some collection $\omega = \omega(n) = \{w_i(n)\}_{i=0}^{n}$ and set

$$\xi(n, x, \omega) = \pi e^{A(\tau_n - \tau_0)} x + \sum_{i=0}^{n} w_i(n).$$

We introduce the functions

$$\begin{aligned} &\tilde{\alpha}_i(n, x, \{v_j\}, \omega) \\ &\quad = \sup\{\alpha \geq 0 : \alpha(M - \xi(n, x, \omega)) \cap (W_i(n, \{v_j\}) - w_i(n)) \neq \varnothing\}. \end{aligned} \tag{34}$$

Let us set

$$\begin{aligned} k &= k(n, x, \{v_j\}, \omega) \\ &= \min \left\{ j \in \{0, \ldots, n\} : \sum_{i=0}^{j} \tilde{\alpha}_i(n, x, \{v_j\}, \omega) \geq 1 \right\}. \end{aligned} \tag{35}$$

If the inequality in braces fails for all j, $j \in \{0, \dots, n\}$, then we set $k = n + 1$. Let us define the resolving function [8]. If $\xi(n, x, \omega) \in M$, we set

$$\alpha_i(n, x, \{v_j\}, \omega) = \frac{1}{n+1}, \quad i = 0, \dots, n.$$

If, otherwise, $\xi(n, x, \omega) \notin M$, then we set

$$\alpha_i(n, x, \{v_j\}, \omega) = \begin{cases} \tilde{\alpha}_i(n, x, \{v_j\}, \omega), & i = 0, \dots, k-1, \\ 1 - \sum_{j=0}^{k-1} \tilde{\alpha}_j(n, x, \{v_j\}, \omega), & i = k, \\ 0, & i = k+1, \dots, n, \end{cases} \tag{36}$$

where $0 < k < n$.

We will analyze separately the cases $k = 0$, $k = n$, and $k = n+1$. If $\xi(n, x, \omega) \notin M$ and $k = 0$, then

$$\alpha_i(n, x, \{v_j\}, \omega) = \begin{cases} 1, & i = 0, \\ 0, & i = 1, \dots, n. \end{cases} \tag{36$'$}$$

If $\xi(n, x, \omega) \notin M$ and $k = n$, then

$$\alpha_i(n, x, \{v_j\}, \omega) = \begin{cases} \tilde{\alpha}_i(n, x, \{v_j\}, \omega), & i = 0, \dots, n-1 \\ 1 - \sum_{j=0}^{n-1} \tilde{\alpha}_j(n, x, \{v_j\}, \omega), & i = n. \end{cases} \tag{36$''$}$$

In the case when $\xi(n, x, \omega) \notin M$ and $k = n + 1$, we set

$$\alpha_i(n, x, \{v_j\}, \omega) = \tilde{\alpha}_i(n, x, \{v_j\}, \omega), \quad i = 0, \dots, n.$$

One can easily show that if sets M and U are convex, then the relation

$$\alpha_i(n, x, \{v_j\}, \omega)(M - \xi(n, x, \omega)) \cap (W_i(n, \{v_j\}) - w_i(n)) \neq \varnothing \tag{37}$$

is true for the above-defined functions. Here $n \in \mathbb{N} \cup \{0\}$, $i = 0, \dots, n$, $x \in \mathbb{R}^m$, $v_j \in V$, $w_i(n) \in W_i(n)$. The reasonings are similar to those used in the proof of Lemma 3.1.

Let us introduce the function

$$N(x, \omega) = \min \left\{ n \in \mathbb{N} \cup \{0\} : \sum_{i=0}^{n} \inf_{\{v_j\}} \alpha_i(n, x, \{v_j\}, \omega) = 1 \right\}. \tag{38}$$

If the equality in braces fails for all n, we set $N(x, \omega) = +\infty$.

Theorem 5.1. *In the game (1), (2) with impulse controls of the players (31), let the sets M and U be convex, Condition 5.1 hold, and let $N(x_0, \omega) < +\infty$ for given initial state x_0 and some collection ω, and $\eta_j \neq \tau_{N(x_0,\omega)}$, $j \in \mathbb{N} \cup \{0\}$. Then a trajectory of the system (1) starting at the point x_0 can be driven to the terminal set (2) at the time instant $\tau_{N(x_0,\omega)}$.*

Proof. Let us denote $N = N(x_0, \omega)$ and fix some collection of jump vectors $\{v_j\}_{j:\,\eta_j \in [\tau_0, \tau_N]}$ from the set V.

Let us consider the case when $\xi(N, x_0, \omega) \notin M$. We set $K = k(N, x_0, \{v_j\}, \omega)$. In view of formulas (35), (36) we have

$$\sum_{i=0}^{K} \alpha_i(N, x_0, \{v_j\}, \omega) = 1.$$

For $i = 0, \ldots, K$ we select jump vectors u_i from the inclusions

$$\begin{aligned} &\pi e^{A(\tau_N - \tau_i)} u_i - \sum_{j \in J_i} \pi e^{A(\tau_N - \eta_j)} v_j - w_i(N) \\ &\quad \in \alpha_i(N, x_0, \{v_j\}, \omega)(M - \xi(N, x_0, \omega)). \end{aligned} \tag{39}$$

The existence of solutions to inclusion (39) follows from Condition 5.1 and formula (37).

For $i = K+1, \ldots, N$ as jump vectors u_i we take solutions of the equations

$$\pi e^{A(\tau_N - \tau_i)} u_i - \sum_{j \in J_i} \pi e^{A(\tau_N - \eta_j)} v_j = w_i(N). \tag{40}$$

The existence of these solutions follows from Condition 5.1.

The Cauchy formula for system (1) in view of properties of the Dirac delta function, and the assumption $\eta_j \neq \tau_N$, $j \in \mathbb{N} \cup \{0\}$, yield the following formula:

$$\pi x(\tau_N) = \pi e^{A(\tau_N - \tau_0)} x_0 + \sum_{i=0}^{N} \left(\pi e^{A(\tau_N - \tau_i)} u_i - \sum_{j \in J_i} \pi e^{A(\tau_N - \eta_j)} v_j \right). \tag{41}$$

Let us add and subtract from the right-hand side of (41) the value $\sum_{i=0}^{N} w_i(N)$. Then, taking into account convexity of the set M and expressions (39), (40) for the pursuer's control, we obtain the inclusion

$$\pi x(\tau_N) \in \xi(N, x_0, \omega) + \sum_{i=0}^{K} \alpha_i(N, x_0, \{v_j\}, \omega)(M - \xi(N, x_0, \omega)) = M. \tag{42}$$

This is equivalent to the inclusion $x(\tau_N) \in M^*$.

Now suppose that $\xi(N, x_0, \omega) \in M$. As jump vectors u_i, $i = 0, \ldots, N$, we take solutions to (40). Then from formula (41) we immediately obtain the inclusion $\pi x(\tau_N) = \xi(N, x_0, \omega) \in M$, which is equivalent to the inclusion $x(\tau_N) \in M^*$.

□

6 Example

As an illustration consider the differential game of pursuit with simple motion dynamics. Let the dynamics of the system be governed by the following equation:

$$\dot{x} = u - v, \quad x \in \mathbb{R}^m. \tag{43}$$

Assume that the terminal set consists of the single point: $M^* = \{0\}$. Then $M^0 = \{0\}$, $M = \{0\}$. Then $L = \mathbb{R}^m$, π is the identity: $\pi = E$, and A is a null matrix, $e^{At} = E$.

At first, we consider the case when the pursuer employs impulse controls, and the evader employs measurable controls. Let us assume that $\tau_i = iP$, where P is some fixed period of time. Then the pursuer's control has the form

$$u(t) = \sum_{i=0}^{\infty} u_i \delta(t - iP), \ u_i \in U.$$

We also assume that U is a closed ball of radius ρ centered at zero, and V is a ball of radius σ, i.e., $U = \rho S$, $V = \sigma S$, $S = \{x \in \mathbb{R}^m : \ \|x\| \le 1\}$.

Then

$$W_0(n, v(\cdot)) = W_0(n) = W_0 = \rho S,$$

$$W_i(n, v(\cdot)) = W_i(v(\cdot)) = \rho S - \int_{(i-1)P}^{iP} v(\vartheta)d\vartheta,$$

$$W_i(n) = W_i = \rho S \overset{*}{-} \int_{(i-1)P}^{iP} \sigma S d\vartheta = \rho S \overset{*}{-} P\sigma S, \ i = 1, \dots, n.$$

One can see that Condition 3.2 holds if $\rho \ge P\sigma$. In this case $W_0 = \rho S$, $W_i = (\rho - P\sigma)S, i = 1, \dots, n$, and for all i sets W_i contain the null vector. Set $w_i(n) \equiv 0$ for all i, n. Then

$$\xi(n, x, \omega) = x.$$

Denote $I_0 = 0$, $I_i = \int_{(i-1)P}^{iP} v(\vartheta)d\vartheta$, and consider the functions

$$\tilde{\alpha}_i(n, x, v(\cdot), \omega) = \max\{\alpha \ge 0 : \ -\alpha x \in \rho S - I_i\}, \ i = 0, \dots, n.$$

These functions can be found explicitly from the equation

$$\|I_i - \alpha x\| = \rho.$$

Solving it we obtain

$$\tilde{\alpha}_i(n, x, v(\cdot), \omega) = \frac{(x, I_i) + \sqrt{(x, I_i)^2 + \|x\|^2(\rho^2 - \|I_i\|^2)}}{\|x\|^2}.$$

The minimum value of $\tilde{\alpha}_i(n, x, v(\cdot), \omega)$ is attained at $I_i = -P\sigma \frac{x}{\|x\|}$, i.e., when $v(t) = -\sigma \frac{x}{\|x\|}$ almost everywhere on the interval $((i-1)P, iP)$, $i = 1, \dots, n$.

Since

$$\inf_{v(\cdot)} \tilde{\alpha}_0(n, x, v(\cdot), \omega) = \frac{\rho}{\|x\|},$$

$$\inf_{v(\cdot)} \tilde{\alpha}_i(n, x, v(\cdot), \omega) = \frac{\rho - P\sigma}{\|x\|}, \; i = 1, \ldots, n,$$

then for $\|x\| \leq \rho$

$$N(x, \omega) = 0,$$

and for $\|x\| > \rho$

$$N(x, \omega) = N(x) = \begin{cases} \frac{\|x\|-\rho}{\rho-P\sigma}, & \frac{\|x\|-\rho}{\rho-P\sigma} \in \mathbb{N}, \\ \left[\frac{\|x\|-\rho}{\rho-P\sigma}\right] + 1, & \frac{\|x\|-\rho}{\rho-P\sigma} \notin \mathbb{N}. \end{cases}$$

Hereafter $[x]$ denotes the integer part of a number x.

Hence a trajectory of the system (43) starting at the initial state x_0 can be driven to the terminal set not later than at the instant $\tau_{N(x_0)} = N(x_0)P$.

It should be noted that for the above-outlined method of pursuit we have

$$x(\tau_i) = x_0 \left(1 - \sum_{j=0}^{i} \alpha_j(x_0, v(\cdot))\right).$$

We see that at the instants $\tau_i = iP$ the vector $x(\tau_i)$ is parallel to the vector x_0; that is, at these instants of time the strategy of the pursuer is similar to the rule of parallel pursuit [18].

Now consider this game under the assumption that the evader employs impulse controls. Let us assume that the pursuer's control $u = u(t)$ is a measurable function with values in the compact U, and that the evader affects the system only at the discrete instants $\tau_i = iP$. Then his control has the form

$$v(t) = \sum_{i=0}^{\infty} v_i \delta(t - iP), \; v_i \in V.$$

Here also $U = \rho S$, $V = \sigma S$.

In this case $n(t) = [\frac{t}{P}]$. According to (16), (17), we have

$$W_i(t, v) = W_i(v) = \int_{iP}^{(i+1)P} \rho S d\vartheta - v = P\rho S - v, \quad i = 0, \ldots, n(t) - 1,$$

$$W_{n(t)}(t, v) = \int_{[\frac{t}{P}]P}^{t} \rho S d\vartheta - v = \left(t - \left[\frac{t}{P}\right] P\right) \rho S - v.$$

$$W_i(t) = W_i = P\rho S \overset{*}{-} \sigma S, \quad i = 0, \ldots, n(t) - 1.$$

Obviously, if $P\rho \geq \sigma$, these sets are nonempty for all t, $t \geq 0$, and $W_i = (P\rho - \sigma)S$.

$$W_{n(t)}(t) = \left(t - \left[\frac{t}{P}\right] P\right) \rho S \overset{*}{-} \sigma S.$$

The set $W_{n(t)}(t)$ is nonempty for

$$t \geq \left[\frac{t}{P}\right] P + \frac{\sigma}{\rho},$$

and since $t < ([\frac{t}{P}]+1)P$, Condition 4.1 is met, if the following inequality is true:

$$P\rho > \sigma. \tag{44}$$

Then

$$\mathfrak{T} = \bigcup_{i=0}^{\infty} \left[iP + \frac{\sigma}{\rho}, (i+1)P\right).$$

Denote $\eta(t) = t - [\frac{t}{P}]P$. Then $W_{n(t)}(t) = (\eta(t)\rho - \sigma)S$ for $t \in \mathfrak{T}$. Therefore, the sets $W_i(t)$ contain the null vector for all $t \in \mathfrak{T}$, $i = 0, \ldots, n(t)$, and we can set $w_i(t) \equiv 0$. As a result we obtain

$$\xi(t, x, \omega) = x.$$

Let us find the resolving functions. For $i = 0, \ldots, n(t) - 1$

$$\tilde{\alpha}_i(t, x, v, \omega) = \sup\{\alpha \geq 0 : \ -\alpha x \in P\rho S - v\}.$$

The values of these functions can be found explicitly from the equation

$$\|v - \alpha x\| = P\rho,$$

namely,

$$\tilde{\alpha}_i(t, x, v, \omega) = \frac{(x, v) + \sqrt{(x, v)^2 + \|x\|^2(P^2\rho^2 - \|v\|^2)}}{\|x\|^2}.$$

Then

$$\inf_{\|v\| \leq \sigma} \tilde{\alpha}_i(t, x, v, \omega) = \frac{P\rho - \sigma}{\|x\|},$$

where the minimum is attained at $v = -\sigma \frac{x}{\|x\|}$. Consequently

$$N = n(T) = \left[\frac{\|x\|}{P\rho - \sigma}\right],$$

where T, $T \in T(x, \omega)$, is the instant of game termination.

In the case $i = n(t)$ the resolving function has the form

$$\tilde{\alpha}_{n(t)}(t, x, v, \omega) = \sup\{\alpha \geq 0 : \ -\alpha x \in \eta(t)\rho S - v\}.$$

In a similar way we obtain

$$\tilde{\alpha}_{n(t)}(t, x, v, \omega) = \frac{(x, v) + \sqrt{(x, v)^2 + \|x\|^2(\eta(t)^2\rho^2 - \|v\|^2)}}{\|x\|^2},$$

whence we have

$$\inf_{\|v\| \leq \sigma} \tilde{\alpha}_{n(t)}(t, x, v, \omega) = \frac{\eta(t)\rho - \sigma}{\|x\|}.$$

The minimum is attained at $v = -\sigma \frac{x}{\|x\|}$. The time of game termination T can be found from the equation

$$\left[\frac{\|x\|}{P\rho - \sigma}\right] \frac{P\rho - \sigma}{\|x\|} + \frac{\eta(t)\rho - \sigma}{\|x\|} = 1.$$

Under the condition (44) this equation always has a solution.

In conclusion, we study the game (43) in the case where both players employ impulse controls. Assume that $\tau_i = iP$, $\eta_j = jQ$, where P, Q are some fixed periods of time and $Q < P$. The controls of the pursuer and of the evader are of the form

$$u(t) = \sum_{i=0}^{\infty} u_i \delta(t - iP), \quad v(t) = \sum_{j=0}^{\infty} v_j \delta(t - jQ),$$

respectively. Here $u_i \in U = \rho S$, $v_j \in V = \sigma S$.

In this case $J_0 = \varnothing$, $J_i = \{j \in \mathbb{N} : \ jQ \in [(i-1)P, iP)\}$, $i \in \mathbb{N}$, i.e., $|J_i| = [\frac{P}{Q}]$, $i \in \mathbb{N}$, where $|J_i|$ denotes the power of the set J_i.

By definitions of the sets (32), (33), we have

$$W_i(n, \{v_j\}) = \rho S - \sum_{j \in J_i} v_j,$$

$$W_i(n) = W_i = \rho S \overset{*}{-} \sum_{j \in J_i} \sigma S = \rho S \overset{*}{-} |J_i| \sigma S = \rho S \overset{*}{-} \left[\frac{P}{Q}\right] \sigma S.$$

Hence, Condition 5.1 holds when $\rho \geq [\frac{P}{Q}]\sigma S$. Then $W_i(n) = (\rho - [\frac{P}{Q}]\sigma)S$, $i = 1, \ldots, n$, and the sets W_i contain the null vector for all i. Set $w_i(n) \equiv 0$ for all i, n. Then

$$\xi(n, x, \omega) = x.$$

Let us denote $\mathfrak{S}_0 = 0$, $\mathfrak{S}_i = \sum_{j \in J_i} v_j$, $i \in \mathbb{N}$, and find the resolving functions,

$$\tilde{\alpha}_i(n, x, \{v_j\}, \omega) = \sup\{\alpha \geq 0 : \ -\alpha x \in \rho S - \mathfrak{S}_i\}. \tag{45}$$

The functions (45) can be found explicitly from the equations

$$\|\mathfrak{S}_i - \alpha x\| = \rho.$$

Solving them we obtain

$$\tilde{\alpha}_i(n, x, \{v_j\}, \omega) = \frac{(x, \mathfrak{S}_i) + \sqrt{(x, \mathfrak{S}_i)^2 + \|x\|^2(\rho^2 - \|\mathfrak{S}_i\|^2)}}{\|x\|^2}.$$

It is evident that

$$\begin{aligned} \inf_{\{v_j\}} \tilde{\alpha}_0(n, x, \{v_j\}, \omega) &= \frac{\rho}{\|x\|}, \\ \inf_{\{v_j\}} \tilde{\alpha}_i(n, x, \{v_j\}, \omega) &= \frac{\rho - [\frac{P}{Q}]\sigma}{\|x\|}, \end{aligned} \tag{46}$$

and the minimum is attained at $\mathfrak{S}_i = -[\frac{P}{Q}]\sigma \frac{x}{\|x\|}$, that is, when $v_j = -\sigma \frac{x}{\|x\|}$ for all j. It follows from (38), (46), that for $\|x\| \leq \rho$

$$N(x, \omega) = 0,$$

and for $\|x\| > \rho$

$$N(x, \omega) = N(x) = \begin{cases} \frac{\|x\|-\rho}{\rho - [\frac{P}{Q}]\sigma}, & \frac{\|x\|-\rho}{\rho - [\frac{P}{Q}]\sigma} \in \mathbb{N}, \\ \left[\frac{\|x\|-\rho}{\rho - [\frac{P}{Q}]\sigma}\right] + 1, & \frac{\|x\|-\rho}{\rho - [\frac{P}{Q}]\sigma} \notin \mathbb{N}, \end{cases}$$

and the game terminates at the instant $\tau_{N(x_0)}$, provided that $\eta_{N(x_0)} \neq \tau_{N(x_0)}$.

References

[1] Pozharitskii G.K., Game Problem of Impulse Soft Encounter of Two Material Points, *Prikl. Mat. Mech.*, vol. 36, no. 2, 1972 (in Russian).

[2] Pozharitskii G.K., Impulse Pursuit of a Point with Limited Thrust, *Prikl. Mat. Mech.*, vol. 37, no. 2, 1973 (in Russian).

[3] Zavalishchin S.T., Sesekin A.N. *Dynamic Impulse Systems: Theory and Applications*, Kluwer Academic Publishers, Boston–London–Dordrecht, 1997.

[4] Ostapenko E.V., Operator Constructions in Games for Discontinuous Dynamic Systems, *Optimal Decision Theory*, pp. 14–18, 1999 (in Russian).

[5] Ostapenko E.V., Game Problems for Systems with Discontinuous Dynamics, *Cybernetics and Systems Analysis*, no. 4, pp. 183–187, 2000 (in Russian).

[6] Aubin J.-P., Lygeros J., Quincampoix M., Sastry S., Seube N., Impulse Differential Inclusions: A Viability Approach to Hybrid Systems, *IEEE Transactions on Automatic Control*, no. 47, pp. 2–20, 2002.

[7] Aubin J.-P., Seube N., Conditional Viability for Impulse Differential Games, *Proc. of the 41st IEEE Conference on Decision and Control*, vol. 3, pp. 3200–3204, 2002.

[8] Chikrii A.A., *Conflict Controlled Processes*, Kluwer Academic Publishers, Boston–London–Dordrecht, 1997.

[9] Chikrii A.A., Matychyn I.I., On Some Class of Game Problems with Impulse Control, *Dopovidi NAS of Ukraine*, no. 6, pp. 73–77, 2004 (in Russian).

[10] Chikrii A.A., Matychyn I.I., Linear Differential Games with Impulse Control of Evador, *Dopovidi NAS of Ukraine*, no. 10, pp. 80–85, 2004 (in Russian).

[11] Chikrii A.A., Matychyn I.I., Chikrii K.A., Conflict Controlled Processes with Discontinuous Trajectories, *Cybernetics and Systems Analysis*, vol. 40, no. 6, pp. 800–811, 2005.

[12] Krivonos Yu.G., Matychyn I.I., Chikrii A.A., *Dynamic Games with Discontinuous Trajectories*, Naukova Dumka, Kiev, 2005 (in Russian).

[13] Bracewell R., The Impulse Symbol, Ch. 5 in *The Fourier Transform and Its Applications*, 3rd ed., McGraw-Hill, New York, 1999.

[14] Filippov A.F., *Differential Equations with Discontinuous Right-Hand Side*, Nauka, Moscow, 1985 (in Russian).

[15] Joffe A.D., Tikhomirov V.M., *Theory of Extremal Problems*, Nauka, Moscow, 1974 (in Russian).

[16] Pontrjagin L.S., *Selected Scientific Works*, Nauka, Moscow, vol. 2, 1988 (in Russian).

[17] Aumann R.J., Integrals of Set-Valued Functions, *J. Math. Anal. Appl.*, no. 12, pp. 1–12, 1965.

[18] Chikrii A.A., Eidelman S.D., Generalized Matrix Functions of Mittag-Leffler in Game Problems for Evolutionary Equations of Fractional Order, *Cybernetics and Systems Analysis*, no. 3, pp. 3–32, 2000 (in Russian).

On the Instability of the Feedback Equilibrium Payoff in a Nonzero-Sum Differential Game on the Line

Pierre Cardaliaguet
Laboratoire de Mathématiques
Unite' CNRS UMR6205
Université de Bretagne Occidentale
6 avenue Victor le Gorgeu
29285 Brest cedex, France
Pierre.Cardaliaguet@univ-brest.fr

Abstract

For a simple nonzero-sum differential game on the real line, the natural notion of Nash equilibrium payoff in feedback form turns out to be extremely unstable. Two examples of different types of instability are discussed.

1 Introduction

The aim of this chapter is the study of the stability of Nash equilibria for some nonzero-sum two-player differential games on the real line. The dynamics of the differential game is the simplest one can imagine:

$$x' = u + v, \text{ with } u \in [-1, 1], \; v = [-1, 1] \; . \tag{1}$$

The payoff is the terminal payoff $g(x(T)) = (g_1(x(T)), g_2(x(T)))$, for some function $g : \mathcal{R} \to \mathcal{R}^2$. Player I, playing with u, wants to maximize $g_1(x(T))$, while Player II, playing with v, wants to maximize $g_2(x(T))$.

In [4] we have defined a notion of **feedback equilibrium payoff (FEP)** for this game, and we have proved that our game has a unique FEP provided (g_1, g_2) are regular in some sense (the exact statement is given in Section 2). Our aim here is to investigate the dependence of this FEP with respect to the terminal payoffs (g_1, g_2).

For *zero-sum* differential games, i.e., when $g_2 = -g_1$, the situation is very well understood: the game has a value (which is our FEP), and this value depends in a continuous way on g_1 for the (locally) uniform convergence (see for instance [5]). Actually, for our elementary example, the value function V is simply given by $V(t, x) = g_1(x)$ for any $(t, x) \in [0, T] \times \mathcal{R}$.

Very surprisingly, the situation for our nonzero-sum game is completely different. The FEP is extremely unstable: we give two examples showing that small uniform perturbations of (g_1, g_2) generally lead to a very strong change in behaviour of the FEP.

Results in the same spirit have been obtained recently by Bressan and Shen [2], [3] for another class of nonzero-sum differential games. Using the characterization of the FEP as solutions of some Hamilton–Jacobi system of equations, the authors show that, in general, this system of equations is ill posed, whence the expected instability (and even nonexistence) of the FEP.

In order to explain why such a situation arises, we have to recall the notion of "Nash payoffs." Nash payoffs are equilibrium payoffs that can be realized when the game is played with memory strategies. The basic idea is that memory strategies incorporate a threat which will be used if the opponent does not observe the agreement, memory allowing each player to recall a possible deviation from the agreement. In [7], [9] (see also [6]), the Nash payoffs for such strategies are completely characterized and, under the well-known Isaacs conditions, are proved to exist for any initial position (and much more general dynamics than ours).

In general there are infinitely many such Nash payoffs for a given initial position. For instance, in our game and for continuous functions g_1 and g_2, a pair $(e_1, e_2) \in \mathcal{R}^2$ is a Nash payoff for the initial position (t_0, x_0) if and only if

i) there is some $y_0 \in [x_0 - 2(T - t_0), x_0 + 2(T - t_0)]$ such that $e_1 = g_1(y_0)$ and $e_2 = g_2(y_0)$,
ii) for any y between x_0 and y_0, we have $e_1 \geq g_1(y)$ and $e_2 \geq g_2(y)$.

The first condition states that the payoff (e_1, e_2) can be realized by the players by using suitable admissible controls, while the second one gives to each player the possibility to punish his opponent in case of deviation.

In [4] (see also Section 2) we show that, for the simple game described above, there is a natural way to select a "good" Nash equilibrium payoff so that it enjoys a dynamic programming property. This is what we call a feedback equilibrium payoff (FEP), because it can be realized by using feedback strategies.

We now go back to our stability problem. When trying to understand the behaviour, under perturbation of the terminal payoff, of the FEP, it is of course crucial to investigate the behaviour of the set of Nash equilibrium payoffs: indeed the FEP is just a selection of these Nash payoffs. It turns out that Nash equilibrium payoffs depend in a very unstable way on the functions g_1 and g_2.

This remark allows us to build two examples of unstable behaviour. The first one shows that instability can come from an oscillatory behaviour of the perturbed terminal payoff. The second example is more subtle: it shows that instability can also appear because of some phenomena of a global nature. More precisely, we build smooth functions (g_1^n, g_2^n) which converge to some smooth (g_1, g_2) "as smoothly as possible," such that g_1 and all the g_1^n (resp. g_2 and all the g_2^n) are increasing and decreasing exactly on the same intervals, and such that the corresponding FEP for (g_1^n, g_2^n) does not converge to the corresponding one for (g_1, g_2). What happens in both examples is that some Nash payoffs are lost by the perturbation.

To complete this introduction, let us briefly indicate some directions of future research: since the instability of the FEP mainly comes from the instability of the

Nash payoffs, it is natural to investigate, first, if there are conditions under which Nash payoffs converge, and, second, if such a convergence leads to the convergence of the FEP. In a work in preparation, we intend to give a positive answer to these questions. The next step—which is very important but yet completely open—is to understand if our continuous game can be approximated by discrete ones, or by stochastic differential games, for which the existence of equilibrium feedbacks is now known (see [8]).

2 Nash Equilibrium Feedbacks

In this section we fix the notation and recall the main definitions and results of [4].

2.1 Feedbacks and Payoffs

The dynamics of the differential game is (1):

$$x'(t) = u(t) + v(t) \qquad u(t) \in [-1, 1] \text{ and } v(t) \in [-1, 1]. \tag{2}$$

The game is of fixed duration and the terminal time is denoted by T. The terminal payoff is a function $g = (g_1, g_2) : \mathcal{R} \to \mathcal{R}^2$: Player I wants to maximize $g_1(x(T))$ while Player II wants to maximize $g_2(x(T))$.

We denote by $\mathcal{U}$ the set of functions $u : [0, T] \times \mathcal{R} \to [-1, 1]$, interpreted as **strategies** for Player I, and by $\mathcal{V}$ the set of functions $v : [0, T] \times \mathcal{R} \to [-1, 1]$, interpreted as **strategies** for Player II. We call a pair $(u, v) \in \mathcal{U} \times \mathcal{V}$ a **feedback**.

Let $(u, v) \in \mathcal{U} \times \mathcal{V}$ be fixed. A solution $x(\cdot)$ to the equations

$$\begin{cases} x'(t) = u(t, x(t)) + v(t, x(t)) & \text{on } [t_0, T] \\ x(t_0) = x_0 \end{cases}$$

is by definition a solution to the differential inclusion

$$\begin{cases} x' \in \tilde{f}(t, x, u, v) & \text{on } [t_0, T] \\ x(t_0) = x_0 \end{cases}, \tag{3}$$

where $\tilde{f}(t, x, u, v)$ is the smallest upper semi-continuous (usc) convex and compact set-valued map containing the map $(t, x) \to u(t, x) + v(t, x)$. It is well known that this differential inclusion has at least one solution (see [1]). We denote by $\mathcal{X}(t_0, x_0, u, v)$ the set of all its solutions.

The **lower and upper payoffs** of the strategies (u, v) for the initial position (t_0, x_0) are given as follows: For $j = 1, 2$, the lower payoff of Player j, denoted by $J_j^\flat(t_0, x_0, u, v)$, is

$$J_j^\flat(t_0, x_0, u, v) = \inf_{x \in \mathcal{X}(t_0, x_0, u, v)} (g_j)_*(x(T)),$$

while the upper payoff $J_j^\sharp(t_0, x_0, u, v)$ is

$$J_j^\sharp(t_0, x_0, u, v) = \sup_{x \in \mathcal{X}(t_0, x_0, u, v)} (g_j)^*(x(T)).$$

Here $(g_j)_*$ and $(g_j)^*$ are respectively the lower semi-continuous and the upper semi-continuous envelopes of g_j (i.e., respectively the largest lower semi-continuous (lsc) function which is not larger than g_j and the the smallest upper semi-continuous (usc) function which is not smaller than g_j). We also note that the inf and sup in the preceding definitions are actually min and max (see Proposition 2.1 of [4]).

We use below the notation

$$J^\sharp(t_0, x_0, u, v) = (J_1^\sharp(t_0, x_0, u, v), J_2^\sharp(t_0, x_0, u, v))$$

and

$$J^\flat(t_0, x_0, u, v) = (J_1^\flat(t_0, x_0, u, v), J_2^\flat(t_0, x_0, u, v)).$$

2.2 Definition of the Feedback Equilibrium Payoff

Definition 2.1. A Nash equilibrium feedback on the time interval $[T_0, T]$ for terminal time T and terminal payoff g is a feedback $(u^*, v^*) \in \mathcal{U} \times \mathcal{V}$ such that for all $t_0 \in [T_0, T]$ there exists a set $S_{t_0} \subset \mathcal{R}$ of zero measure such that for any $x_0 \in \mathcal{R} \backslash S_{t_0}$, we have

$$\forall u \in \mathcal{U}, \ \ J_1^\flat(t_0, x_0, u^*, v^*) \geq J_1^\sharp(t_0, x_0, u, v^*)$$

and

$$\forall v \in \mathcal{V}, \ \ J_2^\flat(t_0, x_0, u^*, v^*) \geq J_2^\sharp(t_0, x_0, u^*, v).$$

Remark: Setting $(u, v) = (u^*, v^*)$ in the previous inequality shows that, for $j = 1, 2$, we have

$$\forall x_0 \notin S_{t_0}, \quad J_j^\flat(t_0, x_0, u^*, v^*) = J_j^\sharp(t_0, x_0, u^*, v^*).$$

In order to simplify the notation, we denote by $J_1(t_0, x_0, u^*, v^*)$ (resp. $J_2(t_0, x_0, u^*, v^*)$) this common value. We say that

$$J(t_0, x_0, u^*, v^*) = (J_1(t_0, x_0, u^*, v^*), J_2(t_0, x_0, u^*, v^*))$$

is the payoff of the Nash equilibrium feedback (u^*, v^*) at the point (t_0, x_0). It is defined for any $t_0 \in [0, T]$ and for any $x_0 \in \mathcal{R} \backslash S_{t_0}$ where S_{t_0} has a zero measure in $\mathcal{R}$.

There are in general infinitely many Nash equilibrium feedbacks, and most of them are completely uninteresting: see the examples in [4], for instance. In order to select "a natural one," we now introduce the notion of maximal payoff. A maximal payoff at some position (t, x) is at the same time an admissible payoff *and* the maximum that each player can hope to get when starting from (t, x).

Let us first introduce the notion of ess lim sup : For a map $g : \mathcal{R} \to \mathcal{R}^2$ and a point $x \in \mathcal{R}$, we set

$$\begin{aligned}
&\operatorname{ess}\limsup_{x' \to x} g(x') \\
&\quad = \left\{a \in \mathcal{R}^2 \mid \forall \epsilon > 0,\ \forall r > 0,\ \left|g^{-1}(B_\epsilon(a)) \cap]x - r, x + r[\right| > 0\right\},
\end{aligned}$$

where $|A|$ denotes the outward Lebesgue measure of a subset A of $\mathcal{R}$ and where $B_\epsilon(a)$ denotes the ball of center a and radius ϵ in $\mathcal{R}^2$. Of course, if g is continous, then $\operatorname{ess}\limsup_{x' \to x} g(x') = g(x)$.

Definition 2.2 (Maximal payoff). Let $(t_0, x_0) \in [0, T] \times \mathcal{R}$ and $(e_1, e_2) \in \mathcal{R}^2$. We say that (e_1, e_2) is a maximal payoff at the point (t_0, x_0) for the game with terminal time T and terminal payoff g if there is some solution $x_0(\cdot)$ to (2) with $x_0(t_0) = x_0$ and such that $(e_1, e_2) \in \operatorname{ess}\limsup_{x' \to x_0(T)} g(x')$ and such that, for any $x(\cdot) \in \mathcal{X}(t_0, x_0)$, we have

$$\text{for } j = 1, 2,\ e_j \geq (g_j)^*(x(T)).$$

Remarks

(1) Loosely speaking, this means that both functions g_1 and g_2 have a maximum at the same point $x_0(T)$ among the points $x(T)$ that one can reach starting from (t_0, x_0).

(2) For a given point (t_0, x_0) such a maximal equilibrium payoff does not necessarily exist (and in fact seldom exists). However, if it exists, it is clearly unique.

(3) If g is continuous, we have of course $(e_1, e_2) = g(x_0(T))$. However, for later use, we need to have this definition at points where the function g can be discontinuous.

Definition 2.3. Let $(u^*, v^*) \in \mathcal{U} \times \mathcal{V}$ be a Nash equilibrium feedback on some time interval $[T_0, T]$ (with $T_0 < T$) for the game with terminal time T and terminal payoff g.

We say that (u^*, v^*) is **maximal** on $[T_0, T]$ for the game with terminal time T and terminal payoff g if (u^*, v^*) is a Nash equilibrium feedback on $[T_0, T]$ and if (u^*, v^*) satisfies the following additional requirement: For any $(t_0, x_0) \in [T_0, T) \times \mathcal{R}$, if there is some maximal payoff (e_1, e_2) for the game with terminal time T and terminal payoff g at the point (t_0, x_0), then

$$\text{for } j = 1, 2,\ e_j = J_j^\sharp(t_0, x_0, u^*, v^*).$$

Moreover, we say that (u^*, v^*) is **completely maximal** on the time interval $[T_0, T]$ if, for any $T_1 \in (T_0, T]$, (u^*, v^*) is a maximal Nash equilibrium feedback for the game with terminal time T_1 and terminal payoff $J^\sharp(T_1, \cdot, u^*, v^*)$.

The payoff associated to a completely maximal Nash equilibrium feedback is called a **feedback equilibrium payoff**: in short FEP.

Remarks

(1) The condition of maximality allows us to rule out the uninteresting Nash equilibrium feedbacks. It is a very weak condition because there are in general few initial conditions at which a maximal payoff exists.
(2) The complete maximality assumption guarantees the time consistency of the solution.

2.3 The Existence and Uniqueness Result for the FEP

In order to state the main result of [4], we have to define the class $\mathcal{G}$ of terminal payoffs for which we can solve the problem. Basically, it is the set of functions which have only a finite number of discontinuities and a finite number of extrema.

Definition 2.4. The class of admissible payoffs $\mathcal{G}$ is the set of maps $g = (g_1, g_2) : \mathcal{R} \to \mathcal{R}^2$ for which there is a partition $\sigma_0 = -\infty < \sigma_1 < \cdots < \sigma_k < \sigma_{k+1} = +\infty$ such that

(1) for any $i = 0, \ldots, k$,
– either g_1 (resp. g_2) is continuous and (strictly) increasing or decreasing on (σ_i, σ_{i+1}),
– or g_1 and g_2 are simultaneously constant on the interval (σ_i, σ_{i+1}),
(2) g_1 and g_2 are usc on $\mathcal{R}$, and, for $i = 1, \ldots, k$ and $j = 1, 2$, g_j is right or left continuous at σ_i.
(3) if g_1 (resp. g_2) has a strict local maximum at the point σ_i, then g_1 and g_2 are continuous at σ_i.

Terminology: We say below that $\Sigma = \{\sigma_0 < \sigma_1 < \cdots < \sigma_k < \sigma_{k+1}\}$ is the partition associated with the map g.

Remark: In particular, if $g = (g_1, g_2)$ is continuous and g_1 and g_2 have a finite number of local extrema, then g belongs to $\mathcal{G}$.

We are now ready to give the main result of [4].

Theorem 2.1. *Assume that the terminal payoff g belongs to the class $\mathcal{G}$. Then there exists a completely maximal Nash equilibrium feedback for the game with terminal time T and terminal payoff g on the time interval* $[0, T]$.

Moreover, the payoff of any two completely maximal Nash equilibrium feedbacks—i.e., two FEPs—coincide almost everywhere.

Finally, if (u^*, v^*) is a completely maximal Nash equilibrium feedback, then the associated FEP $E(\cdot,\cdot) := J(\cdot,\cdot,u^*,v^*)$, belongs to $\mathcal{G}$ at any time t:

$$E(t,\cdot) \in \mathcal{G} \qquad \forall t \in [0, T].$$

In the sequel, we are mainly interested in the behaviour under perturbations of the FEP. We usually denote such an FEP by

$$E(t,x) = (E_1(t,x), E_2(t,x)) := J(t,x,u^*,v^*),$$

where (u^*, v^*) is a completely maximal Nash equilibrium feedback.

2.4 Link with Nash Payoffs for Memory Strategies

Nash payoffs are equilibrium payoffs that can be realized when the game is played with memory strategies. In order to avoid the lengthy introduction of the notion of memory strategies, we skip the definition of such payoffs, and only give their characterization as a definition, as provided in [7] and [9].

Definition 2.5. Let us consider the game with dynamics (1) and with continuous terminal payoff $g = (g_1, g_2)$. Let $(t_0, x_0) \in [0, T) \times \mathcal{R}$ be a fixed initial position. A pair $(e_1, e_2) \in \mathcal{R}^2$ is a **Nash payoff** for this game, played with memory strategies for the point (t_0, x_0), if there is some solution $x(\cdot)$ to (2) with $x(t_0) = x_0$, such that

$$\forall t \in [t_0, T], \text{ for } j = 1, 2, \ g_j(x(t)) \le e_j = g_j(x(T)) \ .$$

We call the solution $x(\cdot)$ a **Nash trajectory.**

Remark 2.1

- The simplest example of a Nash payoff at a point (t_0, x_0) is the payoff $g(x_0)$. To see this, choose $x(t) = x_0$ for $t \in [t_0, T]$ in Definition 2.5. We call this Nash payoff the *trivial one*.
- It is easy to check that a pair (e_1, e_2) is a Nash payoff at (t_0, x_0) if and only if
  - there is some $y_0 \in [x_0 - 2(T - t_0), x_0 + 2(T - t_0)]$ such that $e_1 = g_1(y_0)$ and $e_2 = g_2(y_0)$,
  - for any y between x_0 and y_0, we have $e_1 \ge g_1(y)$ and $e_2 \ge g_2(y)$.

The next proposition states that the FEPs are always Nash payoffs.

Proposition 2.1. *We assume that g is continuous. Let (u^*, v^*) be some Nash equilibrium feedback on the time interval $[0, T]$ and $(S_t)_{t\in[0,T]}$ be its associated set of zero measure. Then, for any $t \in [0, T)$ and $x \notin S_t$, the Nash equilibrium feedback $J(t, x, u^*, v^*)$ is a Nash equilibrium payoff and any solution $x^*(\cdot) \in \mathcal{X}(t, x, u^*, v^*)$ is a Nash trajectory.*

3 First Example of Instability

In a first example, we consider a terminal payoff $g = (g_1, g_2)$ with g_1 and g_2 strictly increasing and continuous. Let us first note that, for any $(t, x) \in [0, T) \times \mathcal{R}$, the payoff $g(x+2(T-t))$ is a maximal payoff because g_1 and g_2 are strictly increasing. Hence the FEP E for this game is given by

$$E(t,x) = g(x+2(T-t)) \qquad \forall (t,x) \in [0,T] \times \mathcal{R}.$$

Proposition 3.1. *Let $g \in \mathcal{G}$ be as above. There is a sequence (g^n) of continuous functions which belong to $\mathcal{G}$, which locally uniformly converges to g, and such that at each point $(t,x) \in [0,T) \times \mathcal{R}$ and for n sufficiently large, the unique Nash payoff is $g^n(x)$. Then, the FEP E^n associated to g^n satisfies*

$$\forall (t,x) \in [0,T] \times \mathcal{R},\ \exists n_0,\ \forall n \geq n_0, \qquad E^n(t,x) = g^n(x)\ .$$

In particular, E^n does not converge to the FEP E associated to g.

Proof. For any $n \geq 1$, define g_1^n and g_2^n as the unique continuous piecewise affine functions such that, for any $k \in \mathcal{Z}$, $|k| \leq 2n-1$,

$$g_1^n\left(\frac{2k}{n}\right) = g_1\left(\frac{2k}{n}\right) \text{ and } g_1^n\left(\frac{2k+1}{n}\right) = g_1\left(\frac{2k}{n}\right) - \frac{1}{n^2},$$

while

$$g_2^n\left(\frac{2k}{n}\right) = g_2\left(\frac{2k-1}{n}\right) - \frac{1}{n^2} \text{ and } g_2^n\left(\frac{2k+1}{n}\right) = g_2\left(\frac{2k+1}{n}\right).$$

Note that, for any $n \geq 1$ and any $k \in \mathcal{Z}$, $|k| \leq 2n-1$ g_1^n is decreasing on $[\frac{2k}{n}, \frac{2k+1}{n}]$ while g_2^n is increasing on this interval, and that g_1^n is increasing on $[\frac{2k+1}{n}, \frac{2k+2}{n}]$ while g_2^n is decreasing on this interval, because g_1 and g_2 are increasing on $\mathcal{R}$. Moreover, the $g^n = (g_1^n, g_2^n)$ converge locally uniformly to g.

We first claim that, for any $t \in [0,T)$ and for any $x \in \mathcal{R}$, for n sufficiently large, the unique Nash equilibrium payoff for the game with terminal payoff g^n at (t,x) is $g^n(x)$. Indeed, from Definition 2.5, $g^n(x)$ is a Nash equilibrium payoff for this game. Let $e = g^n(y)$ be another Nash equilibrium payoff. This means that

$$g_j^n(z) \leq g_j^n(y) \quad \text{for any } z \text{ between } x \text{ and } y \text{ and for } j = 1, 2. \tag{4}$$

Suppose first that $y > x$. Then, if $y \in \left(\frac{2k}{n}, \frac{2k+1}{n}\right]$ for some k, we have that $g_1^n(z) > g_1^n(y)$ for any $z \in \left[\frac{2k}{n}, y\right)$ since the function g_1^n is decreasing on $\left[\frac{2k}{n}, \frac{2k+1}{n}\right]$ (provided n is sufficiently large). This is in contradiction with (4). Similarly, if $y \in \left(\frac{2k+1}{n}, \frac{2k+2}{n}\right]$ for some k, then $g_2^n(z) > g_2^n(y)$ for any $z \in \left[\frac{2k+1}{n}, y\right]$ because g_2^n is decreasing on $\left[\frac{2k+1}{n}, \frac{2k+2}{n}\right]$, and again there is a contradiction. We can prove in the same way that the case $y < x$ is impossible. Therefore $y = x$ and $g^n(x)$ is the unique Nash equilibrium payoff for the game with terminal payoff g^n at (t,x).

Let now $E^n(t,x)$ be the FEP for g^n. From Proposition 2.1, we know that $E^n(t,x)$ is a Nash payoff at (t,x) for any $t \in [0,T)$ and almost every $x \in \mathcal{R}$. Therefore $E^n(t,x) = g^n(x)$ for n sufficiently large, because, from the first part of the proof, $g^n(x)$ is the unique Nash payoff. □

Remark: The main reason why the E^n do not converge to E is that, at any $(t, x) \in [0, T] \times \mathcal{R}$, there is a maximal payoff $g(x + 2(T - t))$ at (t, x) for the game with terminal payoff g, and this payoff does not exist—and cannot even be approximated by Nash payoffs—for the game with terminal payoff g^n.

4 Second Example of Instability

In the first example, the function g^n was much more oscillating than the function g. In particular, the intervals on which g_j^n were increasing (respectively decreasing) did not coincide (by far!) with the intervals on which g_j were increasing (respectively decreasing). On the contrary, in the second example that we present now, these intervals do coincide. More precisely, g^n and g are smooth, the g^n converge "as smoothly as possible" to g (for the $\mathcal{C}^\infty$ topology) and g and all the g^n are increasing and decreasing exactly on the same intervals. Despite this, the corresponding FEP for g^n does not converge to the corresponding one for g. Once again the reason for this strange behaviour is the existence for g of a maximal payoff which does not exist for the g^n.

Let $g_1 = g_1^n$ be a smooth strictly increasing function and g_2 and g_2^n be defined by

$$g_2(x) = -(1 - x^2)^2 \qquad \text{and} \qquad g_2^n(x) = g_2(x) + \tfrac{1}{n}\phi(x),$$

where $\phi : \mathcal{R} \to [0, +\infty]$ is any smooth function with support in $[-2, 0]$, increasing on $[-2, -1]$ and decreasing on $[-1, 0]$. Note that g_2 and the g_2^n are increasing on $(-\infty, -1]$ and on $[0, 1]$ and decreasing on $[-1, 0]$ and on $[1, +\infty)$, and that the g_2^n converge to g_2 uniformly in $\mathcal{R}$. Finally, since ϕ has a strict, global maximum at -1, so has g_2^n.

Let $E = (E_1, E_2)$ and $E^n = (E_1^n, E_2^n)$ be the FEPs for the games with terminal payoffs g and g^n respectively.

Proposition 4.1. *For almost any $(t, x) \in [0, T] \times \mathcal{R}$ with $x < -1$ and $x + 2(T - t) > 1$, we have*

$$E_1^n(t, x) \;\leq\; g_1(-1) \;<\; g_1(1) \;=\; E_1(t, x).$$

Remark: In particular, as soon as $T > 1$, the set of points (t, x) such that $x \leq -1$ and $x + 2(T - t) \geq 1$ is nonempty and contains an open set. Hence the set of points (t, x) for which $E^n(t, x)$ does not converge to $E(t, x)$ is "large."

Proof. Let (u^*, v^*) and (u_n^*, v_n^*) be completely maximal Nash equilibrium feedbacks for the games with terminal payoffs g and g^n respectively, and let $\Sigma(t)$ and $\Sigma^n(t)$ be the associated partitions. Recall that $E(t, x) = J(t, x, u^*, v^*)$ for any t and any $x \notin \Sigma(t)$, while $E^n(t, x) = J(t, x, u_n^*, v_n^*)$ for any t and any $x \notin \Sigma_n(t)$.

We first prove that

$$E_1(t, x) \geq g_1(1) \qquad \text{if } x < -1,\ x + 2(T - t) > 1,\ x \notin \Sigma(t). \tag{5}$$

Proof of (5). Let us first notice that, for any $t \in [0, T]$, the payoff $(g_1(0), 0) = g(1)$ is a maximal payoff for the game with terminal payoff g at $(t, 1-2(T-t))$ because g_1 is strictly increasing on $[1-2(T-t), 1]$ and g_2 has a maximum at $y = 1$. Hence $J^\sharp(t, 1-2(T-t), u^*, v^*) = g(1)$ because (u^*, v^*) is completely maximal. From the definition of $J^\sharp$, there is a solution $y(\cdot) \in \mathcal{X}(t, 1-2(T-t), u^*, v^*)$ such that $g_1(y(T)) = g_1(1)$. So $y_1(T) = 1$ because g_1 is strictly increasing, which implies that $y(s) = 1-2(T-s)$ for any $s \in [t, T]$.

Let now (t, x) be such that $x < -1$ with $x + 2(T-t) > 1$, $x \notin S(t)$. Since we are in the real line, the solutions of the equation

$$x'(s) = u^*(s, x(s)) + v^*(s, x(s)) \tag{6}$$

(understood as in Section 2.1) have the following property: if x_1, x_2 are two solutions of this equation on some time interval $[0, t]$, then $\max\{x_1, x_2\}$ is also a solution on $[0, t]$. Let now $x(\cdot) \in \mathcal{X}(t, x, u^*, v^*)$. Then $x_1 := \max\{x, y\}$ is also a solution to (6) with initial condition $x_1(t) = x$: hence $x_1 \in \mathcal{X}(t, x, u^*, v^*)$. Since $x \notin S(t)$ and g_1 is increasing, we have

$$E_1^\sharp(t, x) = g_1(x_1(T)) \geq g_1(y(T)) = g_1(1).$$

So (5) is proved.

We complete the proof by showing that

$$E_1^n(t, x) \leq g_1(-1) \qquad \text{if } x < -1,\ x \notin \Sigma^n(t). \tag{7}$$

Proof of (7). Let $(t, x) \in [0, T) \times \mathcal{R}$ with $x < -1$, $x \notin \Sigma^n(t)$, and let $x(\cdot) \in \mathcal{X}(t, x, u_n^*, v_n^*)$. We know from Proposition 2.1 that $x(\cdot)$ is a Nash trajectory, which means in particular that $g_2^n(x(T)) \geq g_2^n(x(s))$ for any $s \in [t, T]$. Since $x(t) = x \leq -1$ and -1 is the unique point of global maximum of g_2^n, this implies that $x(s) \leq -1$ for any $s \in [t, T]$. Therefore $E_1^n(t, x) = g_1(x(T)) \leq g_1(-1)$ since g_1 is increasing. □

References

[1] Aubin J.-P. (1992) VIABILITY THEORY. Birkhaüser, Boston.

[2] Bressan A. & Shen W. (2004) *Small BV solutions of hyperbolic noncooperative differential games*, SIAM J. Control Optim. 43, no. 1, 194–215.

[3] Bressan A. & Shen W. (2004) *Semi-cooperative Strategies for Differential Games*, Internat. J. Game Theory 32, 4, 561–593.

[4] Cardaliaguet P. & Plaskacz S. (2003) *Existence and uniqueness of a Nash equilibrium feedback for a simple nonzero-sum differential game*, International J. Game Theory 32, pp. 33–71.

[5] Evans L.C. & Souganidis P.E. (1984) *Differential games and representation formulas for solutions of Hamilton-Jacobi Equations*, Indiana Univ. Math. J., 33, 773–797.

[6] Kleimenov A.F. (1993) NONANTAGONIST DIFFERENTIAL GAMES (in Russian) “Nauka” Uralprime skoe Otdelenie, Ekaterinburg.

[7] Kononenko, A.F. (1976) *On equilibrium positional strategies in nonantagonistic differential games.* Dokl. Akad. Nauk SSSR 231, 285–288.

[8] Mannucci P. (2004/2005) *Nonzero-sum stochastic differential games with discontinuous feedback*, SIAM J. Control Optim. 43 4, 1222–1233 (electronic).

[9] Tolwinski B., Haurie A. & Leitmann G. (1986) *Cooperative equilibria in differential games.* J. Math. Anal. Appl. 119, 182–202.

Constructing Robust Control in Differential Games: Application to Aircraft Control During Landing

S. A. Ganebny
Institute of Mathematics and Mechanics
Russian Academy of Sciences
S. Kovalevskaya Street 16 Ekaterinburg 620219, Russia

S. S. Kumkov
Institute of Mathematics and Mechanics
Russian Academy of Sciences
S. Kovalevskaya Street 16 Ekaterinburg 620219, Russia

V. S. Patsko
Institute of Mathematics and Mechanics
Russian Academy of Sciences
S. Kovalevskaya Street 16 Ekaterinburg 620219, Russia
patsko@imm.uran.ru

S. G. Pyatko
Academy of Civil Aviation
Pilotov Street 38
St. Petersburg 196210, Russia
nita@nita.ru

Abstract

An approach for constructing a robust feedback control for problems having linear dynamics with disturbances is suggested. A useful control is assumed to be scalar and bounded.

The approach can be applied to conflict-controlled systems, where the constraint for the disturbance is unknown in advance. Adjustment of the method is based on results of the theory of linear differential games with fixed terminal time and geometric constraints for the players' controls.

The algorithm for constructing the robust control is fulfilled as a computer program for the case when the quality of the process is defined only by two components of the phase vector at the terminal instant. The paper presents simulation results for the problem of lateral motion control of an aircraft during landing under wind disturbance.

Key words. Differential game, stable sets, robust control, switching surfaces, numerical constructions, aircraft landing, wind disturbance.

Introduction

In the theory of antagonistic differential games (see, for example, [18, 22, 3]), formulations are typical, where constraints both for the useful control and disturbance are given *a priori*. But in many practical situations, introducing the exact constraint for a disturbance is unnatural. For example, when a problem of aircraft landing is formulated, it is difficult to explain why the contstraint on the possible deviation of the wind velocity from some average value is taken to be 10 m/sec, rather than, for example, 12 m/sec. With that, the optimal strategy obtained from the solution of a corresponding differential game depends on the taken disturbance level.

Let us agree that a control is called *robust* if, in the case of a "low" disturbance (which is unknown in advance), it provides good quality of the process by some "low" level useful control. With an increase in the disturbance level, the level of useful control guaranteeing good quality of the process grows too. This sense of the concept "robust control" coheres with that used in the mathematical literature (for example, see [14, 13, 38]).

An approach to constructing a linear robust control (for an H^{∞}-problem) on the basis of the theory of differential games with linear-quadratic cost functional is described in [2].

This chapter suggests an alternative approach to constructing a nonlinear robust control. This method is oriented to problems with linear dynamics, where the constraint for the useful control is prescribed. The method is based on results of the theory of differential games with geometric constraints for the players' controls.

The capacities of the method suggested are illustrated in the final part of the paper by a control problem of lateral motion of an aircraft during landing under wind disturbance. Its dynamics is described by a linearized system.

1 Construction of a Robust Control

1.1 Formulation of the Problem

Let us consider a linear differential game with fixed terminal time:

$$\begin{gathered}\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)u + \mathbf{C}(t)v,\\ \mathbf{x} \in \mathbb{R}^m, \quad t \in T, \quad u \in P = \{u \in \mathbb{R} : |u| \leq \mu\}, \quad v \in \mathbb{R}^q.\end{gathered} \tag{1}$$

Here, P is the constraint for the first player's scalar control, and $T = [\vartheta_0, \vartheta]$ is the time interval of the game. The matrix functions $\mathbf{A}$ and $\mathbf{C}$ are continuous. The vector function $\mathbf{B}$ is Lipschitzian in the time interval T.

The first player tries to lead system (1) to a set M at the terminal instant ϑ. The second player hinders this. The set M is assumed to be a convex compactum in a subspace of n chosen components of the vector $\mathbf{x}$. Let us assume that the set M includes a neighborhood of the origin of this subspace.

Unlike the standard formulation [18, 21, 22] of a differential game, system (1) does not include any constraint for the second player's control v.

Informally, a first player's robust control can be understood as a feedback control $U(t, \mathbf{x})$ obeying the following conditions:

- if the second player applies a "low level" control, the first player should lead the system to the terminal set closely to its center. Moreover, the realization of the first player's control should also be of a "low level;"
- if the second player's control is "stronger," the first player should still lead the system to the terminal set, maybe by a "stronger" or even maximal control;
- in the case when the second player involves "very strong" control and the first player (acting within the framework of his constraint) cannot guarantee reaching the terminal set, then he may allow some terminal miss, but he tries to minimize it.

It is necessary to elaborate a method for constructing such a robust feedback control for system (1).

1.2 Differential Game Without Phase Variable on the Right-Hand Side

By using a standard change of variables [21, 22], let us pass to a system whose right-hand side does not include the phase vector:

$$\begin{gathered} \dot{x} = B(t)u + C(t)v, \\ x \in \mathbb{R}^n, \quad t \in T, \quad u \in P, \quad v \in \mathbb{R}^q. \end{gathered} \tag{2}$$

The passage is provided by the following relations:

$$x(t) = X_{n,m}(\vartheta, t)\mathbf{x}(t), \quad B(t) = X_{n,m}(\vartheta, t)\mathbf{B}(t), \quad C(t) = X_{n,m}(\vartheta, t)\mathbf{C}(t),$$

where $X_{n,m}(\vartheta, t)$ is the matrix combined of n corresponding rows of the fundamental Cauchy matrix of the system $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$.

Here again, the first player tries to lead system (2) to the set M at the terminal instant ϑ, and the second one hinders this. The set M is a convex compact set in $\mathbb{R}^n$ including a neighborhood of the origin.

All following considerations will be given for system (2). The robust control obtained for system (2) will be adopted for system (1).

1.3 Stable Bridges: Addition and Multiplication by a Coefficient

Let us consider a standard differential game with fixed terminal time ϑ, a terminal set $\mathcal{M}$, and geometric constraints $\mathcal{P}$ and $\mathcal{Q}$ for the players' controls:

$$\begin{gathered} \dot{x} = B(t)u + C(t)v, \\ x \in \mathbb{R}^n, \quad t \in T, \quad \mathcal{M}, \quad u \in \mathcal{P}, \quad v \in \mathcal{Q}. \end{gathered} \tag{3}$$

Convex compact sets $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{M}$ will be considered as parameters of the game. The authors use a formalization of game (3) according to [21, 22].

For an arbitrary set $E \subset T \times \mathbb{R}^n$, let us define its section at an instant t by the formula $E(t) = \{x \in \mathbb{R}^n : (t, x) \in E\}$.

Below, operations of addition and multiplication by a nonnegative scalar will be introduced for a set from the space $T \times \mathbb{R}^n$ having nonempty sections at any instant $t \in T$.

Definition 1.1. The sum of two sets $E_1, E_2 \subset T \times \mathbb{R}^n$ is a set

$$E_1 + E_2 = \big\{(t, x) \in T \times \mathbb{R}^n : x \in E_1(t) + E_2(t)\big\}.$$

Definition 1.2. Multiplication of a set $E \subset T \times \mathbb{R}^n$ by a real number $k \geq 0$ is a set

$$kE = \{(t, x) \in T \times \mathbb{R}^n : x \in kE(t)\}.$$

On the basis of [21, 22], let us give definitions of stable and maximal stable bridges. Below, $u(\cdot)$ and $v(\cdot)$ will denote some measurable functions of time with their values in the sets $\mathcal{P}$ and $\mathcal{Q}$, respectively. The symbol $x\big(\cdot; t_*, x_*, u(\cdot), v(\cdot)\big)$ will denote a motion of system (2) under controls $u(\cdot)$ and $v(\cdot)$ emanating from the point x_* at the instant t_*.

Definition 1.3. A set $W \subset T \times \mathbb{R}^n$ is called a stable bridge for system (3) with some fixed $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{M}$ if $W(\vartheta) = \mathcal{M}$ and it possesses the following property. For any position $(t_*, x_*) \in W$ and any second player's control $v(\cdot)$, the first player can choose his control $u(\cdot)$ such that the pair $\big(t, x(t)\big) = \big(t, x(t; t_*, x_*, u(\cdot), v(\cdot))\big)$ stays in W at any instant $t \in (t_*, \vartheta]$, and, consequently, the motion $x(t)$ reaches the set $\mathcal{M}$ at the terminal instant ϑ: $x(\vartheta) \in \mathcal{M}$.

Definition 1.4. A set possessing the stability property and maximal by inclusion in the space $T \times \mathbb{R}^n$ is called the maximal stable bridge for system (3).

A concept close to the maximal stable bridge is *Pontryagin's alternating integral* [33]. In the framework of the viability approach [1], the notion of *discriminating domain* corresponds to the notion of stable bridge, and the notion of *discriminating kernel* is the analogue to the maximal stable bridge [12, pp. 215–219].

We introduce the following notation: $\mathcal{W}(\mathcal{P}, \mathcal{Q}, \mathcal{M})$ is the collection of all stable bridges for system (3) with parameters $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{M}$; $\mathbb{W}(\mathcal{P}, \mathcal{Q}, \mathcal{M})$ is the maximal stable bridge for game (3) with parameters $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{M}$.

Proposition 1.1. *Let $F \in \mathcal{W}(\mathcal{P}, \mathcal{Q}, \mathcal{M})$ with some $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{M}$. Then $kF \in \mathcal{W}(k\mathcal{P}, k\mathcal{Q}, k\mathcal{M})$ for any $k \geq 0$.*

Proof. When $k = 0$, the proposition is evident. Below, it is supposed that $k > 0$.

Fix an arbitrary point $(t_*, x_*) \in kF$ and an instant $t^* \in [t_*, \vartheta]$. Let the second player choose some control $v(t) \in k\mathcal{Q}$ in the interval $[t_*, t^*]$. We shall show how a first player's control $u(t) \in k\mathcal{P}$, $t \in [t_*, t^*]$, can be constructed such that the inclusion $\big(t^*, x(t^*)\big) \in kF$ holds for the motion $x(\cdot) = x\big(\cdot; t_*, x_*, u(\cdot), v(\cdot)\big)$.

Denote $z_* = {}^1\!/_k \cdot x_*$, $\bar{v}(t) = {}^1\!/_k \cdot v(t)$. One has $(t_*, z_*) \in F$. Since F is a stable set, for any second player's control $\bar{v}(t) \in \mathcal{Q}$, $t \in [t_*, t^*]$, it is possible to find a control $\bar{u}(t) \in \mathcal{P}$, $t \in [t_*, t^*]$, such that the motion $z(\cdot) = x\big(\cdot; t_*, z_*, \bar{u}(\cdot), \bar{v}(\cdot)\big)$ gives the inclusion $\big(t^*, z(t^*)\big) \in F$.

Let $u(t) = k\bar{u}(t)$, $t \in [t_*, t^*]$. Taking into account linearity of system (2) on the controls and absence of the phase variable in its right-hand side, one gets that $x(t) = kz(t)$ for any $t \in [t_*, t^*]$. Thus, $\big(t^*, x(t^*)\big) \in kF$, which means stability of the set kF. □

Proposition 1.2. *Multiplication of the maximal stable bridge, corresponding to parameters* $(\mathcal{P}, \mathcal{Q}, \mathcal{M})$*, by a number* $k \geq 0$ *is the maximal stable bridge, corresponding to parameters* $(k\mathcal{P}, k\mathcal{Q}, k\mathcal{M})$*:*

$$k\mathbb{W}(\mathcal{P}, \mathcal{Q}, \mathcal{M}) = \mathbb{W}(k\mathcal{P}, k\mathcal{Q}, k\mathcal{M}).$$

Proof. When $k = 0$, the proposition is evident. Below, we suppose that $k > 0$.

Let $F = \mathbb{W}(\mathcal{P}, \mathcal{Q}, \mathcal{M})$. Denote $\widetilde{F} = kF$. From Proposition 1.1 one gets $\widetilde{F} \in \mathcal{W}(k\mathcal{P}, k\mathcal{Q}, k\mathcal{M})$.

Assume that $\widetilde{F} \neq \mathbb{W}(k\mathcal{P}, k\mathcal{Q}, k\mathcal{M})$. Let $\widetilde{\mathbf{F}} = \mathbb{W}(k\mathcal{P}, k\mathcal{Q}, k\mathcal{M})$. One has $\widetilde{\mathbf{F}} \supset \widetilde{F}$, $\widetilde{\mathbf{F}} \neq \widetilde{F}$.

Consider the set $\mathbf{F} = {}^1\!/_k \cdot \widetilde{\mathbf{F}}$. Then $\mathbf{F} \supset F$, $\mathbf{F} \neq F$. Proposition 1.1 gives that the inclusion $\mathbf{F} \in \mathcal{W}(\mathcal{P}, \mathcal{Q}, \mathcal{M})$ is true. But $F = \mathbb{W}(\mathcal{P}, \mathcal{Q}, \mathcal{M})$, a contradiction. □

Proposition 1.3. *The sum of two stable bridges* F_1 *and* F_2*, corresponding to parameters* $(\mathcal{P}_1, \mathcal{Q}_1, \mathcal{M}_1)$ *and* $(\mathcal{P}_2, \mathcal{Q}_2, \mathcal{M}_2)$ *respectively, is a stable bridge corresponding to the parameters* $(\mathcal{P}_1 + \mathcal{P}_2, \mathcal{Q}_1 + \mathcal{Q}_2, \mathcal{M}_1 + \mathcal{M}_2)$*:*

$$F_1 + F_2 \in \mathcal{W}(\mathcal{P}_1 + \mathcal{P}_2, \mathcal{Q}_1 + \mathcal{Q}_2, \mathcal{M}_1 + \mathcal{M}_2). \tag{4}$$

Proof. Let $\widetilde{F} = F_1 + F_2$. Take an arbitrary point $(t_*, x_*) \in \widetilde{F}$ and an instant $t^* \in [t_*, \vartheta]$. Let the second player choose a control $v(t) \in \mathcal{Q}_1 + \mathcal{Q}_2$, $t \in [t_*, t^*]$. Below, a first player's control $u(t) \in \mathcal{P}_1 + \mathcal{P}_2$, $t \in [t_*, t^*]$, will be constructed such that for the motion $x(\cdot) = x\big(\cdot; t_*, x_*, u(\cdot), v(\cdot)\big)$ the inclusion $\big(t^*, x(t^*)\big) \in \widetilde{F}$ is held.

Let us choose points z_*^1 and z_*^2 such that $(t_*, z_*^1) \in F_1$, $(t_*, z_*^2) \in F_2$, and $z_*^1 + z_*^2 = x_*$.

Take some controls $v_1(\cdot)$ and $v_2(\cdot)$ such that $v_1(t) \in \mathcal{Q}_1$, $v_2(t) \in \mathcal{Q}_2$, and $v_1(t) + v_2(t) = v(t)$, $t \in [t_*, t^*]$.

Let $i = 1, 2$. Using the stability of sets F_i, one can find controls $u_i(t) \in \mathcal{P}_i$, $t \in [t_*, t^*]$, so that the motions $z^i(\cdot) = x\big(\cdot; t_*, z_*^i, u_i(\cdot), v_i(\cdot)\big)$ give the inclusions $\big(t^*, z^i(t^*)\big) \in F_i$.

Denote $u(t) = u_1(t) + u_2(t)$, $t \in [t_*, t^*]$. Then using linearity of system (2) on the controls and absence of the phase variable in its right-hand side, one gets that $x(t) = z^1(t) + z^2(t)$. So, $\big(t^*, x(t^*)\big) \in \widetilde{F}$, which implies the inclusion (4). □

Remark 1.1. Let $F_1 = \mathbb{W}(\mathcal{P}_1, \mathcal{Q}_1, \mathcal{M}_1)$ and $F_2 = \mathbb{W}(\mathcal{P}_2, \mathcal{Q}_2, \mathcal{M}_2)$, i.e., the sets F_1 and F_2 are maximal stable bridges. Then the inclusion (4) is true. An example can be easily constructed when $F_1 + F_2 \neq \mathbb{W}(\mathcal{P}_1 + \mathcal{P}_2, \mathcal{Q}_1 + \mathcal{Q}_2, \mathcal{M}_1 + \mathcal{M}_2)$. So, under the addition of maximal stable bridges, the result is stable, but generally speaking, not maximally stable.

1.4 Robust Feedback Control

We now describe the procedure for constructing robust controls for systems (2) and (1).

1) Let us choose a set $Q_{\max} \subset \mathbb{R}^q$ that will represent the "maximal" constraint on control of the second player, which the first player "agrees" to consider reasonable for the problem of guiding system (2) to the set M. In other words, we choose the set $Q_{\max}$ quite large, but such that if $v(t) \in Q_{\max}$ for any $t \in T$, the first player guarantees leading the system to the terminal set M. The set $Q_{\max}$ should include the origin. Denote by W the maximal stable bridge corresponding to the parameters P, $Q_{\max}$, and M:

$$W = \mathbb{W}(P, Q_{\max}, M).$$

Let $\mathbb{B}(\varepsilon) = \{x \in \mathbb{R}^n : \|x\| \leq \varepsilon\}$ be a ball in $\mathbb{R}^n$ with radius ε and the center at the origin.

Let us agree that the set $Q_{\max}$ is such that

$$\exists\, \varepsilon > 0 : \forall\, t \in T \;\; \mathbb{B}(\varepsilon) \subset W(t). \tag{5}$$

This demands inclusion of the origin with some neighborhood into each time section $W(t)$ of the bridge W.

2) To construct the robust control, one also needs a maximal stable bridge satisfying the following conditions: the first player's control is absent ($P = \{0\}$), the second player's control is constrained by $Q_{\max}$, and the terminal set is ρM, where $\rho > 0$. Let

$$\widehat{W} = \mathbb{W}(\{0\}, Q_{\max}, \rho M). \tag{6}$$

In other words, the set $\widehat{W}$ is collection of all positions (t, x), from which the terminal set ρM can be reached at the terminal instant ϑ without control and under any disturbance bounded by $Q_{\max}$.

Let the chosen multiplier ρ be small enough, but such that

$$\exists\, \varepsilon > 0 : \forall\, t \in T \;\; \mathbb{B}(\varepsilon) \subset \widehat{W}(t). \tag{7}$$

This assumption gives that time sections $\widehat{W}(t)$ of the set $\widehat{W}$ for any instant in the game interval are not empty and are not degenerated to a set of a dimension less than n.

3) We construct a family of sets

$$W_k = \begin{cases} kW, & \text{if } 0 \le k \le 1, \\ W + (k-1)\widehat{W}, & \text{if } k > 1. \end{cases}$$

Due to (5) and (7), one has inclusions

$$W_{k_1} \subset W_{k_2} \subset W \subset W_{k_3} \subset W_{k_4}$$

for any $0 < k_1 < k_2 < 1 < k_3 < k_4$.

Proposition 1.2 gives that the sets W_k, when $0 \le k \le 1$, are maximal stable bridges corresponding to parameters $(kP, kQ_{\max}, kM)$. Propositions 1.1 and 1.3 imply that the sets W_k, when $k > 1$, are stable bridges, constructed with parameters $\bigl(P, kQ_{\max}, M + (k-1)\rho M\bigr)$.

So, with increasing coefficient k, one obtains a growing collection of stable bridges, where any larger bridge corresponds to a larger constraint for the second player's control.

The main idea of the suggested method for constructing a robust feedback control is the following: when the second player's control $v(t)$ for any t belongs to some constraint $k^* Q_{\max}$ and the initial position is in the bridge W_{k^*} corresponding to this number k^*, then the system will not leave this bridge.

4) Define a function $V : T \times \mathbb{R}^n \to \mathbb{R}$ in the following way:

$$V(t, x) = \min\bigl\{k \ge 0 : (t, x) \in W_k\bigr\}.$$

The level sets (Lebesgue sets) of this function coincide with the stable bridges.

Because the set $Q_{\max}$ and the number ρ are chosen such that for some $\varepsilon > 0$ relations (5) and (7) hold, then the function $x \mapsto V(t, x)$ for any $t \in T$ satisfies the Lipschitz condition with the constant $\lambda = 1/\varepsilon$.

Denote by $\mathcal{A}(t, x)$ a line in the space $\mathbb{R}^n$ parallel to the vector $B(t)$ and passing through the point x:

$$\mathcal{A}(t, x) = \{z \in \mathbb{R}^n : z = x + \alpha B(t),\ \alpha \in \mathbb{R}\}.$$

Let

$$\mathcal{V}(t, x) = \min_{z \in \mathcal{A}(t,x)} V(t, x).$$

The minimum is reached, since the function $x \mapsto V(t, x)$ is continuous and tends to infinity with $|x| \to \infty$. Because the function is quasiconvex (i.e., all its Lebesgue sets are convex), the set of points, where the minimum is reached, is either a point or a segment.

If $B(t) = 0$, it is assumed $\mathcal{V}(t, x) \equiv V(t, x)$.

5) For any $t \in T$ let

$$\Pi(t) = \bigl\{x \in \mathbb{R}^n : V(t, x) = \mathcal{V}(t, x)\bigr\},$$
$$\Pi_-(t) = \bigl\{x \in \mathbb{R}^n : x + \alpha B(t) \notin \Pi(t),\ \forall \alpha \ge 0\bigr\},$$
$$\Pi_+(t) = \bigl\{x \in \mathbb{R}^n : x + \alpha B(t) \notin \Pi(t),\ \forall \alpha \le 0\bigr\}.$$

The set $\Pi(t)$ is closed, the sets $\Pi_-(t)$ and $\Pi_+(t)$ are on different sides of $\Pi(t)$. So, these three sets divide the space $\mathbb{R}^n$ into three parts.

6) We define a function

$$\overline{V}(t, x) = \min\{V(t, x), 1\}$$

and a multifunction

$$\mathbf{U}^0(t, x) = \begin{cases} -\overline{V}(t, x)\mu, & \text{if } x \in \Pi_-(t), \\ \overline{V}(t, x)\mu, & \text{if } x \in \Pi_+(t), \\ \big[-\overline{V}(t, x)\mu, \overline{V}(t, x)\mu\big], & \text{if } x \in \Pi(t). \end{cases}$$

As the strategy U of the first player, let us take an arbitrary one-valued selection from the multifunction $\mathbf{U}^0$:

$$U(t, x) \in \mathbf{U}^0(t, x) \qquad \forall (t, x) \in T \times \mathbb{R}^n.$$

Thus, the control $U(t, x)$ "switches" at the set $\Pi(t)$. For demonstrativeness, the set $\Pi(t)$ is called a *switching surface* corresponding to the instant t.

7) Suppose that the first player applies the strategy U in a discrete scheme of control [21, 22] with a time step Δ. In any interval of the discrete scheme, the generated control is constant. Taking an open-loop control $v(\cdot)$ of the second player and an initial position (t_0, x_0), one gets a motion $t \mapsto x(t)$ of system (2) and a realization $u(\cdot)$ of the first player's control.

Let the symbol β denote the Lipschitz constant of the function $B(t)$ and $\sigma = \max\big\{|B(t)| : t \in T\big\}$. As it was introduced above, λ denotes the Lipschitz constant of the function $V(t, x)$ on x.

The following theorem about guarantee is true.

Theorem. *Let U be some strategy of the first player such that $U(t, x) \in \mathbf{U}^0(t, x)$ for all $(t, x) \in T \times \mathbb{R}^n$. Choose arbitrary $t_0 \in T$, $x_0 \in \mathbb{R}^n$, and $\Delta > 0$. Suppose that in the interval $[t_0, \vartheta]$ the second player's control is bounded by a set $k^* Q_{\max}$, $k^* \geqslant 0$. Denote*

$$c^* = V(t_0, x_0), \qquad s^* = \max(k^*, c^*).$$

Let $x^*(\cdot)$ be the motion of system (2) emanating from the point x_0 at the instant t_0 under the control U in a discrete scheme with the time step Δ and some control $v(\cdot)$ of the second player. Then the realization $u(t) = U\big(t, x^*(t)\big)$ of the first player's control obeys to the inclusion

$$u(t) \in \min\big(s^* + \Lambda(t, t_0, \Delta), 1\big)P, \qquad t \in [t_0, \vartheta].$$

With that, the value $V\big(t, x^*(t)\big)$ of the function V satisfies the inequality

$$V\big(t, x^*(t)\big) \leqslant s^* + \Lambda(t, t_0, \Delta), \qquad t \in [t_0, \vartheta].$$

Here,

$$\Lambda(t, t_0, \Delta) = 2\lambda\mu\sqrt{2\sigma\Delta\beta}(t - t_0) + 4\lambda\sigma\mu\Delta.$$

A proof of the theorem is given in [15]. The proof significantly uses scheme of speculations from works [11, 29, 30]. There, the switching surfaces were used to build optimal feedback control of the minimizing player in linear antagonistic differential games with fixed terminal time and geometric constraints on players' controls.

So, any one-valued selection U from the multifunction $\mathbf{U}^0$ gives a robust feedback control for system (2).

8) Returning to system (1), let us introduce a multifunction

$$\widetilde{\mathbf{U}}^0(t, \mathbf{x}) = \mathbf{U}^0\big(t, X_{n,m}(\vartheta, t)\mathbf{x}\big).$$

Its one-valued selection $\widetilde{U}(t, \mathbf{x})$ gives a robust control for system (1).

The construction of robust feedback control is described. It essentially uses the ordering of the stable sets W_k, when $k \geq 0$, and is based on the concept of a switching surface $\Pi(t)$, which changes in time.

For numerical construction of robust control, one should keep sections $W(t)$ of the bridge W and switching surfaces $\Pi(t)$ for some grid $\{t_i\}$ of time instants. Having at the instant t a position $\mathbf{x}(t)$ of system (1), one transforms it to the coordinates of system (2) by the mapping $x(t) = X_{n,m}(\vartheta, t)\mathbf{x}(t)$. The sign of the control $\widetilde{U}\big(t, \mathbf{x}(t)\big) = U\big(t, x(t)\big)$ is defined by the relative position of the point $x(t)$ with respect to the switching surface $\Pi(t)$. Analyzing the position of the point $x(t)$ with respect to the boundary of the section $W(t)$ of the bridge W, one computes the absolute value $\big|\widetilde{U}\big(t, \mathbf{x}(t)\big)\big|$. Here, homothety of sets $W_k(t)$ for $k \leq 1$ is used.

Remark 1.2. The authors do not state that the suggested method for robust control realizes any optimality criterion. Note also that when constructing the robust control, one has some freedom in the choice of the set $Q_{\max}$ and the number ρ.

Remark 1.3. One can see from the given explanation of constructing robust control that its central place is the procedure for building maximal stable bridges for linear differential games of type (3). Some variants of backward constructing time sections of maximal stable bridge or, what is the same, Pontryagin's alternating integral are described in many works. The papers [17, 5, 36, 28, 23] deal with two-dimensional problems ($n = 2$), the papers [41, 42, 37]. concern the three-dimensional case ($n = 3$). The case of an arbitrary dimension is considered in [7, 40, 16, 32]. A method for *a posteriori* evaluation of the numerical construction error in Hausdorff metrics is suggested in [4]. Computer realization of this method in the case $n = 2$ is given in [9]. Some approach for constructing time sections of maximal stable bridge based on elliptical approximation is studied in [24].

1.5 Construction of Robust Control for a Two-Dimensional System (2)

If the original controlled system (1) has the set M defined only by two coordinates of the phase vector $\mathbf{x}$ (i.e., $n = 2$), then after passage to system (2) one gets a new two-dimensional phase vector x. In this case, the sections $W(t)$ and $\widehat{W}(t)$ of maximal stable bridges W and $\widehat{W}$, which correspond to parameters $(P, Q_{\max}, M)$ and $\big(\{0\}, Q_{\max}, \rho M\big)$, are in the plane.

Generation of the robust control consists of two steps: choosing the sign of the control and its absolute value. At any instant t, one has a family of inserted sets $W_k(t)$ (see Figure 1, where $k_1 < 1 < k_2$). Let us find points at the boundary of these sets, where the support line is parallel to the vector $B(t)$. Joining the points obtained in this way, one gets the *switching line* $\Pi(t)$ defining the sign of the control. The absolute value of the control can be computed by the formula

$$|U(t, x)| = \begin{cases} \dfrac{l}{L}\,\mu, & \text{if } x \in W(t), \\ \mu, & \text{if } x \notin W(t). \end{cases}$$

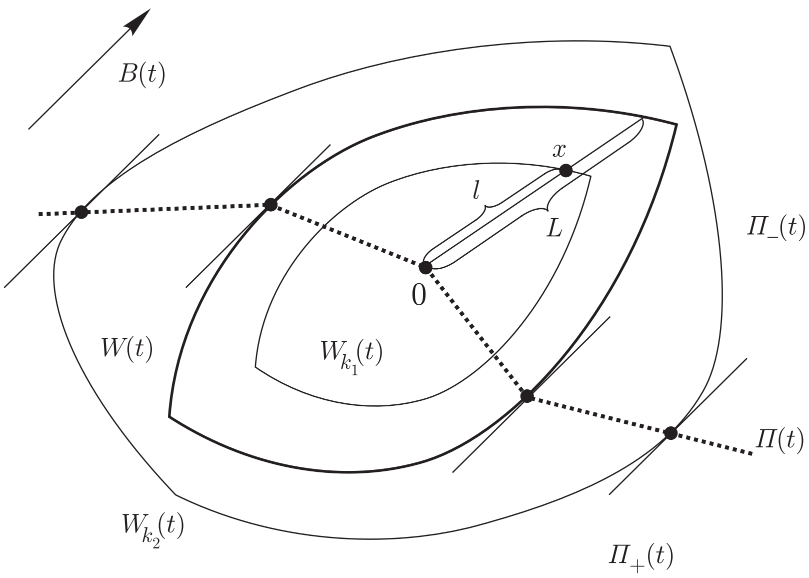


Figure 1: Constructing the robust control in the case $n = 2$.

In this formula, l is the length of the phase vector x, and L is the length of the segment connecting the origin with the boundary of the set $W(t)$ and containing the point x.

To construct robust feedback control one has to store sections $W(t_i)$ of the set W and switching lines $\Pi(t_i)$ in some grid $\{t_i\}$ of time instants. Sections $W(t_i)$ are convex polygons and can be kept in some appropriate form. Each switching line $\Pi(t_i)$ is a polygonal line having four linear parts, and to keep it in computer memory one has to store five of its vertices (one of them is the origin).

Following [17, 28, 23, 16, 32], let us describe the main idea of the procedure for constructing time sections of the maximal stable bridge for systems of type (3) in the case $n = 2$. Actually, this procedure is an application of the dynamic programming method to differential games. Below, we keep the denotation W for the maximal stable bridge.

The system is discretized. In the game interval $T = [\vartheta_0, \vartheta]$, a time grid $\vartheta_0 < \vartheta_1 < \cdots < \vartheta_N = \vartheta$ is introduced. The dynamics of system (3) in each interval $[\vartheta_i, \vartheta_{i+1})$ is changed by a constant one:

$$\dot{x} = B^\sharp(t)u + C^\sharp(t)v, \quad B^\sharp(t) = B(\vartheta_i), \;\; C^\sharp(t) = C(\vartheta_i), \;\; t \in [\vartheta_i, t_{\vartheta+1}).$$

Instead of the sets $\mathcal{M}$, $\mathcal{P}$, and $\mathcal{Q}$, some their convex polygonal (for $\mathcal{P}$ and $\mathcal{Q}$ possibly polyhedral) approximations $\mathcal{M}^\sharp$, $\mathcal{P}^\sharp$, and $\mathcal{Q}^\sharp$ are considered. If the set $\mathcal{P}$ is a line segment, then it is a polygon already.

The procedure produces a collection of convex polygons $W^\sharp(\vartheta_i)$ corresponding to the instants from the taken grid and approximating convex sections $W(\vartheta_i)$ of the maximal stable bridge W. All used convex polygonal sets are described by their support functions. Recall, that the support function $l \mapsto c(l; A)$ of a convex compact set is defined as $c(l; A) = \max\big\{\langle l, a\rangle : a \in A\big\}$. The support function is positively homogeneous and, if the set is polygonal, piecewise-linear in cones defined by outer normals of neighbor edges.

The section $W^\sharp(\vartheta_N)$ at the terminal instant $\vartheta_N = \vartheta$ is assume equal to the approximation of the terminal set: $W^\sharp(\vartheta_N) = \mathcal{M}^\sharp$. The support function of the next (in backward time) section $W^\sharp(\vartheta_i)$ is defined [34] as convex hull of the following positively homogeneous function:

$$\gamma(l, \vartheta_i) = c\big(l; W^\sharp(\vartheta_{i+1})\big) + \delta_i c\big(l; -B^\sharp(\vartheta_i)\mathcal{P}^\sharp\big) - \delta_i c\big(l; C^\sharp(\vartheta_i)\mathcal{Q}^\sharp\big),$$

where $l \in \mathbb{R}^2$, $\delta_i = \vartheta_{i+1} - \vartheta_i$.

With decreasing the diameter of the grid and improvement of approximations $\mathcal{M}^\sharp$, $\mathcal{P}^\sharp$, and $\mathcal{Q}^\sharp$, the produced polygons converge in Hausdorff metrics to the true time sections of the bridge W.

Corresponding procedure with change of polygons by polyhedra can applied in the space $\mathbb{R}^n$ of any dimension. But for the case $n = 2$, the procedures for constructing convex hull are the most simple and effective. Moreover, it is possible to improve them by using specific features of the problem, namely, the information about places of possible violation of the local convexity of the function $\gamma(\cdot, l)$ [17, 28, 23].

2 Robust Control in Problem of Aircraft Landing

2.1 Formulation of the Problem

During the last 20 years, many publications have dealt with the application of methods of modern control theory and differential game theory to problems of landing and takeoff of an aircraft under wind disturbances (see, for example, [26, 27, 25, 10, 31, 35], and references therein).

In this chapter, the problem of lateral motion of midsize transport aircraft during landing is considered. The problem is investigated in the time interval $[0, \vartheta]$, where ϑ is the instant of passing the runway threshold. The instant ϑ is assumed to be fixed. Such an assumption is reasonable if to suppose that the longitudinal motion of the aircraft does not depend on the lateral one and obeys to a given program.

A linear approximation of the motion dynamics is described [20, 6, 8] by the following system of differential equations:

$$
\begin{aligned}
\dot{\mathbf{x}}_1 &= \mathbf{x}_2,\\
\dot{\mathbf{x}}_2 &= -0.0762\mathbf{x}_2 - 5.34\mathbf{x}_3 + 9.81\mathbf{x}_5 + 0.0762v,\\
\dot{\mathbf{x}}_3 &= \mathbf{x}_4,\\
\dot{\mathbf{x}}_4 &= -0.0056\mathbf{x}_2 - 0.392\mathbf{x}_3 - 0.0889\mathbf{x}_4 - 0.0378\mathbf{x}_5 - 0.17\mathbf{x}_6\\
&\quad + 0.0378\mathbf{x}_7 + 0.0056v,\\
\dot{\mathbf{x}}_5 &= -\mathbf{x}_5 + \mathbf{x}_7,\\
\dot{\mathbf{x}}_6 &= -0.0129\mathbf{x}_2 - 0.9016\mathbf{x}_3 - 0.2045\mathbf{x}_4 - 0.0869\mathbf{x}_5 - 0.89\mathbf{x}_6\\
&\quad + 0.0869\mathbf{x}_7 + 0.0129v,\\
\dot{\mathbf{x}}_7 &= -\mathbf{x}_7 + u.
\end{aligned}
\tag{8}
$$

Components of the phase vector $\mathbf{x}$ have the following physical sense:

$\mathbf{x}_1$ — lateral deviation of the mass center of the aircraft from the central axis of the runway;
$\mathbf{x}_2$ — velocity of the lateral deviation;
$\mathbf{x}_3$ — yaw angle counted clockwise from the runway axis;
$\mathbf{x}_4$ — angular velocity of yaw angle;
$\mathbf{x}_5$ — roll angle;
$\mathbf{x}_6, \mathbf{x}_7$ — auxiliary variables.

The control u can be treated as the required roll angle. The disturbance parameter v is the lateral component of the wind velocity. The lateral deviation is measured in meters, angles are measured in radians, and time is measured in seconds.

One can see that the control u and the roll angle $\mathbf{x}_5$ are connected by two simple differential equations (written in the 5th and 7th rows of system (8)), which describe the process of transformation of the control signal in the roll channel. Due to this, the variable $\mathbf{x}_7$ involved in the mentioned rows is called auxiliary.

The variable $\mathbf{x}_6$ is composite. It connects the deviation of rudder and the roll angle. During the landing stage till passing the runway threshold, the rudder is used to stabilize the sideslip angle near zero. The corresponding control law for the rudder is included to system (8).

In more details, some similar model is discussed in [19, pp. 180–181].

The required roll angle is bounded:

$$|u| \leq \mu = 0.2613\,\text{rad}.$$

This constraint defines the set P.

In the subspace of phase variables $\mathbf{x}_1$ and $\mathbf{x}_2$, let us define a set

$$M = \left\{ (\mathbf{x}_1, \mathbf{x}_2) : \frac{\mathbf{x}_1^2}{216} - \frac{2\mathbf{x}_1}{9} - \frac{3}{2} \leq \mathbf{x}_2 \leq -\frac{\mathbf{x}_1^2}{216} - \frac{2\mathbf{x}_1}{9} + \frac{3}{2} \right\}.$$

If at the instant ϑ the lateral deviation $\mathbf{x}_1(\vartheta)$ and its velocity $\mathbf{x}_2(\vartheta)$ are such that $\big(\mathbf{x}_1(\vartheta), \mathbf{x}_2(\vartheta)\big) \in M$, then it is supposed that successful landing is provided after the instant ϑ. Otherwise, if $\big(\mathbf{x}_1(\vartheta), \mathbf{x}_2(\vartheta)\big) \notin M$, there is no such a guarantee. So, the set M is the tolerance for the coordinates $\mathbf{x}_1$, $\mathbf{x}_2$ at the instant ϑ.

After the instant ϑ, there are 3 final stages of landing (see Fig. 2): descent till contact the runway by main wheels, landing run on the main wheels till contact the runway by the front wheel, and run on all wheels. Sequential backward consideration of these stages and accounting all constraints for the state of aircraft allow [19, pp. 210–216] to construct the tolerance set at the instant ϑ. Some reasonable approximation of this set is taken as the terminal one in the considered problem.

Orientation of the terminal set M (see the last picture in Fig. 3) can be explained in the following way. If the lateral deviation from the runway central axis is positive at the instant ϑ, then the lateral velocity should be negative to keep the aircraft within the runway. And vice versa, when the deviation is negative, the velocity should be positive.

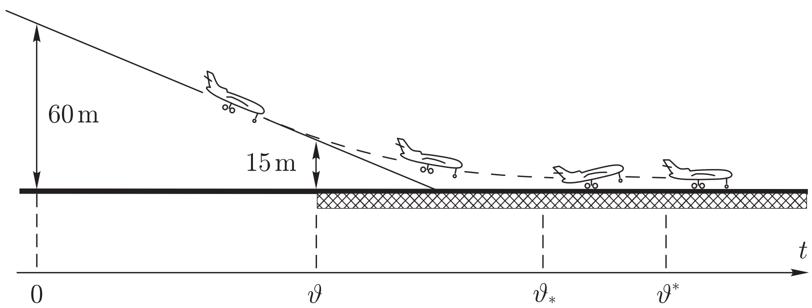


Figure 2: Four stages of landing: $[0, \vartheta]$ — the last stage of descending along the rectilinear glide path; $[\vartheta, \vartheta_*]$ — the stage of flare; $[\vartheta_*, \vartheta^*]$ — the run on the main wheels; $t \geqslant \vartheta^*$ — the run on all wheels.

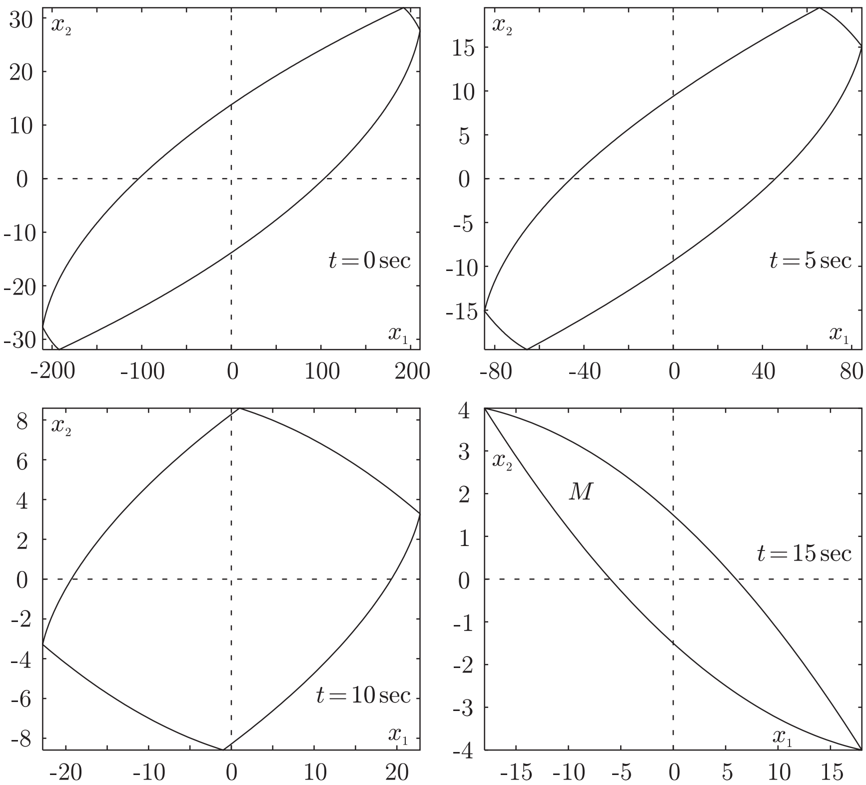


Figure 3: Sections $W(t)$ of the maximal stable bridge W at some instants, $W(15) = M$.

In examples below, $\vartheta = 15\,\mathrm{sec}$.

As the constraint $Q_{\max}$, let us take

$$|v| \le \nu = 10\,\mathrm{m/sec}.$$

The multiplier ρ, used in the definition of the stable bridge $\widehat{W}$ in formula (6), was taken equal to 57. The time step δ of the backward procedure constructing the time sections of the maximal stable bridges W and $\widehat{W}$ is equal to 0.05. With that, the set M was approximated by a convex polygon with 120 vertices.

In Figure 3, time sections of the maximal stable bridge W corresponding to parameters P, $Q_{\max}$, and M are drawn for instants $t = 0$, 5, 10, and 15 sec. In Figure 4, sections for instants $t = 5$ and 10 sec are shown together with the switching lines defining the sign of the first player's control. These figures use the coordinates x_1, x_2 of game (2).

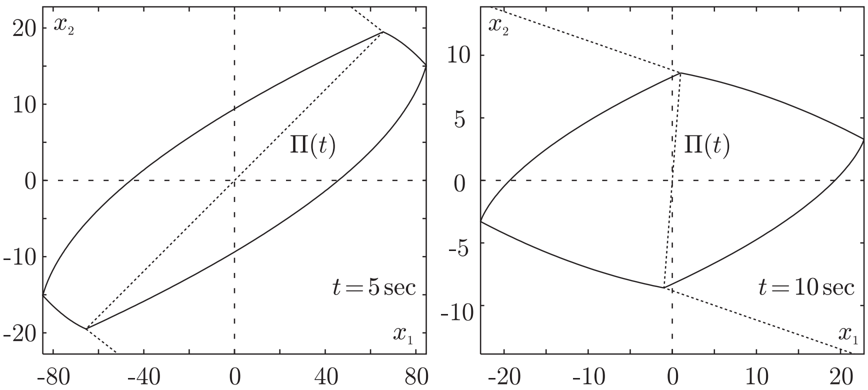


Figure 4: Sections $W(t)$ of the maximal stable bridge W shown together with switching lines $\Pi(t)$.

2.2 Simulation of Motions

It is supposed that the control u in system (8) is generated by the robust feedback law $\widetilde{U}(t, \mathbf{x})$ in a discrete scheme of control with some time step Δ.

To model a motion of the system, one should set the initial position and the second player's control.

In the following simulations, the initial point has lateral deviation equal to 30 m:

$$\mathbf{x}(0) = (30, 0, 0, 0, 0, 0, 0).$$

One can check that

$$x(0) = X_{2,7}(15, 0)\mathbf{x}(0) \in W(0).$$

Here, $X_{2,7}$ is the matrix for passage to a system of type (2). It consists of two first rows of the fundamental Cauchy matrix for system (8).

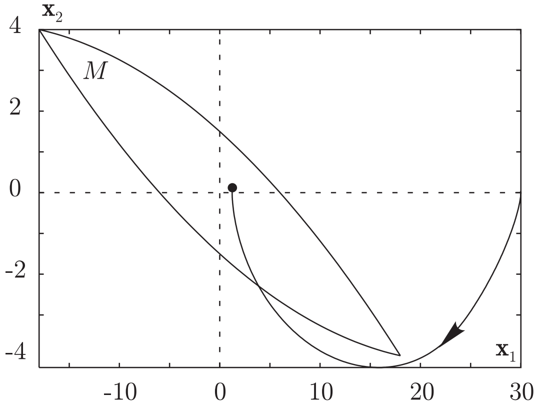
Simulations have been carried out for two types of the second player's control: as sinusoids having the same frequency and different amplitudes, and as an optimal strategy from an auxiliary differential game. This game has fixed terminal time, a given constraint for the second player's control, and a terminal payoff function of Minkowski type generated by the set M. Such an optimal strategy can be computed by means of algorithms from the works [6, 8, 39].

Simulation results are given for two levels of the disturbance:

- 5 m/sec — a level less than the chosen maximal one equal to 10 m/sec;
- 15 m/sec — a level greater than the chosen maximal one; so, in this case there is no guarantee of reaching the terminal set M.

So, in total, there are four variants of the disturbance. For these variants, Figures 5–8 show the trajectory in the plane of phase variables $\mathbf{x}_1$ and $\mathbf{x}_2$, the system position $\big(\mathbf{x}_1(\vartheta), \mathbf{x}_2(\vartheta)\big)$ at the terminal instant $\vartheta = 15$ sec, and the realizations of the players' controls.

Type of disturbance:
sinusoidal
Maximum of disturbance:
5 m/sec (0.5ν)
Termination:
successful
Maximum of control:
0.08 (0.31μ)



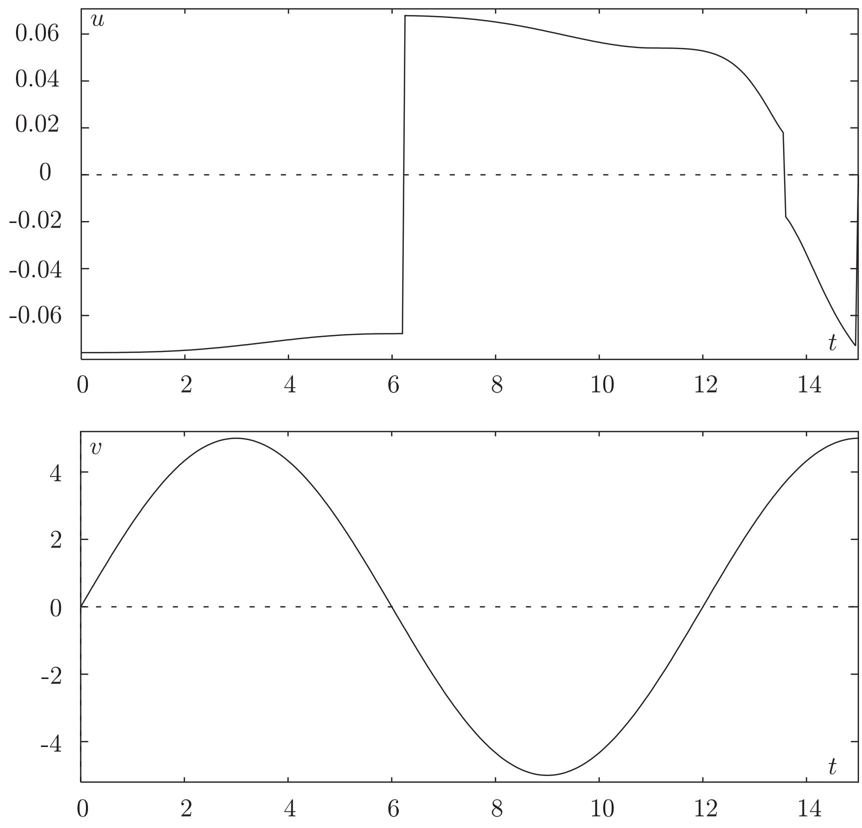


Figure 5: Trajectory of the system (in the coordinates of lateral deviation $\mathbf{x}_1$ and lateral velocity $\mathbf{x}_2$) and its state at the terminal instant (the circle in the upper figure); graphs of players' controls: first player (the middle figure) and second player (the lower figure).

Type of disturbance:
sinusoidal
Maximum of disturbance:
15 m/sec (1.5ν)
Termination:
successful
Maximum of control:
0.25 (0.96μ)

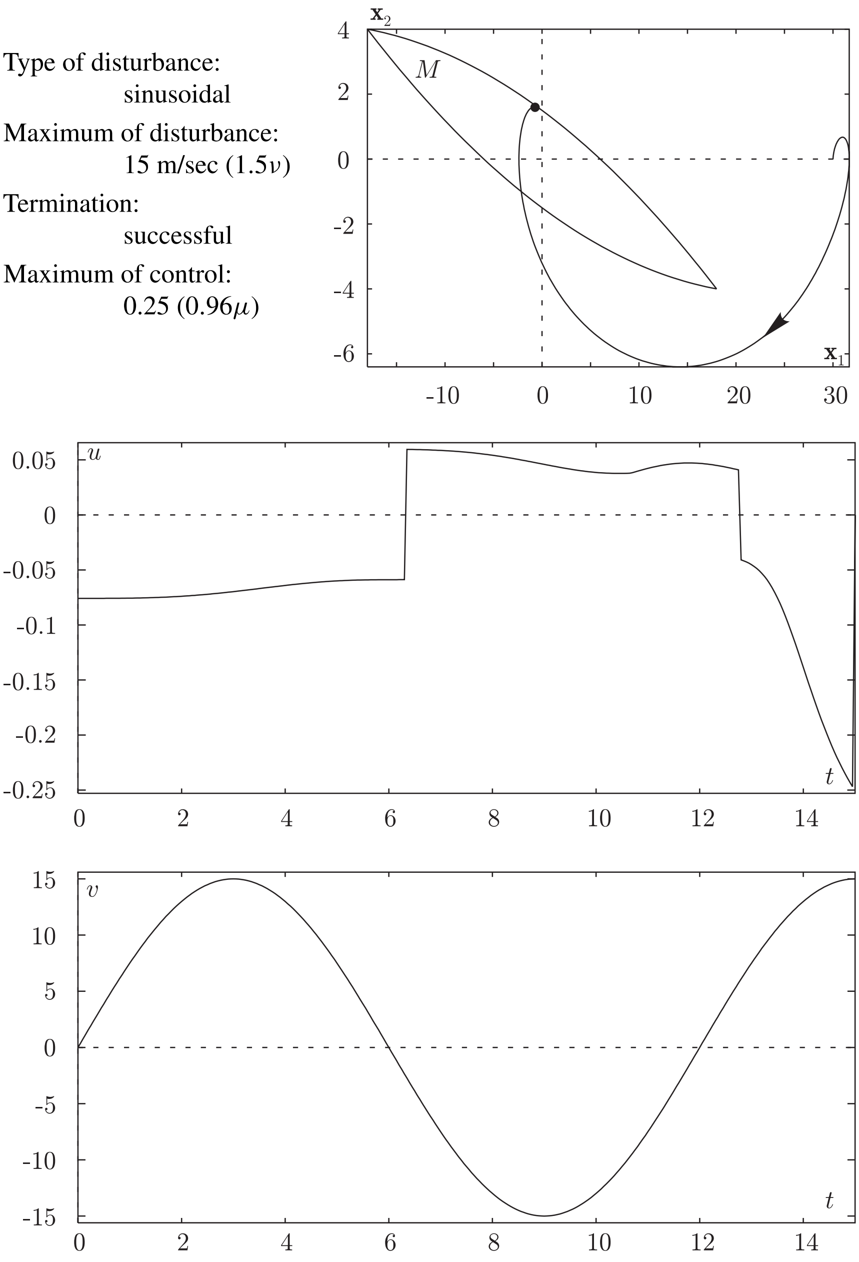


Figure 6: Trajectory of the system (in the coordinates of lateral deviation $\mathbf{x}_1$ and lateral velocity $\mathbf{x}_2$) and its state at the terminal instant (the circle in the upper figure); graphs of players' controls: first player (the middle figure) and second player (the lower figure).

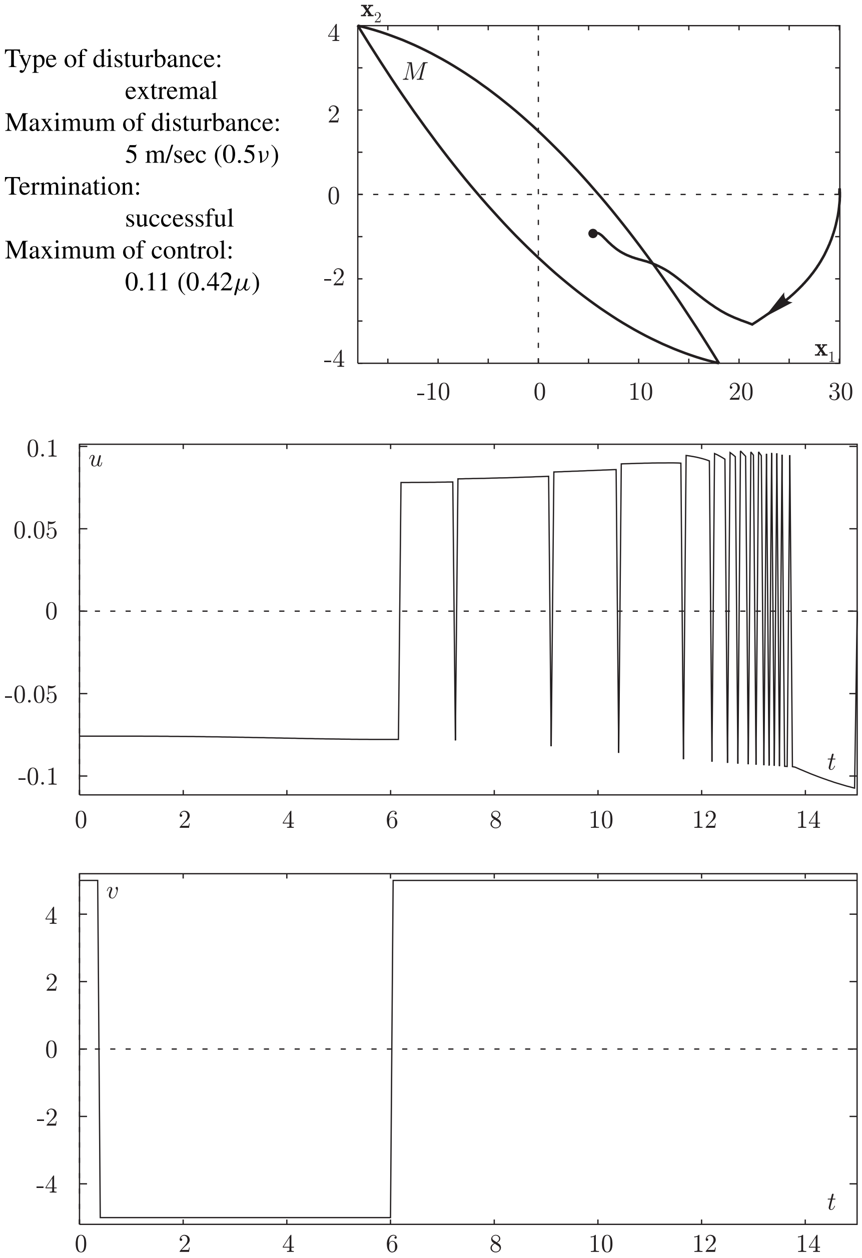


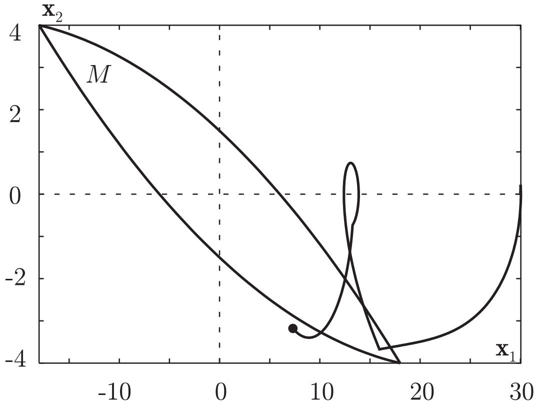
Figure 7: Trajectory of the system (in the coordinates of lateral deviation $\mathbf{x}_1$ and lateral velocity $\mathbf{x}_2$) and its state at the terminal instant (the circle in the upper figure); graphs of players' controls: first player (the middle figure) and second player (the lower figure).

Type of disturbance:
extremal

Maximum of disturbance:
15 m/sec (1.5ν)

Termination:
almost successful

Maximum of control:
0.26 (μ)



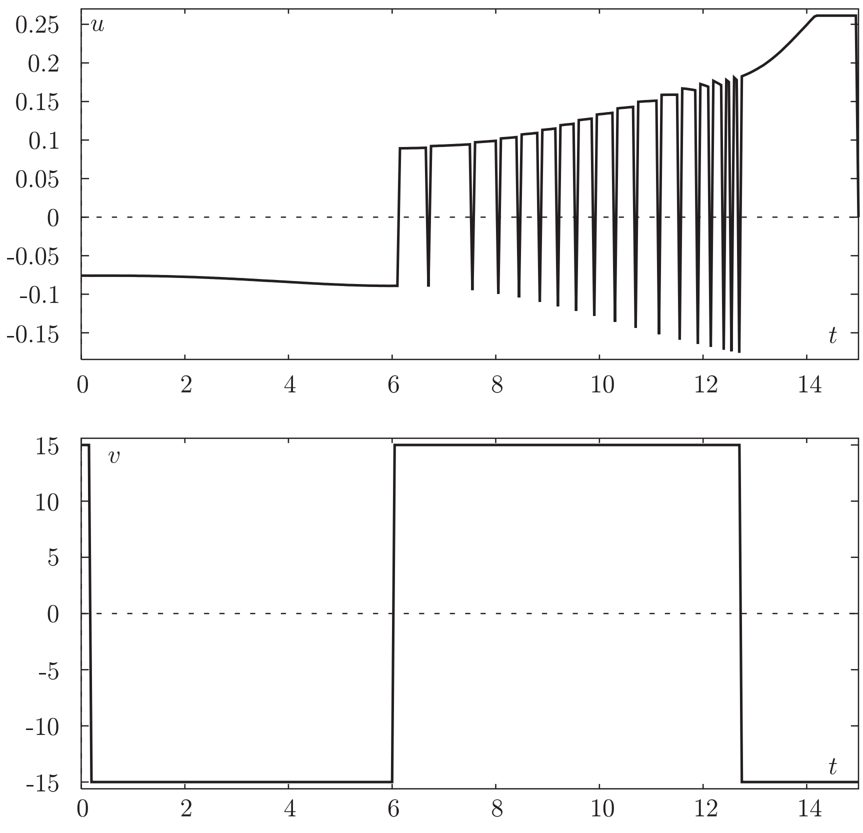


Figure 8: Trajectory of the system (in the coordinates of lateral deviation $\mathbf{x}_1$ and lateral velocity $\mathbf{x}_2$) and its state at the terminal instant (the circle in the upper figure); graphs of players' controls: first player (the middle figure) and second player (the lower figure).

Figures 5 and 6 involve the situation of sinusoidal disturbance, and Figures 7 and 8 correspond to the extremal disturbance generated by a feedback strategy taken from an auxiliary differential game.

When modeling motions, the time step Δ in the discrete scheme of control was taken equal to 0.05 sec.

2.3 Discussion of Simulation Results

Simulation results show that the suggested robust control successfully parries sinusoidal disturbance, including the case when it is stronger than the chosen maximal level. With that, when the disturbance level is less than the maximal one (i.e., for the variant 5 m/sec), the maximal value of the first player's control $u(t)$ is sufficiently less than the maximal level 0.2613 rad. In the case 15 m/sec, the extremal value of the first player's control is reached only in the final stage of the process close to the terminal instant $\vartheta = 15$ sec.

The results become worse when the sinusoidal disturbance is changed to extremal one. In this case, when the disturbance level is 15 m/sec, the phase coordinates $\mathbf{x}_1(15)$, $\mathbf{x}_2(15)$ at the terminal instant do not belong to the terminal set M (see Figure 8). However, the deviation from the terminal set is not too large. Again, the realization of the first player's control reaches its extremal value 0.2613 rad at the final stage of the process only.

For both variants with extremal second player's control, the realization of control u has frequent switches. This means that the phase trajectory in the space of system (2) goes near the switching surface and passes from one side to another. But since $u(t)$ is the required roll angle, which affects the angle of ailerons via some servomechanism, there is nothing bad in these switches. They are smoothed in the mechanism. In the simplest form, the inertiality of the servomechanism is taken into account in system (8) by the dynamics of passing from the control u to the roll angle $\mathbf{x}_5$ via the auxiliary variable $\mathbf{x}_7$.

For the variant of sinusoidal disturbance with level 5 m/sec (15 m/sec), the instant of entrance of the trajectory to the terminal set M in the plane $\mathbf{x}_1$, $\mathbf{x}_2$ equals 12.10 sec (12.95 sec). So, if the instant of passing the runway threshold due to some unaccounted reason becomes a bit less then the nominal one $\vartheta = 15$ sec, then the termination of the landing process can be treated as successful: the phase state in coordinates $\mathbf{x}_1$, $\mathbf{x}_2$ at the instant of passing is inside the tolerance. In the case of the extremal disturbance with level 5 m/sec, the instant of entrance to the set M is equal to 10.30 sec.

Conclusion

In this chapter, a conflict-controlled system with linear dynamics and fixed instant of termination is considered. The useful control is assumed to be scalar and bounded. A method for constructing nonlinear robust control is elaborated. The robust control is built on the basis of a special switching surface changing in time.

For the case when the quality of the process is defined by two components of the phase vector only, the algorithm is realized as a computer program. The elaborated method is applied to a problem of lateral motion control of an aircraft during landing under wind disturbance.

In the future, we plan to expand the suggested method for constructing robust control to the case of a linear-controlled system, where the useful control is a vector, whose components are independently constrained. The authors plan to include the corresponding algorithm into a software system that simulates control of an aircraft during landing in the framework of the full original nonlinear system [31].

Acknowledgment

This work is supported by the Russian Foundation for Basic Researches, projects Nos. 06-01-00414, 04-01-96099.

References

[1] Aubin, J.-P.: Viability Theory. Birkhauser, Basel (1991).

[2] Basar, T., Bernhard, P.: H^∞-Optimal Control and Related Minimax Design Problems. A Dynamic Game Approach. Birkhäuser, Boston (1991).

[3] Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhäuser, Boston (1997).

[4] Botkin, N.D.: Evaluation of numerical construction error in differential game with fixed terminal time. Problems of Control and Information Theory, **11**(4), 283–295 (1982).

[5] Botkin, N.D.: Numerical construction of cross-sections of positional absorption sets for linear differential games. In: Subbotin, A.I., Patsko, V.S. (Eds.) Algorithms and Programs for Solving Linear Differential Games, Institute of Mathematics and Mechanics, Sverdlovsk, 5–38 (1984) (in Russian).

[6] Botkin, N.D., Kein, V.M., Krasov, A.I., Patsko, V.S.: Control of Aircraft Lateral Motion During Landing in the Presence of Wind Disturbances. Report No. 81104592/02830078880. VNTI Center and Civil Aviation Academy, Leningrad, Russia (1983) (in Russian).

[7] Botkin, N.D., Ryazantseva, E.A.: An algorithm of constructing the sovability set for linear differential games of high dimension. Trudy Inst. Mat. i Mekh., **2**, 128–134 (1992) (in Russian).

[8] Botkin, N.D., Kein, V.M., Patsko V.S.: The model problem of controlling the lateral motion of an aircraft during landing. Journal of Applied Mathematics and Mechanics, **48**(4), 395–400 (1984).

[9] Botkin, N.D., Zarkh, M.A. Estimation of error for numerical constructing absorption set in linear differential game. In: Subbotin, A.I., Patsko, V.S. (Eds.) Algorithms and Programs for Solving Linear Differential Games, Institute of Mathematics and Mechanics, Sverdlovsk, 39–80 (1984) (in Russian).

[10] Bulirsch, R., Montrone, F., Pesch, H.J.: Abort landing in the presence of windshear as a minimax control problem, part 1: necessary conditions. Journal of Optimization Theory and Applications, **70**(1), 1–23 (1991).

[11] Botkin, N.D., Patsko, V.S.: Positional control in linear differential game. Engineering Cybernetics, **21**(4), 69–76 (1983).

[12] Cardaliaguet, P., Quincampoix, M., Saint-Pierre, P.: Set-valued numerical analysis for optimal control and differential games. In: Annals of the International Society of Dynamic Games, Vol. 4, Bardi, M., Raghavan, T.E., and Parthasarathy, T. (Eds.), 177–247 (1999).

[13] Dorato, P. (Ed.): Robust Control. IEEE Press, New York (1987).

[14] Fleming, W.H.: Future directions in control theory. A mathematical perspective. Report of the Panel on Future Directions in Control Theory, SIAM Reports on Issues in the Mathematical Sciences. Society for Industrial and Applied Mathematics, Philadelphia (1988).

[15] Ganebny, S.A., Kumkov, S.S., Patsko, V.S., Pyatko S.G.: Robust Control in Game Problems with Linear Dynamics, Preprint. Institute of Mathematics and Mechanics, Ekaterinburg, Russia, 53 p. (2005).

[16] Grigorenko, N.L., Kiselev, Yu.N., Lagunova, N.V., Silin, D.B., Trin'ko, N.G.: Methods for solving differential games. In: Mathematical Modelling, Moscow State University, Moscow, 296–316 (1993) (in Russian).

[17] Isakova, E.A., Logunova, G.V., Patsko, V.S.: Computation of stable bridges for linear differential games with fixed time of termination. In: Subbotin, A.I., Patsko, V.S. (Eds.) Algorithms and Programs for Solving Linear Differential Games, Institute of Mathematics and Mechanics, Sverdlovsk, 127–158 (1984) (in Russian).

[18] Isaacs, R.: Differential Games. John Wiley & Sons, New York (1965).

[19] Kein, V.M.: Optimization of Control Systems by Minimax Criterion. Nauka, Moscow (1985) (in Russian).

[20] Kein, V.M, Parikov, A.N., Smurov, M.Iu.: On means of optimal control by the extremal aiming method. Journal of Applied Mathematics and Mechanics, **44**(3), 306–310 (1980).

[21] Krasovskii, N.N., Subbotin, A.I.: Positional Differential Games. Nauka, Moscow (1974) (in Russian). Transl. as: Krassovski, N.; Soubbotine, A.: Jeux differentiels, Mir, Moscow (1977) (in French).

[22] Krasovskii, N.N., Subbotin, A.I.: Game-Theoretical Control Problems. Springer-Verlag, New York (1988).

[23] Kumkov, S.S., Patsko, V.S.: Construction of singular surfaces in linear differential games. In: Annals of the International Society of Dynamic Games, Vol. 6: Advances in Dynamic Games and Applications, Altman, E., Pourtallier, O., (Eds.), Birkhäuser, Boston, 185–202 (2001).

[24] Kurzhanski, A.B., Valyi, I.: Ellipsoidal Calculus for Estimation and Control. Birkhauser, Boston (1997).

[25] Leitmann, G., Pandey, S.: Aircraft control for flight in an uncertain environment: take-off in windshear. Journal of Optimization Theory and Applications, **70**(1), 25–55 (1991).

[26] Miele, A., Wang, T., Melvin, W.W.: Optimal take-off trajectories in the presence of windshear. Journal of Optimization Theory and Applications, **49**(1), 1–45 (1986).

[27] Miele, A., Wang, T., Wang, H., Melvin, W.W.: Optimal penetration landing trajectories in the presence of windshear. Journal of Optimization Theory and Applications, **57**(1), 1–40 (1988).

[28] Patsko, V.S.: Special aspects of convex hull constructing in linear differential games of small dimension. In: Control Applications of Optimization 1995. A Postprint Volume from the 10th IFAC Workshop, Haifa, Israel, 19–21 December 1995, Shinar, J., Kirillova, F.M. (Eds.), Pergamon, 19–24 (1996).

[29] Patsko, V.S.: Switching surfaces in linear differential games with a fixed instant of termination. Journal of Applied Mathematics and Mechanics, **68**, 583–595 (2004).

[30] Patsko, V.S.: Switching Surfaces in Linear Differential Games. Preprint, Institute of Mathematics and Mechanics, Ekaterinburg, Russia, 80 p. (2004) (in Russian).

[31] Patsko, V.S., Botkin, N.D., Kein, V.M., Turova, V.L., Zarkh, M.A.: Control of an aircraft landing in windshear. Journal of Optimization Theory and Applications, **83**(2), 237–267 (1994).

[32] Polovinkin, E.S., Ivanov, G.E., Balashov, M.V., Konstantiov, R.V., Khorev, A.V.: An algorithm for the numerical solution of linear differential games. Sbornik: Mathematics, **192**(10) 1515–1542 (2001).

[33] Pontryagin, L.S.: Linear differential games. 2. Soviet Math. Dokl., **8**, 910–912 (1967).

[34] Pschenichnyi, B.N., Sagaidak, M.I.: Differential games of prescribed duration. Cybernetics, **6**(2), 72–83 (1970).

[35] Seube, N., Moitie, R., Leitmann, G.: Aircraft taking-off in windshear: a viability approach. Set-Valued Analysis, **8**, 163–180 (2000).

[36] Tarasiev, A.M., Ushakov, V.N.: Construction of set collection approximating a maximal u-stable bridge. In: Subbotin, A.I., Patsko, V.S. (Eds.) Algorithms and Programs for Solving Linear Differential Games, Institute of Mathematics and Mechanics, Sverdlovsk, 159–190 (1984) (in Russian).

[37] Tarasiev, A.M., Uspenskii, A.A., Ushakov, V.N.: Approximate construction of positional absorption set in linear approach problem with convex target in R^3. In: Subbotin, A.I, Ushakov, V.N. (Eds.) Control in Dynamic Systems. Sverdlovsk, 93–100 (1990) (in Russian).

[38] Turetsky, V., Glizer, V.Y.: Robust state-feedback controllability of linear systems to a hyperplane in a class of bounded controls. Journal of Optimization Theory and Applications, **123**(3), 639–668 (2004).

[39] Zarkh, M.A.: A universal second-player strategy in a linear differential game. Journal of Applied Mathematics and Mechanics, **54**(3), 326–330 (1990).

[40] Zarkh, M.A., Ivanov, A.G.: Construction of the value function in the linear differential game with fixed terminal time. Trudy Inst. Mat. i Mekh., **2**, 140–155 (1992) (in Russian).

[41] Zarkh, M.A., Patsko, V.S.: Construction of maximal stable bridges in linear differential game. In: Subbotin, A.I., Kleimenov, A.F. (Eds.) Synthesis of Optimal Control in Game Systems, Ural Scientific Center of AS USSR, Sverdlovsk, 46–61 (1986) (in Russian).

[42] Zarkh, M.A., Patsko, V.S.: Numerical solution of a third-order directed game. Engineering Cybernetics, **26**(4), 92–99 (1988).

Games, Incompetence, and Training

Justin Beck
Defence Systems Analysis Division
Defence Science and Technology Organisation
P.O. Box 1500
Edinburgh, South Australia 5111, Australia
justin.beck@dsto.defence.gov.au

Jerzy A. Filar
Centre for Industrial and Applied Mathematics
University of South Australia
G.P.O. Box 2471
Adelaide, South Australia 5001, Australia
jerzy.filar@unisa.edu.au

Abstract

In classical noncooperative matrix games the payoffs are determined directly by the players' choice of strategies. In reality, however, a player may not be capable of executing his or her chosen strategy due to a lack of skill that we shall refer to as incompetence. A method for analysing incompetence in matrix games is introduced, examined and demonstrated. Along with the derivation of general characteristics, a number of interesting special behaviours are identified. These special behaviours are shown to be the result of particular forms of the game and/or of the incompetence matrices. The development of this simple model was motivated by applications where a decision to increase competence is to be evaluated. Investments that decrease incompetence (such as training) need to be related to the change in the value of the game. A game theory approach is used since possible changes in strategies used by other players must be considered. An analogy with the game of tennis is discussed, as is an analogy with capability investment decisions in the military.

1 Introduction

Game theory is appropriately used to examine situations of competition and conflict in a wide range of contexts. Even some popular games such as poker and tennis have been studied formally (e.g., see [1] and [2]). Note that one important aspect in which tennis is different from poker is that the action a player attempts to take is not necessarily the same as the outcome that is achieved. For example, although a player may choose to play a "down the line passing shot" the actual "outcome" might be to hit the ball into the net, which has a very different payoff. In such situations—where there is incompetence in the player's ability—it is natural to ask how the value of the game varies with the level of incompetence. Related to this issue is the problem of governments investing in the military. A military contest may be treated as a game, such as in references [3] and [4], and it is clear that the realisation of a battle plan may be very different from what the players intended, as a direct result of incompetence. The ability to invest in the military is a mechanism by which the government aims to increase the competence of its forces. This naturally leads to two questions: (i) How much to invest? and (ii) What type of competence to invest in? These two questions motivate the work in this chapter, which must be regarded as a preliminary and an exploratory study.

The usual interpretation of a classical two-person, zero-sum game is that a choice of an action (a pure strategy) by each of the players results in a deterministic payoff. Implicitly, there is an underlying assumption that players are capable of carrying out the actions that they have chosen. However, when incompetence is introduced this is no longer a valid assumption. We shall refer to such a game as the "incompetent game" and to the original as the "competent game."

In such a case, it is necessary to calculate an expected value across the possible "outcomes" (which may differ from intended actions). We introduce a method of doing this that maintains the separation between the parts of the game representing the game itself and those representing it when incompetence is incorporated. This is achieved by introducing "incompetence matrices" that capture the levels of competence of each of the players. The incompetent game is, ultimately, reduced to an equivalent matrix game by a suitable combination of the players' incompetence matrices and the payoff matrix of the original matrix game.

We note that player incompetence is intuitively related to the concept of a "trembling hand equilibrium," which was introduced by Selten [5] for extensive form games and has been studied by many authors. However, it is not clear that this concept lends itself to the explicit, strategic analysis of impacts of varying levels of competence that we have proposed here.

The chapter is divided into a number of sections. Section 2 describes the formulation of incompetent games and the necessary notation. Following this, in Section 3, there is a discussion of the general behaviour of the value of the game with changing incompetence. Additionally, there is a discussion of the question as to whether training is always beneficial.

A number of interesting properties of incompetent games were observed during empirical studies. These properties are shown to be the result of a special structure of the game and/or of the incompetence matrices (see Section 4). Nearly always, we illustrate the various properties with reference only to single player incompetence. Despite this, the reader will undoubtedly recognise that these behaviours naturally extend to the incompetence of both players. In Section 5 a tennis analogy of the model is discussed as an illustrative numerical example. It relates to a match of tennis where an incompetent player has the choice of two different training methods.

Section 6 discusses an extension to the basic model. This extension is motivated by more complex problems and is illustrated by an analogy to increasing military capability to achieve a better outcome. The application of the work to this domain will be the focus of further work on the subject. Finally, Section 7 provides a summary of the paper.

2 Defining Incompetent Matrix Games

2.1 Notation

Consider a two-player, zero-sum matrix game Γ such that there is a set of n^i actions for players 1 and 2 ($i = 1$ and $i = 2$). The superscript i will always be used to index the player. Let these actions be labelled by $A = \{a_1, a_2, \ldots, a_{n^1}\}$ and $B = \{b_1, b_2, \ldots, b_{n^2}\}$. The rewards $r(a_i, b_j)$ are determined by the actions selected by the players, in this case a_i and b_j. The payoff matrix R contains all of the $r(a_i, b_j)$ values. Each player can choose a mixed strategy (X and Y) which describes his intended probability distribution on the actions. The payoff of the game under strategies X and Y is given by $v(X, Y)$, and the value of the game when optimal strategies for the two players are used is given by $\bar{v}$.

Unlike the above matrix game, in the incompetent game (Γ_Q), the rows and columns of the payoff matrix are labelled by *outcomes* and not actions. The outcomes refer to the realizations of chosen actions, rather than the chosen actions themselves. Thus, an outcome is a function of the action selected, and the action selected is not necessarily the same as the outcome. Outcomes are labelled in a similar way to actions, using α and β instead of a and b. There are m^i actions for player i.

The probability of an outcome k when action j is selected by player i is denoted by $q^i(j, k)$, and all of these values are known to both players. For a single action (a_j) of player 1, the set of $q^1(j, k)\ \forall\, k$ forms a probability distribution $\mathbf{q}^1(a_j)$ over the set of $\alpha_j,\ j = 1, \ldots, m^1$. By construction the probability distribution $\mathbf{q}^1(a_j)$ satisfies $\sum_{k=1}^{m^1} q^1(j, k) = 1$ and $q^1(j, k)\ \geq 0,\ \forall\, k$. This is done similarly for all other actions. Analogous notation with the superscript $i = 2$ applies to player 2.

This leads to a formal definition of player i's *incompetence matrix* $\mathcal{Q}^i$ whose rows are the vectors $\mathbf{q}^i$, for $i = 1, 2$. The value of the incompetent game will be denoted by $\bar{v}_{\mathcal{Q}}$. Strictly speaking, the latter should have been written as $\bar{v}_{\mathcal{Q}^1,\mathcal{Q}^2}$ to indicate the dependence on the incompetence matrices of both players. However, $\bar{v}_{\mathcal{Q}}$ is a convenient shorthand notation in all cases where the differentiation between $\mathcal{Q}^1$ and $\mathcal{Q}^2$ is not essential.

In general, there need be no special relationship between a_j and α_j. Indeed, even the number of outcomes and actions need not be the same, that is, $m^i \neq n^i$ is permitted. However, for the purposes of this preliminary study we make the following restrictive assumption.

Assumption 2.1. There is a 1:1 relationship between actions and outcomes, hence $m^i = n^i$, for $i = 1, 2$. Additionally, the action and outcome sets have the special relationship that the level of incompetence is defined to be zero for player 1's action a_j when the probability of executing α_j is 1. That is,

$$q^1(j,k) = \begin{cases} 0 & j \neq k \\ 1 & j = k, \end{cases}$$

and similarly for player 2.

Two immediate consequences of this assumption will be that $\mathcal{Q}^i$ will be $n^i \times n^i$ and that the incompetent game will reduce to the standard matrix game when $\mathcal{Q}^i = I$, the identity matrix for $i = 1, 2$. More generally this may not be the case and $\mathcal{Q}^i$ will have dimension $m^i \times n^i$. Possible extensions of this sort are considered in Section 6. In all other sections, we will use the term actions to refer to both actions and outcomes since the assumption above makes the terms essentially interchangeable.

2.2 Varying the Level of Incompetence

The level of incompetence of player i is defined using a "starting" $\mathcal{S}^i$ and "final" $\mathcal{F}^i$ levels of incompetence. To capture the notion of training influencing the level of incompetence we must introduce a method of moving from $\mathcal{S}^i$ to $\mathcal{F}^i$. Hence, a trajectory joining these two points is defined. Player i's competence increases as the trajectory is traversed from $\mathcal{S}^i$ to $\mathcal{F}^i$. This curve is then parameterised by an *incompetence parameter* λ in the domain [0, 1], which describes the level of competence. In general, the curve need not be a straight line and $\mathcal{S}^i$ and $\mathcal{F}^i$ are arbitrary, as long as they obey the criteria of being incompetence matrices. In order to simplify the problem for our purposes here a linear trajectory has been chosen.[1] It is such that for a given λ the Q^i that corresponds to this λ is given by

$$\mathcal{Q}^i(\lambda) = (1-\lambda)\mathcal{S}^i + \lambda\mathcal{F}^i. \tag{1}$$

In this way λ defines the level of incompetence in the game. When λ is zero, the level of incompetence is at its maximum level; when it is one, it is at its minimum. In practice the value of λ could be increased through mechanisms such as training or the purchase of better equipment.

Hence, the game can be considered a function of the λ's for each player. Of course, with two players, we shall have two incompetence parameters: λ and μ for players 1 and 2, respectively. De facto, we shall now consider a parameterised family of incompetent games $\Gamma_{\mathcal{Q}}(\lambda, \mu)$. Correspondingly, all the notation developed for a single game can be extended to this parameterised family of games. For

[1] In some contexts, a transition representing skill acquisition of the form $1 - \nu\lambda^{-\gamma}$ [6] might be more appropriate.

example, the value of these games will be denoted by $\bar{v}_{\mathcal{Q}}(\lambda, \mu)$. In this chapter, if no confusion is possible, the explicit dependence on (λ, μ) will often be suppressed or reduced to single player form by using (λ). The superscript identifying a particular player will also be suppressed unless required.

It is worth mentioning here that when—for both players—$\mathcal{F}^i$ is the identity matrix (I) and the level of competence is one, the game will revert to the fully competent and standard matrix game. Thus, the value of the original matrix game Γ would in this case coincide with $\bar{v}_{\mathcal{Q}}(1, 1)$. Henceforth, *complete competence* for player 1 refers to the case $\mathcal{F}^1 = I$ and $\lambda = 1$ and similarly for player 2.

It is possible to model the players' levels of incompetence without a reference to a starting and/or finishing point. However, it will be seen in subsequent sections that these "boundary" incompetence matrices enable us to conveniently capture what might be best called "structured incompetence." The latter shows that some actions can be easier to master than others.

2.3 The Reduced Normal Form of the Incompetent Game $\Gamma_{\mathcal{Q}}$

It is possible to transform the incompetent game $\Gamma_{\mathcal{Q}}$ into a matrix game in the classical sense. While this new game will depend on λ and/or μ, standard methods of finding solutions will be available. As discussed in Section 2.1, for a given $\mathcal{Q}$ the probability of player i obtaining outcome k given the action j is selected is simply $q^i(j, k)$. If the two players select actions a_i and b_j the probability that they will execute outcomes α_k and β_h is given by

$$P(\alpha_k, \beta_h | a_i, b_j) = q^1(a_i, \alpha_k) q^2(b_j, \beta_h).$$

This would result in a reward of $r(\alpha_k, \beta_h)$. The expected reward if the players select these actions is therefore given by

$$\bar{r}(a_i, b_j) = \sum_{k=1}^{n^1} \sum_{h=1}^{n^2} q^1(a_i, \alpha_k) q^2(b_j, \beta_h) r(\alpha_k, \beta_h).$$

Let the values $\bar{r}(i, j)$ define the entries of the matrix $R_{\mathcal{Q}}$. Hence, in matrix form, we obtain

$$R_{\mathcal{Q}} = \mathcal{Q}^1 R (\mathcal{Q}^2)^T. \tag{2}$$

From now on, $R_{\mathcal{Q}}$ will define the matrix game that is an equivalent, reduced, normal form of the incompetent game $\Gamma_{\mathcal{Q}}$. Of course, a minimax solution (e.g., see [7]) of the matrix game $R_{\mathcal{Q}}$ will be taken as a solution of the incompetent game $\Gamma_{\mathcal{Q}}$. In order to examine how the value of the incompetent matrix game $\bar{v}_{\mathcal{Q}}(\lambda, \mu)$ varies with the incompetence parameters, it is therefore possible to solve a sequence of equivalent matrix games.

3 General Properties of $\bar{v}_Q$

This section deals mostly with general properties of $\bar{v}_Q$ in a game where only player 1 is incompetent. However, these properties can be extended to the case where both players are incompetent. In general, no special properties of $\bar{v}_Q$ seem intuitively "obvious," with the possible exception that training should yield a benefit. However, even this will be seen to be false in some examples.

First, consider the issue of continuity of $\bar{v}_Q$. Using (1) and (2), and recalling that $\mathcal{Q}^2 = I$, the payoff matrix at λ will be given by

$$R_Q = ((1-\lambda)\mathcal{S} + \lambda\mathcal{F})R.$$

Hence, R_Q is continuous (indeed, linear) in λ and so also its value $\bar{v}_Q$ (e.g., see [8], pp. 346–347).

Lemma 3.1. *The dependence of $\bar{v}_Q$ on λ will be of the form*

$$\frac{c_1\lambda^k + c_2\lambda^{k-1} + \cdots + c_{k+1}}{d_1\lambda^{k-1} + d_2\lambda^{k-2} + \cdots + d_k}, \tag{3}$$

where k is the dimension of the kernel of R_Q and the c's and d's are arbitrary constants.

Proof. Consider a kernel of the matrix game R_Q in the sense of Shapley and Snow [9]. This kernel contains only the rows and columns that contribute to the given basic solution of the game. Shapley and Snow [9] show that the value of the game is then given by

$$v(\lambda) = \frac{1}{J(R_Q^K)^{-1}J^T}, \tag{4}$$

where J represents a vector of 1's. In this proof the form

$$v(\lambda) = \frac{|R_Q^K|}{J(R_Q^K)^* J^T}, \tag{5}$$

where * represents the adjoint of the matrix, will be used.

A direct result of (1) and (2) is that each element of R_Q is either a constant term or linear in λ. As a result the numerator of (5) will be of the general form $c_1\lambda^k + c_2\lambda^{k-1} + \cdots + c_{k+1}$ where the c_i's represent constants because of the way the determinant is defined.

Note now that the denominator is the sum of the elements of the adjoint. Because of the properties of each element of R_Q, each element of $(R_Q^K)^*$ is of the form $b_1\lambda^{k-1} + b_2\lambda^{k-2} + \cdots + b_k$, and hence the sum of these elements is also of the form $d_1\lambda^{k-1} + d_2\lambda^{k-2} + \cdots + d_k$, noting that the constants may be different.

Combining these two results yields the required general form of $\bar{v}_Q$. □

The other general property of $\bar{v}_Q$ that will be examined is the intuitive (naive?) expectation that $\bar{v}_Q$ should be monotonically nondecreasing in λ. However, it is clearly possible to interchange $\mathcal{S}$ and $\mathcal{F}$ to produce a negative training effect.

In addition to this, it is possible to construct examples where $\bar{v}_Q$ both increases and decreases for different $\lambda \in [0, 1]$. Such an example is defined by the simple game

$$R = \begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix}.$$

Define the incompetence of player 1 by

$$\mathcal{S} = \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0 & 0.3 & 0.7 \\ 0 & 0 & 1 \end{pmatrix}; \mathcal{F} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that $\mathcal{F}$ represents complete competence. Using (2) it is easy to calculate that

$$R_Q = \begin{pmatrix} 0.2 - 0.2\lambda \\ -0.9 + 4.9\lambda \\ -3 \end{pmatrix}.$$

A plot of $\bar{v}_Q$ as a function of λ is shown in Figure 1. The game was solved using standard linear programming techniques. It is trivially shown that $\bar{v}_Q$ decreases for $\lambda \in [0, \frac{11}{51})$ and increases for $\lambda \in (\frac{11}{51}, 1]$.

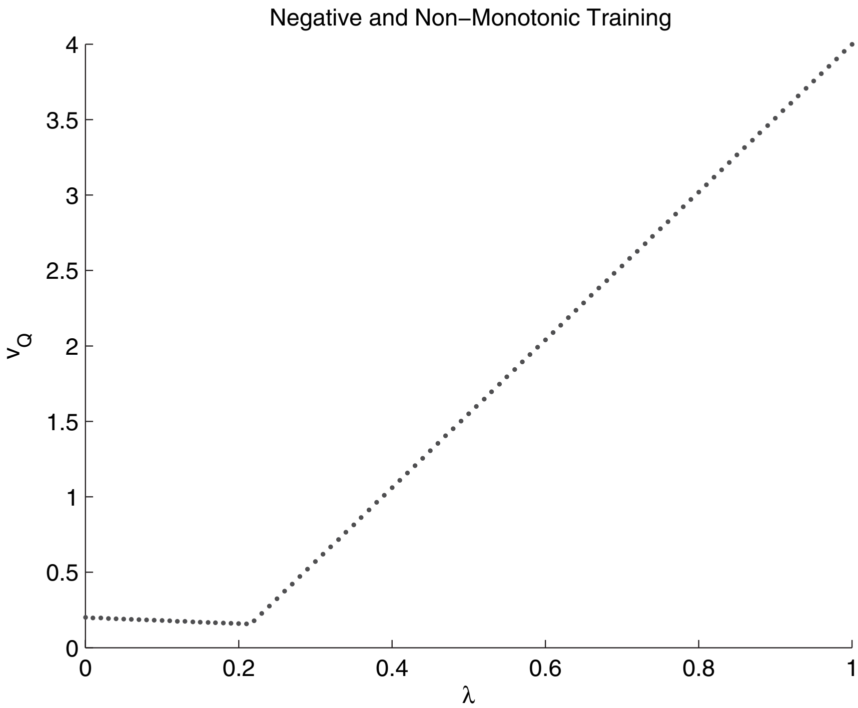


Figure 1: The value of the game decreases from $\lambda = 0$ to $\lambda \approx 0.22$, despite a decrease in the incompetence level and "uniform training."

4 Other Structured Cases

A number of cases—possessing interesting special structures—have been identified. In these cases the behaviour of $\bar{v}_{\mathcal{Q}}(\lambda)$ is much simpler than the general case described in Lemma 3.1. An illustrative example is presented in Section 4.5.

4.1 Uniform Incompetence to Complete Competence

The concept of a uniformly incompetent player is now introduced. Such a player has a $\mathcal{Q}$ matrix where every element has the value $\frac{1}{n}$, recalling that n is the number of actions.

Lemma 4.1. *If a single player increases his or her competence level from uniform incompetence to complete competence, the change in the value with λ is piecewise linear.*

Proof. First, consider the numerator of (5) which gives the value of the game. Equation (7) in [10] shows that the determinant can be written

$$\det R_{\mathcal{Q}}^K = \det(\lambda(I^K - S^K)R^K) + \det(\tilde{r}, \tilde{r}_2 - \tilde{r}_1, \ldots, \tilde{r}_n - \tilde{r}_1),$$

where $\tilde{r}$ is any row of $S^K R^K$ and $\tilde{r}_i$ are the rows of $\lambda(I^K - S^K)R^K$. Now the term

$$\det(\lambda(I^K - S^K)R^K) = \lambda^K \det((I^K - S^K)R^K)$$

and therefore only has terms of the form $c\lambda^K$. The term $\det(\tilde{r}, \tilde{r}_2 - \tilde{r}_1, \ldots, \tilde{r}_n - \tilde{r}_1)$ has only a $c\lambda$ appearing in every element except in the first row where there is no λ dependence. Hence, this term only has terms of the form $c\lambda^{K-1}$. Thus, the numerator of (5) is of the form $c_1\lambda^K + c_2\lambda^{K-1}$.

Note that the matrix $R_{\mathcal{Q}}^K$ is of the form described in [10] (Lemma 2.1) except that rows are interchanged with columns. This can be seen by recalling (1) and hence determining that

$$R_{\mathcal{Q}} = \mathcal{Q}R = \lambda(I - S)R + SR,$$

from (2). Now $R_{\mathcal{Q}}^K$ is constructed by taking

$$R_{\mathcal{Q}}^K = \mathcal{Q}^K R^K = \lambda(I^K - S^K)R^K + S^K R^K,$$

noting that the superscript K refers to the selection of rows and columns corresponding to the kernel of the matrix game $R_{\mathcal{Q}}$, not of R. Now, since in the initial (uniform) incompetence matrix S every element is $1/n$, all elements of S and S^K are identical. Hence, each column of $S^K R^K$ will contain identical elements, that is, any column is a scalar multiple of the vector of $1's$. Hence, the matrix $R_{\mathcal{Q}}^K$ is of the form

$$R_{\mathcal{Q}}^K = (a_{ij} + b_j)_{i,j=1}^n,$$

as required in [10] (Lemma 2.1).

Now, consider the denominator of (5). Here we note that

$$JA = \sum_i A_{ij},$$

and hence [10] (Lemma 2.1 (ii)) can be applied. Thus,

$$J(R_Q^K)^* = J(\lambda(I^K - S^K)R^K)^*.$$

Now, note that $(cA)^* = c^{n-1}A^*$ where n is the dimension of A. Hence, the denominator of (5) contains only terms of the form $c\lambda^{k-1}$, where $k \times k$ are the dimensions of R_Q^K. Thus, (5) is of the form

$$v(\lambda) = \frac{c_1\lambda^k + c_2\lambda^{k-1}}{c_3\lambda^{k-1}} = c_4\lambda + c_5$$

and is linear in λ. The piecewise linearity of the value stems from the fact that the kernels of R_Q^K could change as λ varies from 0 to 1. □

4.2 Completely Mixed R and $\bar{R}_Q$

In the case where both R and $R_Q = Q^1 R(Q^2)^T$ are completely matrix games there is no loss of generality in assuming that both R^{-1} and R_Q^{-1} exist. In such a case it follows that both Q^1 and Q^2 are also invertible.

Lemma 4.2. *If R and $\bar{R}_Q$ are completely mixed, then the value of the game Γ_Q is the same as the value of the game Γ.*

Proof. When R is completely mixed, the kernel of R is the entire matrix, that is $R^K = R$. Using this in (4) results in

$$\bar{v} = \frac{1}{JR^{-1}J^T}. \tag{6}$$

Similarly, the value of the game Γ_Q when it is completely mixed is given by

$$\bar{v}_Q = \frac{1}{JR_Q^{-1}J^T}.$$

Substituting for R_Q from (2) gives

$$\bar{v}_Q = \frac{1}{J(Q^1RQ^2)^{-1}J^T}.$$

This can be rearranged to be

$$\bar{v}_Q = \frac{1}{J((Q^2)^T)^{-1}R^{-1}(Q^1)^{-1}J^T}. \tag{7}$$

Now, since the rows of Q^1 are probability distributions we have $Q^1 J^T = J^T$ and hence $J^T = (Q^1)^{-1} J^T$, and similarly for player 2. Using these results in (7) and recalling (6) we obtain

$$\bar{v}_Q = \frac{1}{JR^{-1}J^T} = \bar{v}$$

as required. □

In general, even if R satisfies the criteria of being square and having no dominated rows or columns, this is no guarantee that $\bar{R}_Q$ has the same properties for all λ. Therefore, this theorem would seem to have limited applicability. However, if $\mathcal{F} = I$ and as λ approaches 1, it might be expected that $\bar{R}_Q$ does have the required properties, and therefore the value of the game would be constant with changing λ.

4.3 Saddlepoints in $\bar{R}_Q$

In general, if there is a saddlepoint in the matrix game $R_{\mathcal{Q}}$, then the dependence of $\bar{v}_{\mathcal{Q}}$ on λ will be (at most) a linear function of λ.

This is easily shown by considering the general form of $\bar{v}_{\mathcal{Q}}$ as given by Lemma 3.1. When the size of the kernel of $R_{\mathcal{Q}}$ is one, as it is when there are pure optimal strategies, the general form will be

$$\frac{c_1\lambda + c_2}{d_1},$$

which is clearly linear in λ. The exact dependence will be determined by the element of $R_{\mathcal{Q}}$ that corresponds to the solution.

4.4 Optimality of Complete Competence

The notation $\bar{v}_I$ is introduced to represent the value of the game when the incompetent player has complete competence. That is, $\bar{v} = \bar{v}_I = \bar{v}_{I,I}$. For the maximising player, we are interested in comparing the value of this game $\bar{v}_I$ with $\bar{v}_{\mathcal{Q}} = \bar{v}_{\mathcal{Q}^1,I}\ \forall\ \mathcal{Q}^1$. Analogously, for the minimising player, we wish to compare $\bar{v}_I$ with $\bar{v}_{\mathcal{Q}} = \bar{v}_{I,\mathcal{Q}^2}\ \forall\ \mathcal{Q}^2$.

It is easy to check that the intuitively obvious inequalities

$$\bar{v}_I \geq \bar{v}_{\mathcal{Q}^1,I} \ \ \forall \mathcal{Q}^1 \quad \text{and} \quad \bar{v}_I \leq \bar{v}_{I,\mathcal{Q}^2} \ \ \forall \mathcal{Q}^2$$

both hold. We refer to these inequalities as "optimality of complete competence." For instance, to see that the first of these inequalities is valid, let Y^0 be optimal for player 2 in the matrix game R. Then it is clear that $RY^0 \leq \bar{v}_I J^T$ and hence (using the fact that rows of $\mathcal{Q}^1$ sum to 1) that

$$\mathcal{Q}^1 RY^0 \leq \bar{v}_I J^T.$$

Now, by multiplying the above, on the left, by X^*—an optimal strategy of player 1 in the matrix game $\mathcal{Q}^1 R$—we immediately obtain the desired inequality.

Remark. It should be clear that, in the preceding discussion, there is no need for the assumption that one of the players is completely competent, as long as it is assumed that this player's incompetence matrix $\mathcal{Q}^1$ or $\mathcal{Q}^2$ is held fixed while we are examining the effects of varying the incompetence matrix of the other player.

Now it is easily seen that the preceding optimality of complete competence inequalities can be generalised as follows:

$$\bar{v}_{\mathcal{Q}}(1, \mu) \geq \bar{v}_{\mathcal{Q}}(\lambda, \mu) \geq \bar{v}_{\mathcal{Q}}(\lambda, 1);$$
$$\bar{v}_{\mathcal{Q}}(1, \mu) \geq \bar{v}_{\mathcal{Q}}(1, 1) \geq \bar{v}_{\mathcal{Q}}(\lambda, 1),$$

for all $\lambda, \mu \in [0, 1]$.

Remark. These relationships suggest that $\bar{v}_{\mathcal{Q}}(1, 1)$ may be a useful reference point. However, it is not clear whether any useful, general, inequality relations exist between $\bar{v}_{\mathcal{Q}}(\lambda, \mu)$, $\bar{v}_{\mathcal{Q}}(0, 0)$ and $\bar{v}_{\mathcal{Q}}(1, 1)$.

4.5 Illustrative Example

The following example illustrates three of the structured special cases described above; namely, the linear behaviour described in Section 4.3, the constant behaviour of Lemma 4.2 and the optimality of complete competence. Consider the payoff matrix defined by

$$R = \begin{pmatrix} 5 & 0 \\ 1 & 8 \end{pmatrix}.$$

Now, assume that player 2 is completely competent and let the initial and final $\mathcal{Q}^1$, for player 1, be defined by

$$\mathcal{S} = \begin{pmatrix} 0.5 & 0.5 \\ 0 & 1 \end{pmatrix}; \mathcal{F} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Using (1) and (2) the $R_{\mathcal{Q}}$ can be calculated to be

$$R_{\mathcal{Q}} = \begin{pmatrix} 3 + 2\lambda & 4 - 4\lambda \\ 1 & 8 \end{pmatrix}.$$

It is easily shown that the matrices R and $R_{\mathcal{Q}}$ satisfy the conditions of Lemma 4.2 for $\lambda \in (\frac{1}{6}, 1]$. The value of the game therefore remains constant in this domain. For $\lambda \in [0, \frac{1}{6}]$ the first column dominates the second and the game has a saddlepoint. The value of the game in this domain is therefore linear; it is given by the entry in $R_{\mathcal{Q}}$ that corresponds to the saddlepoint, in this case, $3 + 2\lambda$. The value of the game as a function of λ is plotted in Figure 2. Note that, in this case, the maximal value

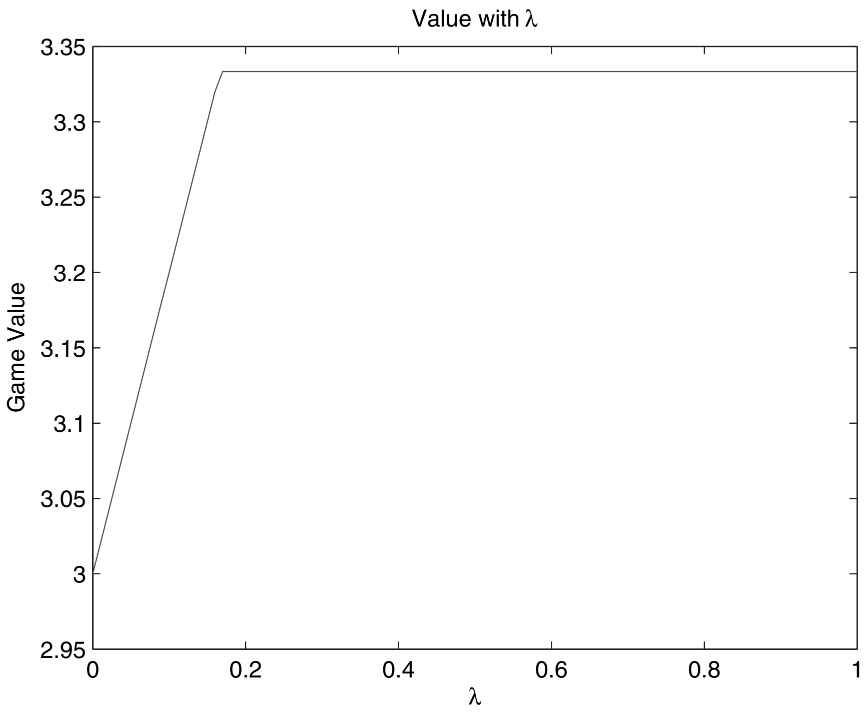


Figure 2: The value of the game changes with λ. The value is constant in the region $\lambda \in (\frac{1}{6}, 1]$ and is simply $3 + 2\lambda$ in the region $\lambda \in [0, \frac{1}{6}]$.

is attained at a level of competence well below complete competence. Of course, at complete competence ($\lambda = 1$) this value is also achieved. However, in practical situations, such a phenomenon raises the question of whether any expense aimed at increasing competence above the level $\lambda = \frac{1}{6}$ is justified.

5 Analogy in Tennis

This section describes an analogy of the theory developed so far. It creates a simple model and then evaluates which of two types of training schemes should be undertaken to achieve maximum benefit.

Consider a match of tennis that can be approximated with three basic actions. First, we have a "good shot" action, which describes a player hitting balls near the lines in some manner that is difficult to return but, presumably, with greater risk. The second action is called "safe," it describes shots where the ball is hit back into play, but into a safe area where it is easier to return. The final action is labelled "out," which results in the loss of the game point. Consider the reward units to be in terms of the percentage of matches won by player 1. For such a game one may expect the reward matrix to look something like

$$R = \begin{pmatrix} 50 & 70 & 100 \\ 30 & 50 & 100 \\ 0 & 0 & 50 \end{pmatrix},$$

where the following criteria have been used to determine the rewards:

- If both players use the same action the reward is 50;
- If one player uses the "out" action against a "good shot" or "safe," then this action always loses;
- If a player uses a "good shot" against a "safe" action, then he or she wins 70% of the time.

If both players are completely competent, then it is easy to see that a saddlepoint exists for this game when both players play the "good shot" strategy. However, even in this simple model, not all tennis players are completely competent. Incompetence for both players is introduced. Initial incompetence is assumed to be the same for both players and is described by

$$\mathcal{S} = \mathcal{S}^1 = \mathcal{S}^2 = \begin{pmatrix} 0.3 & 0.1 & 0.6 \\ 0.1 & 0.6 & 0.3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using (1) and (2), R_Q can be found to be

$$R_Q = \begin{pmatrix} 50 & 38.4 & 70 \\ 61.6 & 50 & 85 \\ 30 & 15 & 50 \end{pmatrix}.$$

This game also has a saddlepoint; however, in this case the saddlepoint corresponds to both players playing the "safe" strategy. This model suggests that incompetent tennis players tend to be more successful when playing a more conservative strategy, a phenomenon personally experienced by the authors.

Having created a model, a typical problem will now be stated in order to demonstrate its utility. The problem relates to a decision about what type of training to receive, and the following assumptions are made:

- Only one player is to receive training;
- Two types of training are available;
- Both types of training cost the same amount per unit λ;
- Training cost is not considered as part of this analysis;
- Both players will know the level of training that has been undertaken and its effects.

The two types of training correspond to two different final incompetence levels and hence different levels for all λ other than $\lambda = 0$. Let "advanced" training

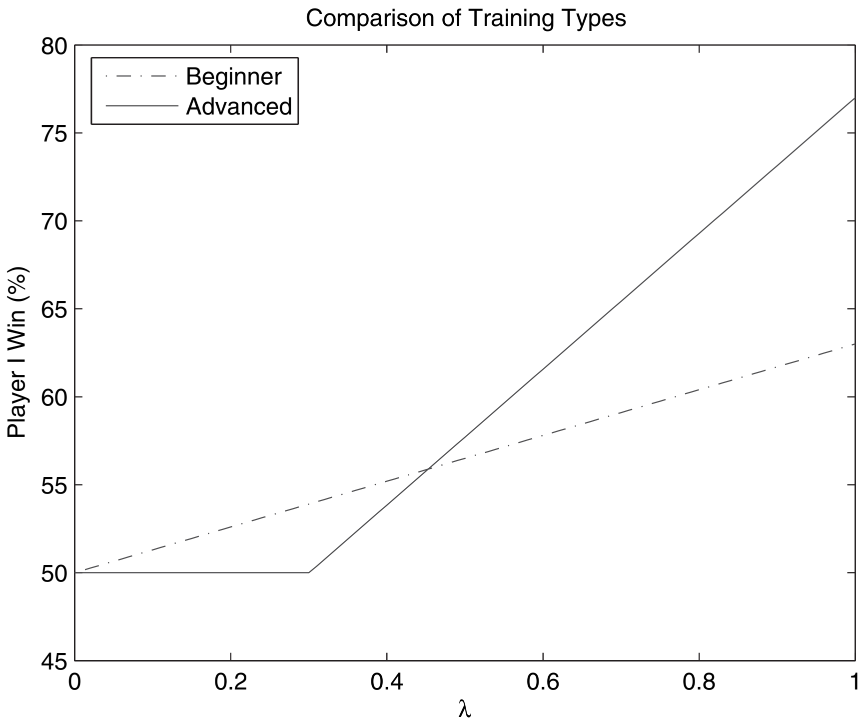


Figure 3: A comparison of two tennis training methods.

correspond to training in the "good shot" strategy and hence

$$\mathcal{F}_{adv} = \begin{pmatrix} 1 & 0 & 0 \\ 0.1 & 0.6 & 0.3 \\ 0 & 0 & 1 \end{pmatrix},$$

and "beginner" training corresponds to training in the "safe" strategy and hence

$$\mathcal{F}_{beg} = \begin{pmatrix} 0.3 & 0.1 & 0.6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The change in $\bar{v}_Q$ with λ for both of the training strategies is shown in Figure 3. From this figure it can be seen that the type of training that should be taken depends on how much training is to be undertaken. If less than $\lambda \approx 0.45$ training is to be taken, then the "beginner" training should be taken, otherwise the "advanced" training should be used.

A more complicated model would introduce a cost for the training and allow both players to train. If training were free, then it is clear that both players would choose to train fully in "advanced" training. However, when the cost of the training is comparable to the benefits of training, more interesting behaviour is expected. The cost functions would cause the game to become nonzero sum, and it would be possible to have multiple equilibria resulting in different rewards to the players. The latter is a subject of continuing research.

6 Model Extension

In this section we relax Assumption 2.1. We consider an analogy to capability acquisition in the military, a topic on which further work is planned.

In particular, we consider the case where the actions and outcomes need not have a 1:1 correspondence, and in fact the number of actions and outcomes are different. Such an extension may be required to build more realistic models. We consider that actions are a set of possible selections. The outcomes are the stochastic consequences of these actions without reference to the other player. The R matrix will now represent the game when the outcomes of the players are considered. The matrix $R_{\mathcal{Q}}$ represents the game when the actions are considered.

If the number of actions is large, an appropriate rectangular $\mathcal{Q}$ could be used to map the actions to a much smaller set of outcomes. Such an approach could be used to reduce the dimensionality of certain problems.

6.1 Analogy to Capability Acquisition in the Military

Decreasing incompetence will not always occur through training. In military applications incompetence may be decreased through the acquisition or upgrading of military hardware. The acquisition of military hardware is not cheap. Most governments spend significant amounts of money on their military. With such large amounts being invested, those making the spending decisions wish to decrease their incompetence in the best way possible.

Consider the problem of one government (player 2) trying to decrease its incompetence versus a completely competent foe. The foe (player 1) may attack using either a conventional or an insurgency strategy. Player 2 can defend with either a counter-conventional or a counter-insurgency strategy. Let the outcomes for player 1 be conventional attack, insurgency attack and for player 2 good conventional defence, good insurgency defence, bad conventional defence, bad insurgency defence. Let the reward matrix of the competent game be

$$R = \begin{pmatrix} 0 & 8 & 6 & 15 \\ 2 & 0 & 10 & 4 \end{pmatrix}$$

so that the good outcomes for player 2 clearly dominate the bad ones. Let the initial incompetence for player 2 be defined by

$$\mathcal{S} = \mathcal{S}^2 = \begin{pmatrix} 0.8 & 0 & 0.2 & 0 \\ 0 & 0.2 & 0 & 0.8 \end{pmatrix}.$$

This indicates that an attempted conventional defence results in a good conventional defence 80% of the time, while an attempted insurgency defence results in a good insurgency defence only 20% of the time. Since we are assuming that player 1 is competent, consistent with equations (1) and (2) we have that the initial incompetent game has the reward matrix

$$R_{\mathcal{Q}} = R\mathcal{S}^T = \begin{pmatrix} 1.2 & 13.6 \\ 3.6 & 3.2 \end{pmatrix}.$$

Now, let there be a large number of possible final incompetence matrices $\mathcal{F}$' for player 2. This is intended to capture policy decisions that may choose to devote resources either to defence capabilities targeting conventional warfare, or to capabilities targeting insurgency warfare. Since defence budgets can be split in many ways between these options, we parameterise $\mathcal{F}$ with a parameter $t \in [0, 0.2]$, as follows:

$$\mathcal{F}(t) = \mathcal{F}^2(t) = \begin{pmatrix} 0.8+t & 0 & 0.2-t & 0 \\ 0 & 1-4t & 0 & 4t \end{pmatrix}.$$

When $t = 0$ the money is spent entirely on counter-insurgency defence and when $t = 0.2$ it is spent entirely on counter-conventional defence. Note that if all the investment is spent on either single action, then player 2 will become completely competent in that action. This means that it is possible to increase competence in counter-insurgency defence four times more cheaply in terms of t. Now, no matter what value of the parameter t is selected, the reward matrix in the incompetent game corresponding to the incompetence level μ for player 2 can now be calculated according to

$$R_{\mathcal{Q}}(t) = R[(1-\mu)\mathcal{S} + \mu\mathcal{F}(t)]^T.$$

It is interesting to consider the benefits (if any) of eliminating the incompetence of player 2, that is, setting $\mu = 1$ above. In such a case the preceding parameterised reward matrix reduces to

$$R_{\mathcal{Q}}(t) = R\mathcal{F}^T(t) = \begin{pmatrix} 1.2-6t & 8+28t \\ 3.6-8t & 16t \end{pmatrix}.$$

The value of the preceding matrix game can now be plotted as a function of t in order to determine the capability investment decision that is most worthwhile. The value of the original game R for a completely competent player 2 can be easily calculated to be $\frac{8}{5}$. For the initial incompetence described above ($\mu = 0$) the value of the game $R\mathcal{S}^T$ is 3.525. The value of the game at the end of the capability acquisition program, $\mu = 1$, is shown as a function of t in Figure 4.

Since player 2 is the minimizer, he or she should choose $t = 0.2$, that is, to invest only in counter-conventional defence, despite the seemingly greater increase in competence by training in counter-insurgency warfare. In this case, player 2 can achieve the minimum value of approximately 2.2, which is still quite a bit greater than the case of complete competence (1.6), but considerably better than the initial

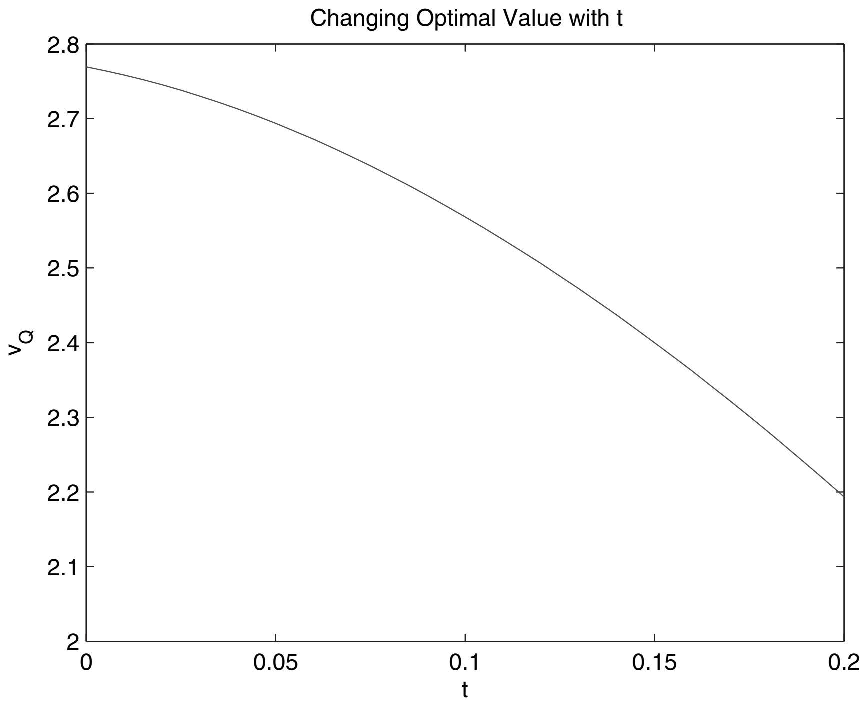


Figure 4: The value of the game after capability acquisition, $\mu = 1$, as the value of t changes. The value is best for player 2 when $t = 0.2$.

level (3.525). If player 2 chooses to spend all funds on counter-insurgency defence, the value of the game will be reduced to only approximately 2.75, a significantly worse result.

A similar method could be used to analyse more complex capability investment decisions.

7 Summary

This preliminary study has introduced a notion of incompetence in matrix games together with a proposed method of analysing such games. The method reduces an incompetent game to an equivalent competent game.

The approach lends itself to possible generalisations to more complicated situations, although this is not done explicitly here. One such extension is illustrated by an example described in Section 6. This and other examples indicate how the model could be used to examine real situations including, perhaps, those involving the determination of an “optimal” level of investment in defence, or training in sporting games such as tennis.

Further work in this area will examine incompetent nonzero-sum and dynamic games. An extension to extensive form games may also be useful for certain applications. In such a case, the concept of a trembling hand equilibrium introduced by Selten [5] may be of use in determining equilibria that are consistent with the concept of incompetent players.

Acknowledgments

Justin Beck is thankful for the support provided to him by DSTO, and in particular by Coen van Antwerpen. We are grateful to Geert-Jan Olsder, Harold Kushner, and Alain Haurie for their useful comments.

References

[1] D. Billings, N. Burch, A. Davidson, R. Holte, J. Schaeffer, T. Schauenberu and D. Szafron. Approximating game-theoretic optimal strategies for full-scale poker. In *Proceedings of the 18th International Conference on Artificial Intelligence*, 2003.

[2] M. Walker and J. Wooders. Minimax play at Wimbledon. *The American Econmic Review*, 91(5):1521–1538, 2001.

[3] O. G. Haywood Jr. Military decision and game theory. *Operations Research*, 2(4):365–385, 1954.

[4] D. R. Fulkerson and S. M. Johnson. A tactical air game. *Operations Research*, 5(5):704–712, Oct 1957.

[5] R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *Intl Journal of Game Theory*, 4(1):25–55, 1975.

[6] C. Stothard and R. Nicholson. Skill acquisition and retention in training. Technical Report DSTO-CR-0218, Defence Science and Technology Organisation, Edinburgh, South Australia, December 2001.

[7] G. Owen. *Game Theory*. W. B. Saunders Company, Philadelphia, PA, 1968.

[8] J. A. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer-Verlag, New York, 1996.

[9] L. S. Shapley and R. N. Snow. Basic solutions of discrete games. *Annals of Mathematics Study*, 1(24):27–35, 1950.

[10] J. A. Filar. Semi-antagonistic equilibrium points and action costs. *Cahiers du C.E.R.O.*, 25(3-4):227–239, 1984.

Stackelberg Well-Posedness and Hierarchical Potential Games

Marco Margiocco
Department of Mathematics
University of Genoa
via Dodecaneso 35, 16146 Genoa, Italy

Lucia Pusillo
Department of Mathematics
University of Genoa
via Dodecaneso 35, 16146 Genoa, Italy
pusillo@dima.unige.it

Abstract

We consider the Stackelberg well-posedness for hierarchical potential games and relate it to the Tikhonov well-posedness of the potential function as a maximum problem. We also make some considerations about the strong and weak Stackelberg approximate equilibria.

1 Introduction

In this chapter we study a Stackelberg problem corresponding to a game with two players: player I is the leader and player II is the follower.

The problem was introduced in 1934 by von Stackelberg [28] to give a result to many economic competitions. Later this problem was deeply studied by Chen and Cruz in 1972 [6] and by Simaan and Cruz in 1973 [26]. An exhaustive bibliography can be found in Bagchi [1], in Basar-Olsder [3], in Leitmann [10] and in Morgan [22].

Player I (the leader) chooses his strategy from the set X and player II (the follower) chooses his from the set Y: their goals are to maximize their utility functions. The follower conforms to the policies of the leader by allowing him to determine his strategy first, so the leader controls the whole game.

For economic applications with approximations and numerical aspects and further details about weak and strong Stackelberg problems, which are also called bilevel optimization problems, see references [2, 4, 7, 9, 11, 12, 14, 15, 16, 17, 23, 27].

In [22] Morgan characterized two-level optimization problems that are well posed.

This property comes from optimization ([8]) and it has economic motivations. In general, by well-posedness we mean that there is existence and uniqueness of the solution and a "maximizing" sequences (to be defined) approach to the solution ([5, 24, 25]).

To give an appropriate definition of "maximizing sequences," Morgan introduced in [22] the approximate Stackelberg solutions (we call them $(\epsilon, \eta)SE$). Intuitively, if player II chooses his best reply at minus then ϵ and player I is unlucky, he does not lose more then η.

In this paper we consider the strong Stackelberg well-posedness and the weak one (in the literature they are also called optimistic and pessimistic well-posedness respectively). We consider the relations between them when the best reply is not a singleton, otherwise they are the same. In Section 2 some definitions and preliminary theorems about Stackelberg well-posedness are given. Here we introduce strong and weak, approximate and exact Stackelberg equilibria. We use many examples to show the relations among them.

We note an interesting and paradoxical behavior of Stackelberg equilibria: it is possible that a Stackelberg equilibrium may not be an approximate equilibrium.

Then we study weak and strong Stackelberg well-posedness (*wSwp* and *sSwp* for short) and we show that these properties are different even if the game considered has a unique strong equilibrium, a unique weak equilibrium and they coincide.

We also give a metric characterization for Stackelberg well-posedness in a similar way as in [19] for another well-posedness property: Tikhonov well-posedness in value ($T^v wp$ for short) introduced in [18].

In Section 3 we consider the hierarchical potential games studied in [30] by Voorneveld, Tijs and Mallozzi. We prove that for these games, strong and weak well-posedness coincide even if the sets of approximate Stackelberg equilibria are not the same.

Finally we prove that these properties of well-posedness are equivalent to Tikhonov well-posedness (*Twp*) as a maximum problem of the potential function P.

This result shows that the behavior of Stackelberg approximate equilibria is very different from that of Nash approximate ones. For more details see [20].

2 Preliminaries

In this section we introduce the Stackelberg equilibrium problem.

Definition 2.1. Let $G = (X, Y, f, g)$ be a game with two players where X, Y are the strategy spaces and f, g are their utility functions. Let us consider the following problem: find $\overline{x} \in X$ such that

$$\inf_{y \in R_{II}(\overline{x})} f(\overline{x}, y) \geq \inf_{y \in R_{II}(x)} f(x, y) \quad \forall x \in X, \tag{1}$$

where $R_{II(\overline{x})} = \text{argmax}_{y \in Y} g(\overline{x}, y)$.

A pair $(\overline{x}, \overline{y}) \in X \times Y$ with $\overline{x}$ satisfying (1) and $\overline{y} \in R_{II(\overline{x})}$ is called a weak Stackelberg equilibrium (*wSE*) and $\overline{x}$ is called a weak Stackelberg solution (or pessimistic Stackelberg solution).

If we write "sup" instead of "inf", we have the strong Stackelberg equilibrium (*sSE*) or optimistic Stackelberg solution.

If we let $\beta(x) = \inf_{y \in R_{II(x)}} f(x, y)$ and $\gamma(x) = \sup_{y \in R_{II(x)}} f(x, y)$, then $(\overline{x}, \overline{y}) \in X \times Y$ is *wSE* if it satisfies $\beta(\overline{x}) = \max \beta(x)$ and $\overline{y} \in R_{II(\overline{x})}$ and $(\tilde{x}, \tilde{y}) \in X \times Y$ is *sSE* if it verifies $\gamma(\tilde{x}) = \max \gamma(x)$ and $\tilde{y} \in R_{II(\tilde{x})}$.

Intuitively the word "strong" ("weak") can be related with a strong (weak) attitude of the leader to influence (or not) the choice of the follower.

There is no link between these two definitions, as we show by the following examples.

Example 2.2. Let G be a game where $X = Y = [0, 1]$, $f(x, y) = 3x+6y-4xy$, $g(x, y) = \min\{y, 1/2\}$, $A = \{(0, y) : y \in [1/2, 1]\}$ is the set of strong Stackelberg equilibria, $B = \{(1, y) : y \in [1/2, 1]\}$ is the set of weak Stackelberg equilibria.

Example 2.3. Let G be a game where $X = Y = (0, 1]$, $f(x, y) = 3x+6y-4xy$, $g(x, y) = \min\{y, 1/2\}$, there are no strong Stackelberg equilibria but the set of weak ones is $B = \{(1, y) : y \in [1/2, 1]\}$.

Example 2.4. Let G be a game where $X = Y = [0, 1)$, $f(x, y) = 3x+6y-4xy$, $g(x, y) = \min\{y, 1/2\}$, there are no weak Stackelberg equilibria but the set of strong ones is $A = \{(0, y) : y \in [1/2, 1)\}$.

Definition 2.5. Given $(\epsilon, \eta) \in \mathbb{R}^2$ with $\epsilon, \eta \geq 0$, $\overline{x} \in X$ is an (ϵ, η) weak Stackelberg solution to problem (1) if, $\forall x \in X$,

$$\inf_{y \in R_{II(x,\eta)}} f(x, y) - \inf_{y \in R_{II(\overline{x},\eta)}} f(\overline{x}, y) \leq \epsilon, \tag{2}$$

where $R_{II(x,\eta)} = \{\tilde{y} \in Y : g(x, y) - g(x, \tilde{y}) \leq \eta \quad \forall y \in Y\}$.

That is, if player I is unlucky, he does not lose more than ϵ.

We say that $(\overline{x}, \overline{y})$ is a weak $(\epsilon, \eta)SE$ if $\overline{x}$ satisfies the condition (2) and $\overline{y} \in R_{II}(\overline{x}, \eta)$ and we say that $(\overline{x}, \overline{y})$ is a strong $(\epsilon, \eta)SE$ if $\overline{x}$ satisfies condition (2) but with "sup" instead of "inf" and $\overline{y} \in R_{II}(\overline{x}, \eta)$.

We write $(\epsilon, \eta)wSE$ for the set of (ϵ, η) weak Stackelberg equilibria and $(\epsilon, \eta)sSE$ for the set of (ϵ, η) strong Stackelberg equilibria.

Remark 2.6. If we define $\beta(x, \eta) = \inf_{y \in R_{II}(x,\eta)} f(x, y)$ and $\gamma(x, \eta) = \sup_{y \in R_{II}(x,\eta)} f(x, y)$, $(\overline{x}, \overline{y}) \in X \times Y$ is $(\epsilon, \eta)wSE$ if and only if it satisfies

$$\beta(\overline{x}, \eta) \geq \sup_{x \in X} \beta(x, \eta) - \epsilon, \quad \overline{y} \in R_{II}(\overline{x}, \eta)$$

and $(\overline{x}, \overline{y}) \in X \times Y$ is $(\epsilon, \eta)sSE$ if and only if it verifies

$$\gamma(\overline{x}, \eta) \geq \sup_{x \in X} \gamma(x, \eta) - \epsilon, \quad \overline{y} \in R_{II}(\overline{x}, \eta).$$

The following notions of maximizing Stackelberg sequences and Stackelberg well-posedness were introduced in [22].

Definition 2.7. The sequence $(x_n, y_n) \in X \times Y$ is a weak maximizing Stackelberg sequence (or an asymptotically w-Stackelberg sequence, or simply a weak *aSE*) if there is a sequence $(\epsilon_n, \eta_n) \in \mathbb{R}_+ \times \mathbb{R}_+$ converging to $(0, 0)$ for $n \to +\infty$ such that

1) x_n is a weak (ϵ_n, η_n) Stackelberg solution
2) $y_n \in R_{II}(x_n, \eta_n)$.

That is, $(x_n, y_n) \in (\epsilon_n, \eta_n)wSE$ for any n.

We define strong maximizing Stackelberg sequences analogously.

Definition 2.8. A game G is said to be weak Stackelberg well posed (*wSwp*) if

i) there is only one *wSE*: $(\overline{x}, \overline{y})$
ii) every (x_n, y_n) weak maximizing Stackelberg sequence converges to $(\overline{x}, \overline{y})$.

In an analogous way we define strong Stackelberg well-posedness (*sSwp*).

The sets of approximate equilibria are nested with respect to ϵ; Example 2.9 shows that, in general, they are not nested with respect to η.

It may happen that the Stackelberg equilibrium does not belong to the approximate equilibria sets, as we see with the following example.

Example 2.9. Let $G = ((0, 1], (0, 1], f, g)$ be a game with

$$f(x, y) = -x - y + 2xy, \; g(x, y) = y^2 - y.$$

It turns out that $R_{II}(x) = \operatorname{argmax}_{y \in (0,1]} g(x, y) = 1$. Then $R_{II}(x, \eta) = \{y : g(x, y) \geq \max_{y \in (0,1]} g(x, y) - \eta\} = (0, \frac{1-\sqrt{1-4\eta}}{2}] \cup [\frac{1+\sqrt{1-4\eta}}{2}, 1]$ if $\eta < 1/4$.

$\beta(x) = \inf_{y \in R_{II}(x)} f(x, y) = f(x, R_{II}(x)) = \sup_{y \in R_{II}(x)} f(x, y) = \gamma(x) = f(x, 1) = x - 1$

$\max \beta(x) = \max \gamma(x) = 0$ when $x = 1$, so $wSE = sSE = \{(1, 1)\}$

$$\beta(x, \eta) = \inf_{y \in (0, \frac{1-\sqrt{1-4\eta}}{2}] \cup [\frac{1+\sqrt{1-4\eta}}{2}, 1]} f(x, y) = \begin{cases} -x & \text{if } x \geq 1/2 \\ x - 1 & \text{if } x < 1/2 \end{cases}$$

$(\epsilon, \eta)wSE = \{(x, y) : \beta(x, \eta) \geq \sup_{x \in X} \beta(x, \eta) - \epsilon\} = \{(x, y) : x \in [1/2 - \epsilon, 1/2 + \epsilon], y \in R_{II}(x, \eta)\}$, if $\eta < 1/4, \epsilon < 1/2$ (see Figure 2.1).

This game is not *wSwp* and the unique $wSE \notin (\epsilon, \eta)wSE$.

$$\gamma(x, \eta) = \sup_{y \in R_{II}(x, \eta)} f(x, y) = \begin{cases} x - 1 & \text{if } x \geq 1/2 \\ -x & \text{if } x < 1/2 \end{cases}$$

$(\epsilon, \eta)sSE = \{(x, y) : \gamma(x, \eta) \geq \sup_{x \in X} \gamma(x, \eta) - \epsilon\} = \{(x, y) : x \in (0, \epsilon] \cup [1-\epsilon, 1], y \in R_{II}(x, \eta)\}$ (see Figure 2.2). This game is not *sSwp* and the unique $sSE \in (\epsilon, \eta)sSE$.

In Figure 2.1 we draw the approximate weak Stackelberg equilibria and in Figure 2.2 the strong ones (for one (ϵ, η) fixed).

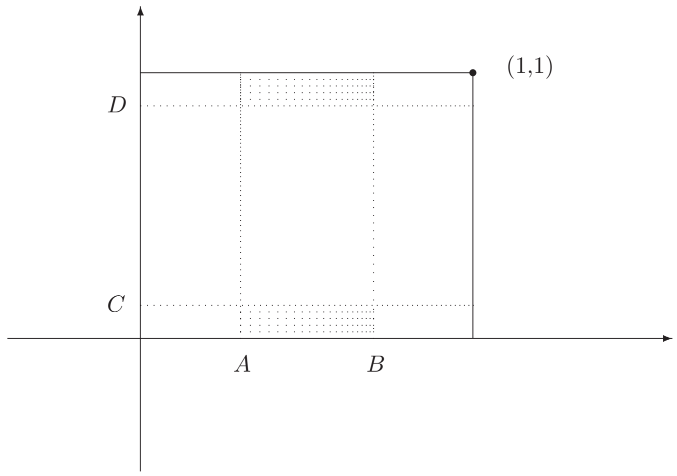


Figure 2.1.

$$(\epsilon, \eta)wSE = [1/2 - \epsilon, 1/2 + \epsilon] \times \left(\left(0, \frac{1-\sqrt{1-4\eta}}{2}\right] \cup \left[\frac{1+\sqrt{1-4\eta}}{2}, 1\right] \right)$$

where

$$A = (1/2 - \epsilon, 0), \quad B = (1/2 + \epsilon, 0),$$
$$C = \left(0, \frac{1-\sqrt{1-4\eta}}{2}\right), \quad D = \left(0, \frac{1+\sqrt{1-4\eta}}{2}\right).$$

$$(\epsilon, \eta)sSE = ((0, \epsilon] \cup [1-\epsilon, 1]) \times \left(\left(0, \frac{1-\sqrt{1-4\eta}}{2}\right] \cup \left[\frac{1+\sqrt{1-4\eta}}{2}, 1\right] \right)$$

where

$$A = (\epsilon, 0), \quad B = (1-\epsilon, 0),$$
$$C = \left(0, \frac{1-\sqrt{1-4\eta}}{2}\right), \quad D = \left(0, \frac{1+\sqrt{1-4\eta}}{2}\right)$$

In the following example we see that strong approximate equilibria are not nested while the weak ones are.

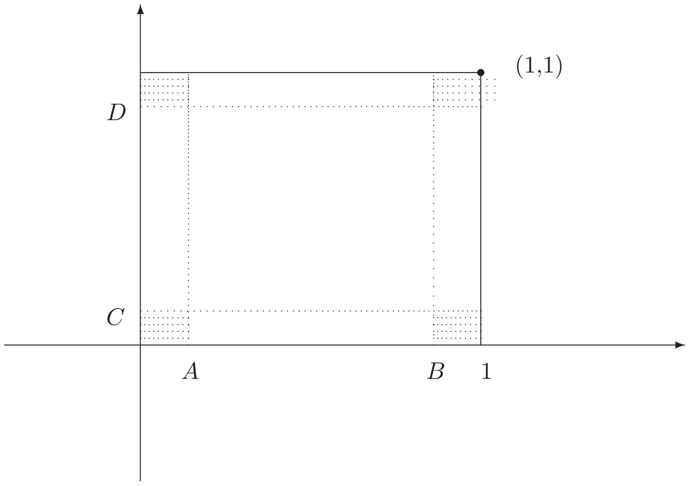


Figure 2.2.

Example 2.10. Let $G = ([0,1), [0,+\infty), f, g)$ with $f(x,y) = \frac{x-x^2}{y+1}$ and $g(x,y) = \frac{-y}{x+1}$.

It turns out that $R_{II}(x) = 0$,

$$\begin{aligned}
\beta(x) &= x - x^2 \quad \text{so} \\
wSE &= sSE = \{(1/2, 0)\} \\
R_{II}(x,\eta) &= \{y \in [0,+\infty) : \ y \le \eta(x+1)\} \\
\gamma(x,\eta) &= \sup_{y \le \eta(x+1)} f(x,y) = x - x^2 \\
\beta(x,\eta) &= \inf_{y \le \eta(x+1)} f(x,y) = \frac{x-x^2}{\eta(x+1)+1} \\
(\epsilon,\eta)sSE &= \{(x,y) : x \in [1/2 - \sqrt{\epsilon}, 1/2 + \sqrt{\epsilon}], \quad y \in [0, \eta(x+1)]\}
\end{aligned}$$

if $\epsilon < 1$, $\eta > 0$.

So $(\epsilon,\eta)sSE$ are nested and contain the equilibrium point, and the game G is *sSwp*.

Now let us determine $(\epsilon,\eta)wSE$.

$$\begin{aligned}
(\epsilon,\eta)wSE &= \{(x,y) \in X \times Y : \beta(x,\eta) \ge \sup_{x\in X} \beta(x,\eta) - \epsilon, \ y \in R_{II}(x,\eta)\} \\
&= \left\{ (x,y) : \ \frac{x-x^2}{\eta(x+1)+1} \right. \\
&\qquad \left. \ge \frac{1}{3\eta + 2 + 2\sqrt{2\eta^2 + 3\eta + 1}} - \epsilon, \ y \in [0, \eta(x+1)] \right\}.
\end{aligned}$$

Besides,

$$(0,\eta)wSE = \left\{(x,y)\ s.t.\ x = \frac{-(\eta+1)+\sqrt{2\eta^2+3\eta+1}}{\eta},\ y \in [0,\eta(x+1)]\right\}.$$

The sets $(0,\eta)wSE$ are disjoint segments, so they are not nested.

In this example we have seen that: the sets $(\epsilon,\eta)sSE$ are nested, they contain the strong equilibrium and the game is *sSwp*; instead $(\epsilon,\eta)wSE$ are not nested, they do not contain the weak equilibrium, if ϵ is small enough, but the game is *wSwp*.

In Figure 2.3 we draw weak Stackelberg approximate equilibria and in Figure 2.4 the strong ones.

To see better the sets of approximate equilibria we draw them in three cases: (ϵ_1,η_1), (ϵ_2,η_2), (ϵ_3,η_3) with $0<\epsilon_1<\epsilon_2<\epsilon_3$, $0<\eta_1<\eta_2<\eta_3$ and

$$r:\ y=\eta_1(x+1)\quad s: y=\eta_2(x+1)\quad t: y=\eta_3(x+1)$$

$(\epsilon,\eta)sSE = \{(x,y) : x \in [1/2-\sqrt{\epsilon}, 1/2+\sqrt{\epsilon}], y \in [0,\eta(x+1)]$ with $\epsilon < 1, \eta > 0\}$

$$A=(1/2-\sqrt{\epsilon_3},0),\ C=(1/2-\sqrt{\epsilon_2},0),\ E=(1/2-\sqrt{\epsilon_1},0),$$
$$B=(1/2+\sqrt{\epsilon_1},0),\ D=(1/2+\sqrt{\epsilon_2},0),\ F=(1/2+\sqrt{\epsilon_3},0),$$

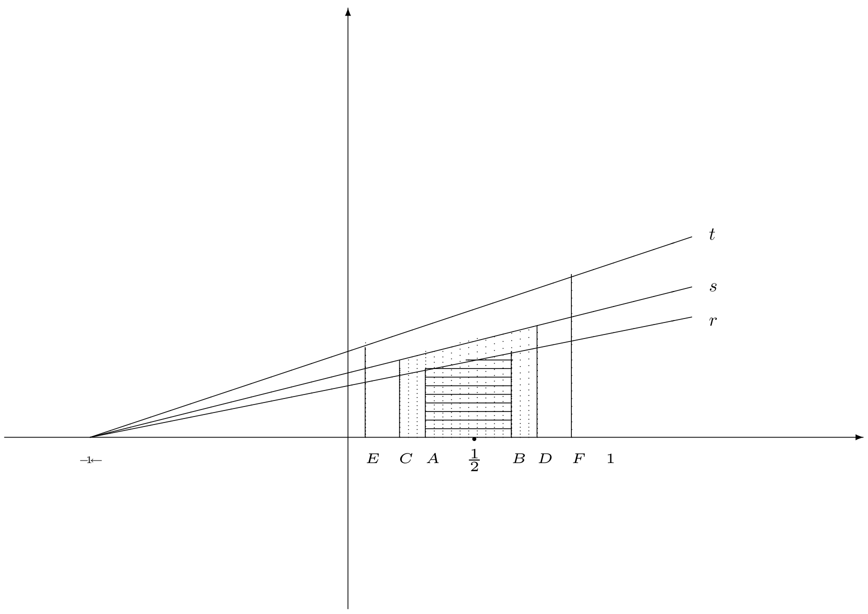


Figure 2.3.

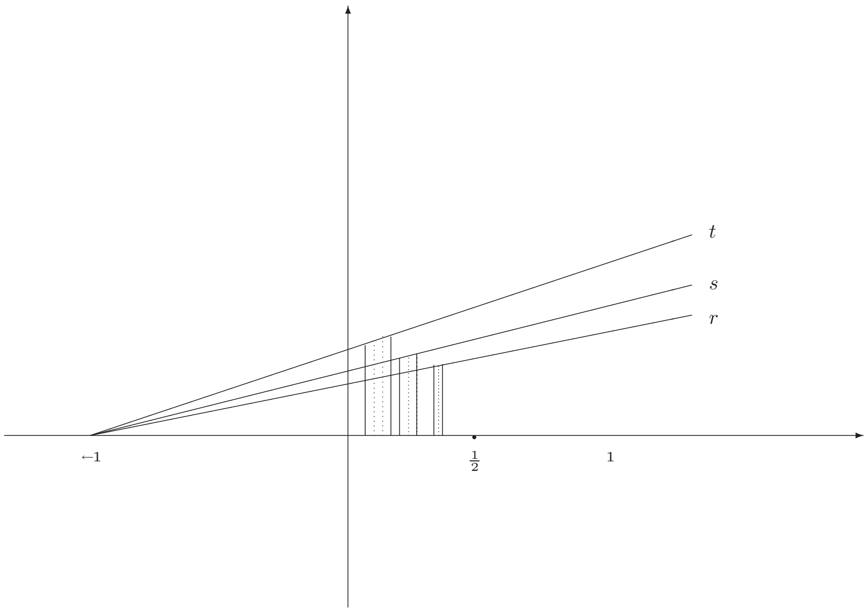


Figure 2.4.

$$(\epsilon,\eta)wSE = \Big\{(x,y) \in [0,1) \times [0,+\infty) : \frac{x-x^2}{\eta(x+1)+1} \geq \frac{\eta+1}{n+1+\sqrt{2\eta^2+3\eta+1}} - \epsilon,\ y \in [0,\eta(x+1)]\Big\}.$$

We can see that the behavior of Stackelberg approximate equilibria is very different from that of Nash approximate ones.

The following theorem gives us a metric characterization of Stackelberg well-posedness.

Theorem 2.11. *Let X, Y be metric spaces.*
Assume that there exists a unique wSE $(\overline{x}, \overline{y})$.
G is wSwp $\Leftrightarrow$ $\operatorname{diam}((\epsilon,\eta)wSE \cup (\overline{x},\overline{y})) \to 0$ *if* $(\epsilon,\eta) \to (0,0)$.
Assume that there exists a unique sSE $(\tilde{x}, \tilde{y})$.
G is sSwp $\Leftrightarrow$ $\operatorname{diam}((\epsilon,\eta)sSE \cup (\tilde{x},\tilde{y})) \to 0$ *if* $(\epsilon,\eta) \to (0,0)$.

Proof. "$\Leftarrow$". Let (x_n, y_n) be a weak maximizing Stackelberg sequence, $(x_n, y_n) \in (\epsilon_n, \eta_n)wSE$, $\forall n$ with $(\epsilon_n, \eta_n) \to (0,0)$. So

$$0 \leq \operatorname{dist}((x_n,y_n),(\overline{x},\overline{y})) \leq \operatorname{diam}(((\epsilon_n,\eta_n)wSE) \cup (\overline{x},\overline{y})).$$

"$\Rightarrow$". By contradiction let us suppose that $\operatorname{diam}((\epsilon,\eta)wSE \cup (\overline{x},\overline{y})) \not\to 0$ when $(\epsilon,\eta) \to (0,0)$. So there are $(\epsilon_n,\eta_n) \to (0,0)$ s.t.
$\operatorname{diam}((\epsilon_n,\eta_n)wSE \cup (\overline{x},\overline{y})) \not\to 0$,

Choosing a subsequence, if it is necessary, it turns out that
$\text{diam}((\epsilon, \eta)wSE \cup (\overline{x}, \overline{y})) \to \ell > 0$,
so $\forall n \in \mathbb{N}$ we choose
$(x_n, y_n) \in (\epsilon_n, \eta_n)wSE$ s.t. $\text{dist}((x_n.y_n), (\overline{x}, \overline{y}))$
$> \text{diam}((\epsilon_n, \eta_n)wSE \cup (\overline{x}, \overline{y})) - 1/n$.
$\text{diam}((\epsilon_n, \eta_n)wSE \cup (\overline{x}, \overline{y})) - 1/n < \text{dist}((x_n, y_n), (\overline{x}, \overline{y}))$
$< \text{diam}((\epsilon_n, \eta_n)wSE \cup (\overline{x}, \overline{y}))$,
from which $\text{dist}((x_n, y_n), (\overline{x}, \overline{y})) \to \ell$.

Because $(x_n, y_n) \in (\epsilon_n, \eta_n)wSE$ by definition it is a weak maximizing Stackelberg sequence but not converging. This is absurd.

In an analogous way we can prove strong well-posedness.

See [13] for a similar result applied to variational inequalities. □

The following examples show that there are games which are *wSwp* but not *sSwp* and conversely, so, in general, there is no link between weak and strong well-posedness even if the set of strong Stackelberg equilibria is the same as that of weak equilibria and it is a singleton.

Example 2.12. Let $G_1 = (X, Y, f_1, g_1)$, $X = Y = (0, 1]$.

$$f_1(x, y) = \begin{cases} x^2 - x & \text{if } y < x \\ x^2 - x + (2x + 1 - x^2)(y - x)/(1 - x) & \text{if } y \geq x, x \neq 1 \\ 2 & \text{if } (x, y) = (1, 1) \end{cases}$$

$$g_1(x, y) = \min(y/x, 1).$$

We have $wSE = sSE = \{(1, 1)\}$, yet this game is *sSwp* and not *wSwp*. In fact,

$$\begin{aligned} R_{II}(x) &= \{\tilde{y} \in (0, 1] : g(x, \tilde{y}) \geq g(x, y)\ \forall y \in (0, 1]\} \\ &= \{\tilde{y} \in (0, 1] : \min(\tilde{y}/x, 1) \geq \min(y/x, 1)\ \forall y \in (0, 1]\} = [x, 1], \end{aligned}$$

$$\beta(x) = \begin{cases} x^2 - x & \text{if } x \neq 1 \\ 2 & \text{if } x = 1. \end{cases}$$

One sees that $\max\beta(x) = \beta(\overline{x}) = \beta(1) = 2$, so $\{(1, 1)\}$ is the unique *wSE*.

Let us determine the strong Stackelberg equilibria:

$$\gamma(x) = \sup_{y \in [x,1]} f_1(x, y) = \begin{cases} x + 1 & \text{if } x \neq 1 \\ 2 & \text{if } x = 1 \end{cases}.$$

So $sSE = \{(1, 1)\}$ is a singleton and it coincides with the unique *wSE*.

We study approximate strong and weak Stackelberg equilibria.

$$R_{II(x,\eta)} = \{\tilde{y} \in Y : y/x \geq 1 - \eta\} = [x(1 - \eta), 1]$$

$$\beta(x, \eta) = \inf_{y \in [x(1-\eta),1]} f_1(x, y) = x^2 - x$$

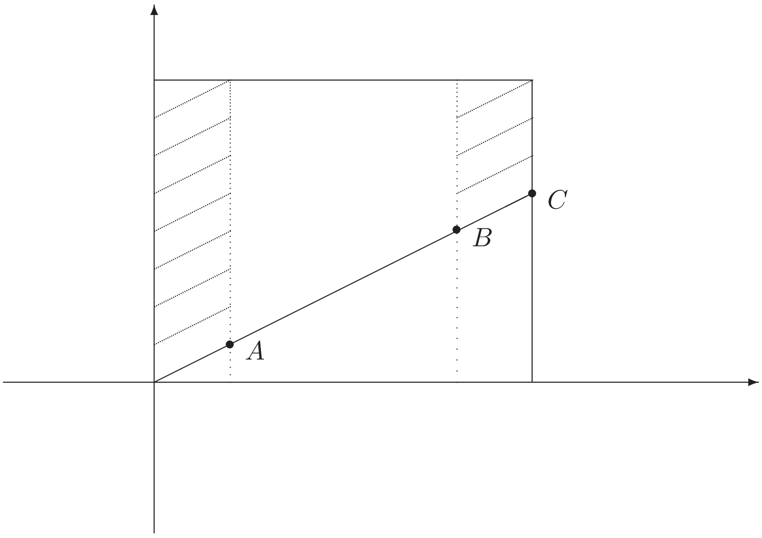


Figure 2.5.

$(\epsilon, \eta)wSE = \left\{(x, y)\ s.t.\ x \in \left(0, \frac{1-\sqrt{1-4\epsilon}}{2}\right] \bigcup \left[\frac{1+\sqrt{1-4\epsilon}}{2}, 1\right], y \in [x(1-\eta), 1]\right\}$ (see Figure 2.5)

So $\operatorname{diam}((\epsilon, \eta)wSE \bigcup (1, 1)) = \sqrt{2}$, hence G is not *wSwp*, but we can see that it is *sSwp*.

Because $\gamma(x, \eta) = \sup_{y \in [x(1-\eta), 1]} f_1(x, y) = x + 1$, it turns out that $(\epsilon, \eta)sSE = \{x \in (0, 1] : x \in [1-\epsilon, 1], y \in [x(1-\eta), 1]\}$ and in this case $\operatorname{diam}((\epsilon, \eta)sSE \bigcup (1, 1)) \to 0$ so G is strong well posed.

In Figure 2.5 we draw weak approximate equilibria and in Figure 2.6 the strong ones (for one (ϵ, η) fixed).

$$A \equiv \left(\frac{1-\sqrt{1-4\epsilon}}{2}, 1-\eta\right)$$
$$B \equiv \left(\frac{1+\sqrt{1-4\epsilon}}{2}, (1-\eta)\frac{1+\sqrt{1-4\epsilon}}{2}\right)$$
$$C \equiv (1, 1-\eta)$$

$$(\epsilon, \eta)wSE = \left\{(x, y) \in \mathbb{R}^2 :\ \mathrm{x} \in \left(0, \frac{1-\sqrt{1-4\epsilon}}{2}\right] \cup \left[\frac{1+\sqrt{1-4\epsilon}}{2}, 1\right],\right.$$
$$\left. y \in [x(1-\eta), 1]\right\}.$$

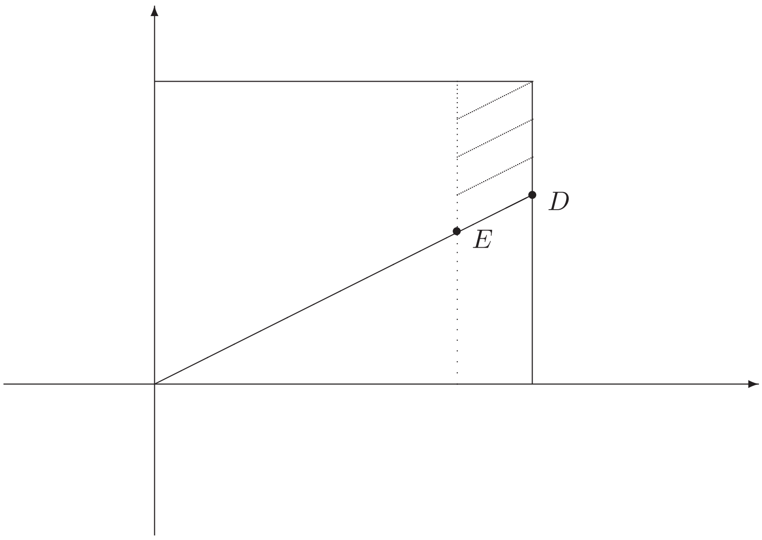


Figure 2.6.

$$D \equiv (1-\epsilon, (1-\eta)(1-\epsilon)),\ E \equiv (1, 1-\eta)$$
$$(\epsilon,\eta)sSE = \{(x,y) \in \mathbb{R}^2 :\ x \in [1-\epsilon, 1],\ y \in [x(1-\eta), 1]\}.$$

Example 2.13. Consider the game $G_2 = ((0.1], (0,1], f_2, g_2)$ where

$$f_2(x,y) = \begin{cases} x & \text{if } y < x \\ x + (x^2 - 2x + 2)(y-x)/(1-x) & \text{if } y \geq x, x \neq 1 \\ 2 & \text{if } (x,y) = (1,1) \end{cases}$$
$$g_2(x,y) = \min(x,y).$$

It turns out that $wSE = sSE = (1,1)$ and this game is *wSwp* and not *sSwp*.

3 Potential Games and Stackelberg Well-Posedness

Definition 3.1. $G = (X, Y, f, g)$ is a hierarchical potential game with two players if there are a potential function $P : X \times Y \to \mathbb{R}$ and a function $h : X \to \mathbb{R}$ such that
$$\begin{cases} f(x,y) = P(x,y) \\ g(x,y) = P(x,y) + h(x). \end{cases}$$

So a hierarchical potential game (studied in [30]) is a particular case of an exact potential game introduced in [21] and studied in detail in [29].

Proposition 3.2. *The following conditions are equivalent for a hierarchical potential game with potential P:*

i) $(\overline{x}, \overline{y}) \in X \times Y$ *is a weak Stackelberg equilibrium (wSE) of G*
ii) $(\overline{x}, \overline{y}) \in X \times Y$ *is a weak Stackelberg equilibrium (wSE) of* $G^P = (X, Y, P, P)$
iii) $(\overline{x}, \overline{y}) \in \operatorname{argmax}_{(x,y)\in X\times Y} P(x, y)$
iv) $(\overline{x}, \overline{y}) \in X \times Y$ *is a strong Stackelberg equilibrium (sSE) of G*
v) $(\overline{x}, \overline{y}) \in X \times Y$ *is a strong Stackelberg equilibrium (sSE) of* $G^P = (X, Y, P, P)$.

Proof. See [30]. □

Proposition 3.3. *A hierarchical potential game G is wSwp (sSwp) if and only if G^P is wSwp (sSwp respectively).*

Proof. The proof is based on the equalities $(\epsilon, \eta)wSE(G) = (\epsilon, \eta)wSE(G^P)$, $(\epsilon, \eta)sSE(G) = (\epsilon, \eta)sSE(G^P)$; that is, the set of approximate Stackelberg equilibria of game G and of game G^P are the same (either in the strong or in the weak case). □

The Stackelberg well-posedness of a game can be studied by the sets of approximate Stackelberg equilibria. These sets, for a hierarchical potential game, do not coincide with the superlevels of the potential function P, as Lemma 3.4 and Example 3.5 show.

Lemma 3.4. *Let $M = sup_{X\times Y} P(x, y) < +\infty$ and let $L(\epsilon) = \{(x, y) \in \mathbb{R}^2 : P(x, y) \geq M - \epsilon\}$, that is, the set of superlevels of P. The following inequalities are valid:*

i) $L(\epsilon) \subset (\epsilon + \eta, \eta)sSE \quad$ *if* $\epsilon \leq \eta$
ii) $(\epsilon, \eta)sSE \subset (\epsilon + \eta, \eta)wSE$
iii) $(\epsilon, \eta)wSE \subset L(2\epsilon + \eta)$.

Proof. (i) Suppose $(\overline{x}, \overline{y}) \in L(\epsilon)$ so $P(\overline{x}, \overline{y}) \geq M - \epsilon$ and $P(\overline{x}, y) \geq P(\overline{x}, \overline{y}) \geq M - \epsilon \ \forall y \in R_{II(\overline{x})}$. Because G is a hierarchical potential game, the following relations are valid: $P(\overline{x}, y) = \beta(\overline{x}) = \gamma(\overline{x})$ if $y \in R_{II(\overline{x})}$ so

$$M - \epsilon - \eta \leq \beta(\overline{x}) - \eta \leq \beta(\overline{x}, \eta) \leq \gamma(\overline{x}, \eta). \tag{3}$$

Because

$$\sup_{x\in X} \gamma(x, \eta) - (\epsilon + \eta) \leq M - (\epsilon + \eta),$$

from (3) it follows that $\sup_{x\in X} \gamma(x, \eta) - (\epsilon + \eta) \leq \gamma(\overline{x}, \eta)$ so $\overline{x}$ is a Stackelberg approximate solution or, more precisely, $\overline{x}$ is an $(\epsilon + \eta, \eta)$ strong Stackelberg solution and $\overline{y} \in R_{II}(\overline{x}, \eta)$ if $\epsilon \leq \eta$.

ii) We have $\forall y \in R_{II(\overline{x},\eta)}$. $P(x, R_{II}(x)) - \eta \leq P(x, y) \leq P(x, R_{II}(x))$ so $\beta(x, \eta), \gamma(x, \eta) \in J = [P(x, R_{II}(x)) - \eta, P(x, R_{II}(x))]$. So the following is valid:

$$\gamma(x, \eta) - \beta(x, \eta) \leq \eta. \tag{4}$$

Let $(\overline{x}, \overline{y}) \in (\epsilon, \eta)sSE$.

$\beta(x, \eta) - \beta(\overline{x}, \eta) = \beta(x, \eta) - \gamma(x, \eta) + \gamma(x, \eta) - \gamma(\overline{x}, \eta) + \gamma(\overline{x}, \eta) - \beta(\overline{x}, \eta) \leq 0 + \epsilon + \eta$, so the conclusion follows.

iii) Let $(\overline{x}, \overline{y}) \in (\epsilon, \eta)wSE$ and for a contradiction we suppose $(\overline{x}, \overline{y}) \notin L(2\epsilon + \eta)$ so $P(\overline{x}, \overline{y}) < M - 2\epsilon - \eta$.

It is possible to choose $(x_1, y_1) \in X \times Y$ s.t. $P(x_1, y_1) > M - \epsilon$. Then

$\inf_{y_2 \in R_{II(x_1,\eta)}} P(x_1, y_2) > M - \epsilon - \eta$

$\inf_{y_3 \in R_{II(\overline{x},\eta)}} P((\overline{x}, y_3) \leq P(\overline{x}, \overline{y}) < M - 2\epsilon - \eta$

$\inf_{y_2 \in R_{II(x_1,\eta)}} P(x_1, y_2) - \inf_{y_3 \in R_{II(\overline{x},\eta)}} P(\overline{x}, y_3)$

$> (M - \epsilon - \eta) - (M - 2\epsilon - \eta) = \epsilon,$

and this is absurd because $(\overline{x}, \overline{y}) \in (\epsilon, \eta)wSE$. □

Let us illustrate by the following example some of the preceding relationships.

Example 3.5. Let $G = (\mathbb{R}, \mathbb{R}, \mathrm{f}, \mathrm{g})$ be a game with

$$f(x, y) = g(x, y) = -x^2 - y^2 = P(x, y).$$

It turns out that

$$L(\delta) = \{(x, y) \in [-1, 1]^2 \text{ s.t. } x^2 + y^2 \leq \delta\}.$$

To determine $(\epsilon, \eta)wSE$, we write

$$\begin{aligned} R_{II(x,\eta)} &= \{\tilde{\mathrm{y}} \in \mathbb{R} \text{ s.t. } -\mathrm{x}^2 - \tilde{\mathrm{y}}^2 \geq -\mathrm{x}^2 - \mathrm{y}^2 - \eta \ \forall \mathrm{y} \in \mathbb{R}\} \\ &= \{\tilde{\mathrm{y}} \in \mathbb{R} \text{ s.t. } \tilde{\mathrm{y}}^2 \leq \eta\} = [-\sqrt{\eta}, \sqrt{\eta}] \end{aligned}$$

so

$$\inf_{y \in R_{II(x,\eta)}} (-x^2 - y^2) - \inf_{y \in R_{II(\overline{x},\eta)}} (-\overline{x}^2 - y^2) \leq \epsilon, \tag{5}$$

which becomes

$$inf_{y \in [-\sqrt{\eta}, \sqrt{\eta}]}(-x^2 - y^2) - \inf_{y \in [-\sqrt{\eta}, \sqrt{\eta}]} (-\overline{x}^2 - y^2) \leq \epsilon. \tag{6}$$

$$\overline{x} \in [-\sqrt{\epsilon}, \sqrt{\epsilon}]$$

and

$$\overline{y} \in [-\sqrt{\eta}, \sqrt{\eta}],$$

therefore

$$L(\delta) \subset (\epsilon, \eta)sSE = (\epsilon, \eta)wSE = [-\sqrt{\epsilon}, \sqrt{\epsilon}] \times [-\sqrt{\eta}, \sqrt{\eta}]$$

if $\delta = \min(\epsilon, \eta)$.

Furthermore, $(\epsilon, \eta)wSE \subset L(\delta_1)$ if $\delta_1 = \epsilon + \eta$.

Theorem 3.6. *Let G be a hierarchical potential game with potential function P. The following conditions are equivalent:*

a) G is weak Stackelberg well posed
b) G is strong Stackelberg well posed
c) P is Tikhonov well posed as a maximum problem.

Proof. $\operatorname{diam}((\epsilon, \eta)wSE \cup (\overline{x}, \overline{y})) \to 0$ so for ii) of Lemma 3.4 we conclude that G is strong *Swp*. If G is strong Swp, diam $((\epsilon, \eta)sSE \cup (\overline{x}, \overline{y})) \to 0$ by hypothesis and

$$(\epsilon, \eta)wSE \subset L(2\epsilon + \eta) \subset (4\epsilon + 2\eta, 2\epsilon + \eta)sSE,$$

and so G is weak Swp.

Because $(\epsilon, \eta)sSE \subset (\epsilon + \eta, \eta)wSE \subset L(2\epsilon + 3\eta)$ and $L(\epsilon) \subset (\epsilon + 2\epsilon)sSE$, the equivalence between the strong *Swp* of the game and the *Twp* of the function P follows. □

The following example illustrates that, in general, $wS(\epsilon, \eta) \neq sS(\epsilon, \eta)$, $\forall\epsilon, \eta \geq 0$; that is, the set of strong approximate Stackelberg equilibria is different from the set of weak approximate Stackelberg equilibria (even for a hierarchical potential game). In spite of this, the strong Stackelberg well-posedness and the weak Stackelberg well-posedness for hierarchical potential games coincide.

Example 3.7. Let $G = (X, Y, P, P)$, $X = Y = [0, 1]$.

$$P(x, y) = \begin{cases} x + y & \text{if } y < 1 \\ (x + 3)/2 & \text{if } y = 1 \end{cases},$$

so $R_{II(x)} = 1\ \forall x \in [0, 1]$

$$R_{II(x,\eta)} = \begin{cases} 1 & \text{if } x \in [0, 1 - 2\eta] \\ [((3 - x)/2) - \eta, 1] & \text{if } x \in (1 - 2\eta, 1] \end{cases}$$

$$\beta(x, \eta) = \inf_{R_{II(x,\eta)}} P(x, y) = \begin{cases} (x + 3)/2 & \text{if } x \in [0, 1 - 2\eta] \\ ((x + 3)/2) - \eta & \text{if } x \in (1 - 2\eta, 1] \end{cases}$$

$$\gamma(x, \eta) = \sup_{R_{II(x,\eta)}} P(x, y) = P(x, 1) = (x + 3)/2$$

$$(\epsilon, \eta)wSE = \{(\overline{x}, \overline{y}) : \overline{x} \in [1 - 2\eta - 2\epsilon, 1 - 2\eta] \cup [1 - 2\epsilon, 1],\ \ \overline{y} \in R_{II(\overline{x},\eta)}\}$$

$$(\epsilon, \eta)sSE = \{(\overline{x}, \overline{y}) : \overline{x} \in [1 - 2\epsilon, 1],\ \ \overline{y} \in R_{II(\overline{x},\eta)}\}.$$

In the following example we show that $(\epsilon, \eta)wSE$ can be nested in a decreasing way with respect to η.

Example 3.8. Let $G = ([0, +\infty), [0, +\infty), f, g)$ where $f(x, y) = g(x, y) = P(x, y)$,

$$f(x, y) = \begin{cases} 1 - x & \text{if } 0 \le y \le x \\ 0 & \text{otherwise} \end{cases}$$

$$R_{II}(x) = \text{argmax}_{y \in [0,+\infty)} g(x, y) = \begin{cases} 0 & \text{if } x = 0 \\ [0, x] & \text{if } x \in (0, 1) \\ [0, +\infty) & \text{if } x = 1 \\ (x, +\infty) & \text{if } x > 1 \end{cases}$$

$wSE = sSE = \{(0, 0)\} = \text{argmax}\, P(x, y)$

$$R_{II}(x, \eta) = \begin{cases} [0, +\infty) & \text{if } x \in [1 - \eta, 1 + \eta], \eta < 1 \\ R_{II}(x) & \text{otherwise} \end{cases}$$

$$\beta(x, \eta) = \begin{cases} 0 & \text{if } 1 - \eta \le x \le 1 \\ 1 - x & \text{otherwise} \end{cases}$$

$(\epsilon, \eta)wSE = \{(x, y) \in \mathbb{R}^2, \text{x} \in [0, 1 - \eta], \text{y} \in [0, \text{x}]\}$ if $\epsilon \in (1 - \eta, 1)$.

So the approximate weak Stackelberg equilibria are nested in this way: $(\epsilon, \eta_2)wSE \subset (\epsilon, \eta_1)wSE$ if $\eta_2 > \eta_1$ and $\epsilon \in (1 - \eta_1, 1)$.

Because

$$\gamma(x, \eta) = \begin{cases} 1 - x & \text{if } x < 1 \\ 0 & \text{if } x \ge 1, \end{cases}$$

it follows that $(\epsilon, \eta)sSE$ are nested in an increasing way (as we were expecting) and $(\epsilon, \eta_2)sSE = \{(x, y) \in \mathbb{R}^2 \text{s.t.x} \in [0, \epsilon], \text{y} \in [0, \text{x}]\}$ if $\epsilon < 1$.

For hierarchical potential games the following relation shows that $(\epsilon, \eta)sSE$ are increasing sets with respect either to ϵ or to η.

Proposition 3.9. *For the strong approximate Stackelberg equilibria, the following relations are valid:*

$$(\epsilon, \eta_1)sSE \subset (\epsilon, \eta_2)sSE \quad \text{if } \eta_1 \le \eta_2.$$

Proof. $\gamma(x, \eta_1) = \sup_{y \in R_{II}(x,\eta_1)} P(x, y) = \sup_{y \in R_{II}(x)} P(x, y) = \gamma(x) = \gamma(x, \eta)$.

If $(\overline{x}, \overline{y}) \in (\epsilon, \eta_1)sSE$ then $\overline{x}$ is a (ϵ, η_2) Stackelberg strong solution, moreover $\overline{y} \in R_{II}(\overline{x}, \eta_1) \subset R_{II}(\overline{x}, \eta_2)$ so $(\overline{x}, \overline{y}) \in (\epsilon, \eta_2)sSE$. □

Concluding, we have seen that approximate Stackelberg equilibria are increasing sets with respect to ϵ, but for potential games the strong ones are increasing with respect to ϵ and η.

4 Conclusions

We would like to point out an interesting feature of the approximation that we consider for Stackelberg problems: a Stackelberg equilibrium need not be an approximate Stackelberg equilibrium. It sounds paradoxical, but we believe that the definition for approximate Stackelberg equilibrium (as given in [22]) is quite natural, since it allows approximate solutions at both decision levels.

Another concluding remark is the following: in paper [20], (in which Nash equilibria are involved), we concluded that there is no relationship between *Twp* of P (as a maximum problem) and *Twp* of an exact potential game (with potential function P).

In this chapter we show that the situation is different for a hierarchical potential game and Stackelberg well-posedness: to say that P is *Twp* as a maximum problem is equivalent to the fact that the corresponding hierarchical potential game G is Stackelberg well posed.

Acknowledgments

This paper has been partially supported by PRIN "Approximations and applications in Game Theory." The authors are gratefully indebted to Professor F. Patrone for many interesting discussions and to Professor L. Mallozzi for an illuminating conference held at the Department of Mathematics in the University of Genoa. Additional thanks for interesting discussions and encouragements go to the usual participants of "Seminario itinerante: Ottimizzazione, Teoria dei Giochi e Applicazioni" organized by F. Patrone and R. Lucchetti. We also wish to acknowledge the kind suggestions of an anonymous referee.

References

[1] Bagchi A., Stackelberg differential games in economic models, *Lecture Notes in Control and Information Sciences*, 64, Springer, Berlin, 1984.

[2] Bard J., An efficient point algorithm for a linear two-stage optimization problem, *Operations Research*, 1983, 670–684.

[3] Başar T. and Olsder G.J., *Dynamic Noncooperative Game Theory*, Academic Press, New York, 31, 1982.

[4] Breton M., Aij A. and Haurie A., Sequential Stackelberg equilibria in two-person games, *Journal of Optimization Theory and Applications*, 59, 1980, 71–97.

[5] Cavazzuti E. and Morgan J., Well-posed saddle point problems, in *Optimization, Theory and Algorithms* (J. B. Hirriart-Urruty, W. Oettli and J. Stoer, eds.), Proc. Conf. Confolant/France 1981, 1983, 61–76.

[6] Chen C.I. and Cruz J.B., Stackelberg solution for two-person games with biased information patterns, *IEEE Transaction on Automatic Control*, 17, 1972, 791–798.

[7] Dempe S., On an algorithm solving two-level programming problems with nonunique lower level solutions, *Comp. Opti. Appl.*, 6, 1996, 227–249.

[8] Dontchev A. and Zolezzi T., Well-posed optimization problems, *Lecture Notes in Mathematics*, 1543, Springer, Berlin, 1993.

[9] Flam S., Mallozzi L. and Morgan J., A new look on Stackelberg Cournot equilibria in oligopolistic markets, *Economic Theory*, 20, 2002, 183–188.

[10] Leitmann G., On generalized Stackelberg strategies, *Journal of Optimization Theory and Applications*, 26, 1978, 637–643.

[11] Lignola M.B. and Morgan J., Topological existence and stability for Stackelberg problem. *Journal of Optimization Theory and Applications*, 84, 1997, 145–169.

[12] Lignola M.B. and Morgan J., Stability for regularized bilevel programming problems, *Journal of Optimization Theory and Applications*, 93, 1997, 575–596.

[13] Lignola M.B. and Morgan J., Well-posedness for optimization problems with constraints defined by variational inequalities having a unique solution, *Journal of Global Optimization*, 6, 2000, 57–67.

[14] Loridan P. and Morgan J., New results on approximate solutions in two-level optimization, *Optimization*, 20, 1989, 819–836.

[15] Loridan P. and Morgan J., Regularization for two-level optimization problems, *Lecture Notes in Economic and Math. Systems*, 382, 1992, 239–255.

[16] Loridan P. and Morgan J., Weak via strong Stackelberg problem: new results, *Journal of Global Optimization*, 8, 1996, 263–287.

[17] Mallozzi L. and Morgan J., Problema di Stackelberg con risposte pesate, *Atti del XIX Convegno della Associazione per la Matematica Applicata alle Scienze Economiche e Sociali*, 1995, 416–425.

[18] Margiocco M., Patrone F. and Pusillo Chicco L., A new approach to Tikhonov well-posedness for Nash equilibria, *Optimization*, 40, 1997, 385–400.

[19] Margiocco M., Patrone F. and Pusillo Chicco L., Metric characterizations of Tikhonov well-posedness in value, *Journal of Optimization Theory and Applications*, 100, n. 2, 1999, 377–387.

[20] Margiocco M. and Pusillo L., Potential games and well-posedness properties, preprint 2005, submitted.

[21] Monderer P. and Shapley L., Potential games, *Games and Economic Behaviour*, 14, 1996, 124–143.

[22] Morgan J., Constrained well-posed two-level optimization problems, *Nonsmooth Optimization and Related Topics*, edited by F. H. Clarke, V. F. Dem'yanov and F. Giannessi, Plenum Press, New York, 1989, 307–325.

[23] Morgan J. and Patrone F., Stackelberg problems: Subgame perfect equilibria via Tikhonov regularization, *Advances in Dynamic Games: Applications to Economics, Management Science, Engineering, and Environmental Management* (Haurie A., Muto S., Petrosjan L.A., and Raghavan T.E.S., eds), Bizkhauser, Boston, 2006, 209–221.

[24] Patrone F., Well-posedness as an ordinal property, *Rivista di Matematica pura ed applicata* 1, 1987, 95–104.

[25] Patrone F., Well posedness for Nash equilibria and related topics, *Recent Developments in Well posed Variational Problems* (Lucchetti R. and Revalski J., eds), Kluwer, Dordrecht, 1995, 211–227.

[26] Simaan M. and Cruz J., On the Stackelberg strategy in nonzero-sum games, *Journal of Optimization Theory and Applications*, 11, 1973, 533–555.

[27] Sheraly H.D., Soyster A.L. and Murphy F.H., Stackelberg–Nash–Cournot equilibria: characterizations and computations, *Operations Research*, 31, 1983, 253–276.

[28] Von Stackelberg H., *Marktform und Gleichgewitch*, Springer, Vienna, 1934.

[29] Voorneveld M., Potential games and interactive decisions with multiple criteria, *CentER Dissertation Series*, Tilburg University, 1999.

[30] Voorneveld M., Tijs S. and Mallozzi L., Sequential production situations and potentials, *Game Practice Book* (F. Patrone, I. Garcia Jurado, S. Tijs, eds.), Kluwer Academic Publishers, Boston, 1999.

PART II
Stochastic Differential Games

Ergodic Problems in Differential Games

Olivier Alvarez
UMR 60-85
Université de Rouen
76821 Mont-Saint Aignan Cedex, France
olivier.alvarez@univ-rouen.fr

Martino Bardi
Dipartimento di Matematica Pura ed Applicata
Università di Padova
via Belzoni 7
35131 Padova, Italy
bardi@math.unipd.it

Abstract

We present and study a notion of ergodicity for deterministic zero-sum differential games that extends the one in classical ergodic control theory to systems with two conflicting controllers. We show its connections with the existence of a constant and uniform long-time limit of the value function of finite horizon games, and characterize this property in terms of Hamilton–Jacobi–Isaacs equations. We also give several sufficient conditions for ergodicity and describe some extensions of the theory to stochastic differential games.

Introduction

We consider a nonlinear system in $\mathbb{R}^m$ controlled by two players

$$\dot{y}(t) = f(y(t), a(t), b(t)), \qquad y(0) = x, \tag{1}$$

and we denote by $y_x(\cdot)$ the trajectory starting at x. We are also given a bounded, uniformly continuous running cost l, and we are interested in the payoffs associated to the *long time average cost* (LTAC), namely,

$$J^\infty(x, a(\cdot), b(\cdot)) := \limsup_{T\to\infty} \frac{1}{T}\int_0^T l(y_x(t), a(t), b(t))\, dt,$$

$$J_\infty(x, a(\cdot), b(\cdot)) := \liminf_{T\to\infty} \frac{1}{T}\int_0^T l(y_x(t), a(t), b(t))\, dt.$$

We denote by $u - \mathrm{val}\, J^\infty(x)$ (respectively, $l - \mathrm{val}\, J_\infty(x)$) the upper value of the zero-sum game with payoff J^∞ (respectively, the lower value of the game with payoff J_∞) which the first player $a(\cdot)$ wants to minimize while the second player $b(\cdot)$ wants to maximize, and the values are in the sense of Variya–Roxin–Elliott–Kalton. We say that the LTAC game is *ergodic* if

$$u - \text{val}\, J^\infty(x) = l - \text{val}\, J_\infty(x) = \lambda \quad \forall x,$$

for some constant λ.

The terminology is motivated by the analogy with classical ergodic control theory, see, e.g., [6, 7, 9, 10, 13, 22, 26, 28]. Note also that if the controls can take only one value the game is ergodic for all continuous l if the dynamical system $\dot{y} = f(y)$ is ergodic with a unique invariant measure (see Proposition 13 of [3] for a precise statement). Similar problems were already studied for some games, in particular by Fleming and McEneaney [18] in the context of risk-sensitive control, Carlson and Haurie [12] within the turnpike theory, and Kushner [27] for controlled nondegenerate diffusion processes. After this research was completed we also learned of the recent paper [21]. There is a large literature on related problems for discrete-time games; see the recent survey by Sorin [34].

In order to have a compact state space we assume that the data f and l are $\mathbb{Z}^m$-periodic. First of all we show the connection between the ergodicity of the LTAC game and the existence of a constant uniform long-time limit of the lower and upper value functions of the finite horizon games with the same running cost. We call this property ergodicity of the lower (respectively, upper) game. Then we prove that the lower game is ergodic with limit λ if and only if the lower value of the discounted infinite horizon game with payoff

$$\delta \int_0^\infty l(y_x(t), a(t), b(t))\, e^{-\delta t}\, dt$$

converges uniformly to λ as the discount rate δ tends to 0. Moreover, this is also equivalent to the existence of a $\mathbb{Z}^m$-periodic viscosity χ to the Hamilton–Jacobi–Isaacs equation

$$\lambda + \min_b \max_a \{-f(y, a, b) \cdot \nabla \chi - l(y, a, b)\} = 0,$$

and similar statements hold for the upper value.

Next we describe two sets of conditions ensuring the previous facts and therefore the ergodicity. The first is a bounded-time controllability property of the system by one of the players, uniformly with respect to the behavior of the opponent. It is a generalization to games of a condition used for systems with a single controller by Grüne [22], Arisawa [6], and Artstein and Gaitsgory [8].

Different from the first, the second set of conditions is symmetric for the two players. We assume that some state variables y^A are asymptotically controllable by the first player, and the remaining variables y^B are asymptotically controllable by the second (see Section 2 for the precise definition). In this case neither player can control the whole state vector $y = (y^A, y^B)$. We further assume that the running cost depends only on y^A and y^B and has a saddle point, namely,

$$\min_{y^A} \max_{y^B} l(y^A, y^B) = \max_{y^B} \min_{y^A} l(y^A, y^B) =: \bar{l}.$$

Then we show that the LTAC game has the value $\lambda = \bar{l}$.

In the last section we also show that for systems affected by a nondegenerate white noise the game is ergodic with no controllability assumptions on either player (see [27] for related results).

Our methods rely heavily on the Hamilton–Jacobi–Isaacs equations associated to the games, in the framework of the theory of viscosity solutions. We follow ideas of authors such as P.-L. Lions and L.C. Evans, see [7, 9, 15, 28, 29], and their developments in our papers [3, 4].

Undiscounted infinite horizon control problems arise in many applications to economics and engineering, see [10, 13, 26] and [12, 18, 34] for games. Our additional motivation is that ergodicity plays a crucial role in the theory of singular perturbation problems for the dimension reduction of multiple-scale systems [10, 20, 24–26, 31, 35, 36] and for the homogenization in oscillating media [1, 5, 11, 15, 16, 23, 29, 30]. A general principle emerging in the papers [2–4, 8] is that an appropriate form of ergodicity of the fast variables (for frozen slow variables) ensures the convergence of the singular perturbation problem, in a suitable sense.

The paper is organized as follows. Section 1 describes the connection between the ergodicity of the LTAC game and the ergodicity of the lower and upper games. Section 2 studies the ergodicity of the finite horizon games. Section 3 presents some examples. In Section 4 we give some extensions of the results of Sections 2 and 3 to diffusion processes controlled by two players, and we prove the ergodicity result for nondegenerate noise.

1 The Long Time Average Cost Game and Ergodicity

For the system (1) and the cost we assume throughout the paper that $f : \mathbb{R}^m \times A \times B \mapsto \mathbb{R}^m$ and $l : \mathbb{R}^m \times A \times B \mapsto \mathbb{R}$ are continuous and bounded, A and B are compact metric spaces, and f is Lipschitz continuous in x uniformly in a, b. In this section we do not assume the compactness of the state space.

We consider the cost functional

$$J(T, x) = J(T, x, a(\cdot), b(\cdot)) := \frac{1}{T} \int_0^T l(y_x(t), a(t), b(t))\, dt,$$

where $y_x(\cdot)$ is the trajectory corresponding to $a(\cdot)$ and $b(\cdot)$. We denote by $\mathcal{A}$ and $\mathcal{B}$, respectively, the sets of open-loop (measurable) controls for the first and the second player, and by Γ and Δ, respectively, the sets of nonanticipating strategies for the first and the second player, see, e.g., [9, 17] for the precise definition.

Finally, we define the upper and lower values for the finite horizon game with average cost

$$u - \text{val}\, J(T, x) := \sup_{\beta\in\Delta} \inf_{a\in\mathcal{A}} J(T, x, a, \beta[a]),$$
$$l - \text{val}\, J(T, x) := \inf_{\alpha\in\Gamma} \sup_{b\in\mathcal{B}} J(T, x, \alpha[b], b),$$

and for the LTAC game

$$u - \text{val}\, J^{\infty}(x) := \sup_{\beta\in\Delta} \inf_{a\in\mathcal{A}} \limsup_{T\to\infty} J(T, x, a, \beta[a]),$$
$$l - \text{val}\, J_{\infty}(x) := \inf_{\alpha\in\Gamma} \sup_{b\in\mathcal{B}} \liminf_{T\to\infty} J(T, x, \alpha[b], b).$$

We say that *the lower game is (uniformly) ergodic* if the long-time limit of the finite horizon value exists, uniformly in x, and it is constant, i.e.,

$$l - \text{val}\, J(T, \cdot) \to \lambda \quad \text{as } T \to \infty \text{ uniformly in } \mathbb{R}^m.$$

Similarly, *the upper game is ergodic* if

$$u - \text{val}\, J(T, \cdot) \to \Lambda \quad \text{as } T \to \infty \text{ uniformly in } \mathbb{R}^m.$$

Theorem 1.1. *If the lower game is ergodic, then*

$$l - \text{val}\, J_{\infty}(x) = \lim_{T\to\infty} l - \text{val}\, J(T, x) = \lambda \quad \forall x \in \mathbb{R}^m; \tag{2}$$

if the upper game is ergodic, then

$$u - \text{val}\, J^{\infty}(x) = \lim_{T\to\infty} u - \text{val}\, J(T, x) = \Lambda \quad \forall x \in \mathbb{R}^m. \tag{3}$$

Proof. We follow the arguments of the proof of Theorem 8.4 in [18]. We begin wih the proof that $l - \text{val}\, J_{\infty}(x) \geq \lambda$. To achieve this goal we fix $x \in \mathbb{R}^m$, $\overline{\alpha} \in \Gamma$, and $\varepsilon > 0$, and construct $\overline{b} \in \mathcal{B}$ such that

$$J_{\infty}(x; \overline{\alpha}[\overline{b}], \overline{b}) \geq \lambda - \varepsilon. \tag{4}$$

We set $v(T, x) := l - \text{val}\, J(T, x)$ and choose T_o such that

$$|v(T_o, y) - \lambda T_o| \leq \frac{\varepsilon T_o}{3} \quad \forall y \in \mathbb{R}^m. \tag{5}$$

By definition of v, for all $\alpha \in \Gamma$ and $z \in \mathbb{R}^m$,

$$\sup_{b\in\mathcal{B}} \int_0^{T_o} l(y_z(t), \alpha[b](t), b(t))\, dt \geq v(T_o, z), \tag{6}$$

where $y_z(\cdot) = y_z(\cdot; \alpha, b)$ solves

$$\dot{y} = f(y, \alpha[b], b), \qquad y(0) = z.$$

Then, for $z = x$ and $\alpha = \overline{\alpha}$, we can choose $\overline{b} \in \mathcal{B}$ such that

$$\int_0^{T_o} l(y_x(t), \overline{\alpha}[\overline{b}](t), \overline{b}(t))\,dt \geq v(T_o, x) - \frac{\varepsilon T_o}{3}.$$

This defines the desired control $\overline{b}$ on the interval $[0, T_o]$. Its definition on the intervals $[nT_o, (n+1)T_o)$ is obtained inductively as follows. Suppose $\overline{b}$ is defined on the interval $[0, nT_o)$. Define the strategy $\alpha_n \in \Gamma$ by $\alpha_n[b] := \overline{\alpha}[b_n]$, where

$$b_n(t) := \begin{cases} \overline{b}(t), & 0 \leq t < nT_o, \\ b(t - nT_o), & t \geq nT_o. \end{cases}$$

In (6) put $z = y_x(nT_o; \overline{\alpha}, \overline{b})$, $\alpha = \alpha_n$, and choose b_n such that

$$\int_0^{T_o} l(y_z(t), \alpha_n[b_n](t), b_n(t))\,dt \geq v(T_o, z) - \frac{\varepsilon T_o}{3}.$$

Now define $\overline{b}(t) := b_n(t + nT_o)$ for $t \in [nT_o, (n+1)T_o)$. Then

$$\int_{nT_o}^{(n+1)T_o} l(y_x(t), \overline{\alpha}[\overline{b}](t), \overline{b}(t))\,dt \geq v(T_o, z) - \frac{\varepsilon T_o}{3}.$$

By adding over n and using (5) we get

$$\int_0^{nT_o} l(y_x(t), \overline{\alpha}[\overline{b}](t), \overline{b}(t))\,dt \geq n\lambda T_o - \frac{2\varepsilon n T_o}{3}$$

and therefore

$$J(nT_o, x, \overline{\alpha}[\beta], \beta) \geq \lambda - \frac{2\varepsilon}{3}.$$

Now we write $T = nT_o + t_o$, where n is the integer part of T/T_o and $0 \leq t_o < T_o$, and observe that, for any bounded integrand $\bar{l}(t)$,

$$\frac{1}{T}\int_0^T \bar{l}(t)\,dt \geq \frac{nT_o}{nT_o + t_o} \cdot \frac{1}{nT_o}\int_0^{nT_o} \bar{l}(t)\,dt + \frac{t_o \inf \bar{l}}{nT_o + t_o}.$$

Then

$$J(T, x, \overline{\alpha}[\beta], \beta) \geq c_n J(nT_o, x, \overline{\alpha}[\beta], \beta) + c_n' \geq \lambda - \varepsilon$$

for all T large enough, because $c_n \to 1$ and $c_n' \to 0$ as $n \to \infty$. Now we let $T \to \infty$ and obtain the desired inequality (4).

In order to prove that $l - \mathrm{val}\, J_\infty(x) \le \lambda$, one fixes $\varepsilon > 0$ and constructs a strategy $\widetilde{\alpha} \in \Gamma$ such that, for all $b \in \mathcal{B}$,

$$\int_0^{nT_o} l(y_x(t), \widetilde{\alpha}[b](t), b(t))\, dt \le nT_o(\lambda + \varepsilon),$$

where T_o satisfies (5). This can be done exactly as in the last part of the proof of Theorem 8.4 in [18]. Then we divide by nT_o, let $n \to \infty$, take $\inf_{\alpha\in\Gamma} \sup_{b\in\mathcal{B}}$, and finally let $\varepsilon \to 0$ to reach the conclusion.

Finally, the statement about the upper value (3) is obtained by observing that

$$-u - \mathrm{val}\, J^\infty(x) = \inf_{\beta\in\Delta} \sup_{a\in\mathcal{A}} \liminf_{T\to\infty} (-J(T,x))$$

and the left-hand side is the lower value of the LTAC game with running cost $-l$ where the second player wishes to minimize and the first player to maximize. Then the conclusion follows by applying (2) to this game. □

We recall the classical Isaacs condition, or solvability of the small game,

$$\begin{aligned} H(y,p) &:= \min_{b\in B} \max_{a\in A} \{-f(y,a,b)\cdot p - l(y,a,b)\} \\ &= \max_{a\in A} \min_{b\in B} \{-f(y,a,b)\cdot p - l(y,a,b)\}, \quad \forall y, p \in \mathbb{R}^m. \end{aligned} \tag{7}$$

It is well known that it implies the equality of the upper and the lower values of the finite horizon game, that is, the existence of the value of that game, which we denote with val $J(T,x)$, see [9, 17]. Therefore, we immediately get the following consequence of Theorem 1.1.

Corollary 1.1. *Assume* (7) *and that either the lower or the upper game is ergodic. Then the LTAC game is ergodic, i.e.,*

$$l - \mathrm{val}\, J_\infty(x) = u - \mathrm{val}\, J^\infty(x) = \lim_{T\to\infty} \mathrm{val}\, J(T,x) = \lambda, \quad \forall x \in \mathbb{R}^m.$$

2 Characterizations of Ergodicity

From now on we add periodicity to the standing assumptions:

$$\begin{aligned} &f(y,a,b) = f(y+k,a,b), \quad l(y,a,b) = l(y+k,a,b), \\ &\qquad \forall k \in \mathbb{Z}^m,\ y \in \mathbb{R}^m,\ a \in A,\ b \in B. \end{aligned} \tag{8}$$

This means that the state space is the m-torus $\mathbb{T}^m = \mathbb{R}^m/\mathbb{Z}^m$. The first result is a consequence of Theorem 4 in [3].

Theorem 2.1. *The following statements on the lower game are equivalent.*

(i) The lower game is ergodic, i.e., $l - \mathrm{val}\, J(T,x) \to \mathrm{const}$ *uniformly in* x *as* $T \to +\infty$.

(ii) $l - \mathrm{val}\, \delta \int_0^\infty l(y_x(t), a(t), b(t))\, e^{-\delta t}\, dt \to \mathrm{const}$ *uniformly in* x *as* $\delta \to 0+$.

(iii) The additive eigenvalue problem

$$\lambda + \min_{b \in B} \max_{a \in A} \{-f(y,a,b) \cdot \nabla\chi - l(y,a,b)\} = 0 \quad \text{in } \mathbb{R}^m, \quad \chi\ \mathbb{Z}^m\text{-periodic} \tag{9}$$

has the property that

$$\begin{aligned} &\sup\{\lambda \mid \text{there is a viscosity subsolution of } (9)\} \\ &\quad = \inf\{\lambda \mid \text{there is a viscosity supersolution of } (9)\}. \end{aligned} \tag{10}$$

If one of the above assertions is true, then the constants in (i) *and* (ii) *are equal and they coincide with the number defined by* (10)*. Moreover, the same result holds for the upper game, after replacing* $l - \mathrm{val}$ *with* $u - \mathrm{val}$ *in* (i) *and* (ii)*, and* (9) *with*

$$\lambda + \max_{a \in A} \min_{b \in B} \{-f(y,a,b) \cdot \nabla\chi - l(y,a,b)\} = 0 \quad \text{in } \mathbb{R}^m, \quad \chi\ \mathbb{Z}^m\text{-periodic}.$$

Proof. If we set $v(t,y) := l - \mathrm{val}\, J(t,y)$ it is well known [9, 17] that $w(t,y) := tv(t,y)$ is the viscosity solution of the Cauchy problem for the Isaacs equation

$$w_t + H(y, D_y w) = 0 \quad \text{in } (0,+\infty) \times \mathbb{R}^m, \qquad w(0,y) = 0, \quad w\ \mathbb{Z}^m\text{-periodic}.$$

The equivalence of (iii) and the uniform convergence of $w(t,\cdot)/t$ to a constant as $t \to \infty$ is stated in Theorem 4 of [3], and it gives the equivalence of (i) and (iii).

Next, $w_\delta(x) := l - \mathrm{val} \int_0^\infty l(y_x(t), a(t), b(t))\, e^{-\delta t}\, dt$ is the viscosity solution of the Isaacs equation

$$\delta w_\delta + H(y, Dw_\delta) = 0 \quad \text{in } \mathbb{R}^m, \qquad w_\delta\ \mathbb{Z}^m\text{-periodic}, \tag{11}$$

and Theorem 4 of [3] states the equivalence of (iii) and the uniform convergence of δw_δ to a constant as $\delta \to 0+$. Therefore, (ii) and (iii) are equivalent.

The equality of the three constants is also given by Theorem 4 of [3]. Finally, the proof for the upper value is the same, with the Hamiltonian $H = \min\max$ replaced by $\max\min$. □

Remark 2.1. Note that (ii) deals with a vanishing discount rate problem for infinite horizon games. The equivalence between (i) and (ii) is a differential game extension of the classical Abelian–Tauberian theorem, stating that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \varphi(t)\, dt = \lim_{\delta \to 0+} \delta \int_0^\infty \varphi(t)\, e^{-\delta t}\, dt$$

whenever one of the two limits exists. The property (iii) is a characterization of the uniform ergodicity of the lower game by a Hamilton–Jacobi–Isaacs equation. In some cases the inf and the sup in the formula (10) are attained, and the number defined by (10) is the unique constant λ such that the additive eigenvalue problem (9) has a continuous viscosity solution, see Remark 3.1. In general, however, even if (iii) holds, (9) may have no continuous solution χ (see Arisawa and Lions [7]). By analogy with the theory of homogenization we call (9) the *cell problem*.

Whenever the conditions of Corollary 1.1 for the ergodicity of the LTAC game are satisfied, we have the following informations on the value of the game, namely, the constant λ.

Proposition 2.1. *Assume* (7) *and that either the lower or the upper game is ergodic. Then* $\lambda = l - \text{val}\, J_\infty(x) = u - \text{val}\, J^\infty(x)$ *satisfies*

$$\min_x \min_{a\in A} \max_{b\in B} l(x,a,b) \le \lambda \le \max_x \min_{a\in A} \max_{b\in B} l(x,a,b).$$

If, moreover,

$$\max_{a\in A} \min_{b\in B}\{-f(x,a,b)\cdot p - l(x,a,b)\} \ge \max_{a\in A} \min_{b\in B}\{-l(x,a,b)\} \quad \forall x, p \in \mathbb{R}^m \tag{12}$$

(respectively, $\le$*), then*

$$\lambda = \min_x \min_{a\in A} \max_{b\in B} l(x,a,b) \tag{13}$$

(respectively, $\lambda = \max_x \min_a \max_b l(x,a,b)$*).*

Proof. First we use the characterization (i) of ergodicity in Theorem 2.1, and we set $v(t,y) := l - \text{val}\, J(t,y)$, $w(t,y) := tv(t,y)$. It is well known [9, 17] that w satisfies, in the viscosity sense,

$$w_t + H(y, D_y w) = 0 \quad \text{in } (0,+\infty)\times\mathbb{R}^m, \qquad w(0,y)=0, \quad w \text{ periodic}.$$

We observe that $-t\max_y H(y,0)$ and $-t\min_y H(y,0)$ are, respectively, a sub- and a supersolution of this Cauchy problem. Therefore, the comparison principle gives

$$-t\max_y H(y,0) \le w(t,y) \le -t\min_y H(y,0).$$

We divide by t and let $t \to +\infty$. Since $w(t,y)/t \to \lambda$, and $H(y,0) = \max_a \min_b\{-l(x,a,b)\}$ by (7), we get the first pair of inequalities.

To prove the second statement we assume by contradiction that $\lambda > -H(y,0)$ in a neighborhood of a minimum point of $-H(y,0) = \min_a \max_b l(x,a,b)$. Now we use the characterization (ii) of ergodicity in Theorem 2.1, as in the proof of Theorem 2.1. With the same notation, the value function w_δ of the infinite horizon discounted game satisfies the Isaacs equation (11). By the uniform convergence of δw_δ to λ we get

$$H(y, Dw_\delta) - H(y, 0) = -\lambda - H(y, 0) + o(1) < 0 \quad \text{as } \delta \to 0$$

in an open set. This is a contradiction to the assumption (12). □

Remark 2.2. Note that, for a running cost independent of the controls, $l = l(x)$, the condition (12) reads

$$\min_a \max_b f(x, a, b) \cdot p = \max_b \min_a f(x, a, b) \cdot p \le 0, \tag{14}$$

in view of (7). This says that the first player has a stronger control on the vector field than the second one. The conclusion is that the LTAC value is

$$\lambda = \min_x l(x),$$

so the minimizing player can drive asymptotically the system near the minimum points of the running cost.

Next we describe some sufficient conditions for the ergodicity of the upper or the lower game. We say that the system (1) is *bounded-time controllable by the first player* if there exists $S > 0$ and for each $x, \widetilde{x} \in \mathbb{R}^m$ there exists a strategy $\widetilde{\alpha} \in \Gamma$ such that for all control functions $b \in \mathcal{B}$ there is a time $t^\# = t^\#(x, \widetilde{x}, \widetilde{\alpha}, b)$ with the properties

$$t^\# \le S \quad \text{and} \quad y_x(t^\#) - \widetilde{x} \in \mathbb{Z}^m, \tag{15}$$

where $y_x(\cdot)$ is the trajectory corresponding to the strategy $\widetilde{\alpha}$ and the control function b, i.e., it solves

$$\dot{y}(t) = f(y(t), \widetilde{\alpha}[b](t), b(t)), \qquad y(0) = x. \tag{16}$$

In other words, the first player can drive the system from any initial position x to any given state $\widetilde{x}$ on the torus $\mathbb{T}^m$ in a uniformly bounded time for all possible behaviors of the second player. Symmetrically, we say that the system (1) is *bounded-time controllable by the second player* if for some $S > 0$ and for all $x, \widetilde{x} \in \mathbb{R}^m$ there is a strategy $\widetilde{\beta} \in \Delta$ such that for all control functions $a \in \mathcal{A}$

$$\exists\, t^\# = t^\#(x, \widetilde{x}, a, \widetilde{\beta}) \le S \text{ such that } y_x(t^\#) - \widetilde{x} \in \mathbb{Z}^m$$

where $y_x(\cdot)$ is the trajectory corresponding to the strategy $\widetilde{\beta}$ and the control function a, i.e., it solves

$$\dot{y}(t) = f(y(t), a(t), \widetilde{\beta}[a](t)), \qquad y(0) = x.$$

For systems with a single player this notion is studied in the literature under various names such as *complete controllability* [14], *uniform exact controllability* [6], and *total controllability* [8].

Theorem 2.2. *If the system* (1) *is bounded-time controllable by the first player (respectively, by the second player), then the lower game (respectively, the upper game) is ergodic.*

Proof. The proof uses the characterization (ii) of ergodicity in Theorem 2.1, and we call w_δ the lower value function of the discounted infinite horizon problem, namely,

$$w_\delta(x) := \inf_{\alpha\in\Gamma} \sup_{b\in\mathcal{B}} \int_0^\infty l(y_x(t), \alpha[b](t), b(t))\, e^{-\delta t}\, dt.$$

The main tool of the proof is the following dynamic programming principle due to Soravia, Remark 4.2 of [33],

$$w_\delta(x) = \inf_{\alpha\in\Gamma} \sup_{b\in\mathcal{B}} \inf_{0\le t<\infty} \left\{ \int_0^t l(y(s), \alpha[b](s), b(s)) e^{-\delta s}\, ds + e^{-\delta t} w_\delta(y(t)) \right\}, \tag{17}$$

where $y(\cdot)$ is the trajectory of (16) with $\widetilde{\alpha}$ replaced by a generic α.

For fixed $x, \widetilde{x}$ we take a strategy $\widetilde{\alpha} \in \Gamma$ such that (15) holds. Then (17) and the periodicity of w_δ give

$$w_\delta(x) \le \sup_{b\in\mathcal{B}} \left\{ \int_0^{t^\#} l(y(s), \alpha[b](s), b(s)) e^{-\delta s}\, ds + e^{-\delta t^\#} w_\delta(\widetilde{x}) \right\},$$

where $y(\cdot)$ is the trajectory of (16). Since l and δw_δ are uniformly bounded, there is a constant C such that

$$\delta w_\delta(x) - \delta w_\delta(\widetilde{x}) \le C(1 - e^{-\delta S}). \tag{18}$$

Now we exchange the roles of x and $\widetilde{x}$ to get

$$\lim_{\delta\to 0+} |\delta w_\delta(x) - \delta w_\delta(\widetilde{x})| = 0 \quad \text{uniformly in } x, \widetilde{x} \in \mathbb{R}^m.$$

If for fixed $\widetilde{x}$ we choose a sequence $\delta_k \to 0$ such that $\delta_k w_{\delta_k}(\widetilde{x}) \to \mu$, we obtain the uniform convergence of $\delta_k w_{\delta_k}$ to μ.

We claim that μ is independent of the sequence δ_k. This implies the uniform convergence of the whole net δw_δ to μ, as desired. To prove the claim we recall the cell problem (9), i.e.,

$$\lambda + H(y, D\chi) = 0 \quad \text{in } \mathbb{R}^m, \qquad \chi \text{ periodic}, \tag{19}$$

where λ is a constant, and use the inequality

$$\begin{aligned} &\lambda_1 := \sup\{\lambda \mid \exists \text{ a u.s.c. subsolution of (19)}\} \\ &\le \lambda_2 := \inf\{\lambda \mid \exists \text{ a l.s.c. supersolution of (19)}\}, \end{aligned}$$

which follows from a standard argument based on the comparison principle for sub- and supersolutions of Hamilton–Jacobi equations (see, e.g., the proof of Theorem 1 in [6] or that of Theorem 4 in [3]). The Isaacs equation satisfied by w_δ in the viscosity sense is (see, e.g., [9])

$$\delta w_\delta + H(y, Dw_\delta) = 0 \quad \text{in } \mathbb{R}^m, \qquad w_\delta \text{ periodic.} \tag{20}$$

Then, for $\lambda < \mu$, w_{δ_k} is a subsolution of (19) for k large enough, so $\mu \le \lambda_1$. The same argument gives $\lambda_2 \le \mu$. Therefore, $\mu = \lambda_1 = \lambda_2$, which proves the claim. □

An immediate consequence of this theorem and of Corollary 1.1 is the following.

Corollary 2.1. *Assume the Isaacs' condition* (7) *and that the system* (1) *is bounded-time controllable either by the first or by the second player. Then the LTAC game is ergodic, i.e.,*

$$l - \text{val}\, J_\infty(x) = u - \text{val}\, J^\infty(x) = \lim_{T\to\infty} \text{val}\, J(T, x) = \lambda, \quad \forall x \in \mathbb{R}^m.$$

We end this section with some sufficient conditions for ergodicity that are symmetric for the two players, unlike the preceding Theorem 2.2 and Corollary 2.1 where one of the two players has a much stronger hold of the system than the other. We take a system of the form

$$\begin{cases} \dot y^A(t) = f_A(y(t), a(t), b(t)), & y^A(0) = x^A \in \mathbb{R}^{m_A}, \\ \dot y^B(t) = f_B(y(t), a(t), b(t)), & y^B(0) = x^B \in \mathbb{R}^{m_B}, \\ y(t) = (y^A(t), y^B(t)), \end{cases} \tag{21}$$

and we assume that the state variables y^A are *(uniformly) asymptotically controllable* by the first player, whereas the variables y^B are asymptotically controllable by the second, in the following sense. There exists a function $\eta : [0, \infty) \to [0, \infty)$ with

$$\lim_{T\to\infty} \eta(T) = 0, \tag{22}$$

and for all $x^A, \widetilde{x}^A \in \mathbb{R}^{m_A}$, $x^B \in \mathbb{R}^{m_B}$, there is a strategy $\widetilde{\alpha} \in \Gamma$ such that, for $x = (x^A, x^B)$,

$$\frac{1}{T}\int_0^T \min_{k^A \in \mathbb{Z}^{m_A}} |y_x^A(t) - \widetilde{x}^A - k^A|\, dt \le \eta(T), \ \forall b \in \mathcal{B}, \tag{23}$$

whereas for all $x^B, \widetilde{x}^B \in \mathbb{R}^{m_B}$, $x^A \in \mathbb{R}^{m_A}$, there is a strategy $\widetilde{\beta} \in \Delta$ such that

$$\frac{1}{T}\int_0^T \min_{k^B \in \mathbb{Z}^{m_B}} |y_x^B(t) - \widetilde{x}^B - k^B|\, dt \le \eta(T), \ \forall a \in \mathcal{A}. \tag{24}$$

Note that the integrand in (23) is the distance between $y_x^A(t)$ and $\widetilde{x}^A$ on the m_A-dimensional torus $\mathbb{T}^{m_A} = \mathbb{R}^{m_A}/\mathbb{Z}^{m_A}$, so (23) and (22) mean that the first player can drive asymptotically y^A near $\widetilde{x}^A$, uniformly with respect to x, $\widetilde{x}^A$, and the control of the other player b. Similarly, (24) says that the second player can drive asymptotically y^B to $\widetilde{x}^B$ on the m_B-dimensional torus $\mathbb{T}^{m_B}$, uniformly with respect to x, $\widetilde{x}^B$, and a.

We will also assume that the running cost does not depend on the controls, $l = l(y^A, y^B)$, and it has a saddle point in $[0, 1]^{m_A} \times [0, 1]^{m_B}$, that is, it satisfies

$$\min_{y^A \in \mathbb{R}^{m_A}} \max_{y^B \in \mathbb{R}^{m_B}} l(y^A, y^B) = \max_{y^B \in \mathbb{R}^{m_B}} \min_{y^A \in \mathbb{R}^{m_A}} l(y^A, y^B) =: \bar{l}. \tag{25}$$

Proposition 2.2. *Assume the system* (1) *is of the form* (21) *with y^A and y^B asymptotically controllable, respectively, by the first and by the second player. Suppose also that $l = l(y^A, y^B)$ satisfies* (25) *and* (7) *holds. Then the LTAC game is ergodic and its value is the value of the static game with payoff l, that is,*

$$l - \text{val}\, J_\infty(x^A, x^B) = u - \text{val}\, J^\infty(x^A, x^B) = \bar{l}, \quad \forall (x^A, x^B) \in \mathbb{R}^m. \tag{26}$$

Proof. We first prove that, for all $x = (x^A, x^B)$,

$$\limsup_{T\to\infty} l - \text{val}\, J(T, x) \le \min_{y^A \in \mathbb{R}^{m_A}} \max_{y^B \in \mathbb{R}^{m_B}} l(y^A, y^B) \quad \text{uniformly in } x \in \mathbb{R}^m. \tag{27}$$

To this goal we fix $x^A, \widetilde{x}^A, x^B$ and consider the strategy $\widetilde{\alpha}$ from the asymptotic controllability assumption. If $y_x(\cdot) = y_x(\cdot, b)$ is the corresponding trajectory, from the periodicity of l we get

$$|l(y_x^A(t), y_x^B(t)) - l(\widetilde{x}^A, y_x^B(t))| \le \omega_l(\min_{k^A \in \mathbb{Z}^{m_A}} |y_x^A(t) - \widetilde{x}^A - k^A|),$$

where ω_l is the modulus of continuity of l, i.e.,

$$|l(x) - l(y)| \le \omega_l(|x - y|),\ \forall x, y \in \mathbb{R}^m, \quad \lim_{r\to 0} \omega_l(r) = 0.$$

We recall that it is not restrictive to assume the concavity of ω_l, so Jensen's inequality implies

$$\frac{1}{T}\int_0^T |l(y_x^A(t), y_x^B(t)) - l(\widetilde{x}^A, y_x^B(t))|\, dt \le \omega_l(\eta(T)).$$

Then

$$\begin{aligned} l - \text{val}\, J(T, x) &\le \sup_{b\in\mathcal{B}} \frac{1}{T}\int_0^T l(y_x^A(t), y_x^B(t))\, dt \\ &\le \sup_{b\in\mathcal{B}} \frac{1}{T}\int_0^T l(\widetilde{x}^A, y_x^B(t))\, dt + \omega_l(\eta(T)) \\ &\le \max_{y^B \in \mathbb{R}^{m_B}} l(\widetilde{x}^A, y^B) + \omega_l(\eta(T)). \end{aligned}$$

Now we take the $\limsup_{T\to\infty}$ of both sides and finally the min over $\widetilde{x}^A \in \mathbb{R}^{m_A}$ to get (27).

Next we observe that a symmetric proof gives

$$\liminf_{T\to\infty} u - \text{val}\, J(T,x) \geq \max_{y^B\in\mathbb{R}^{m_B}} \min_{y^A\in\mathbb{R}^{m_A}} l(y^A, y^B) \quad \text{uniformly in } x \in \mathbb{R}^m.$$

By the Isaacs condition (7) $l - \text{val}\, J(T,x) = u - \text{val}\, J(T,x)$, so the assumptions (25) on l gives

$$\lim_{T\to\infty} \text{val}\, J(T,x) = \bar{l} \quad \forall\, x \in \mathbb{R}^m.$$

The conclusion now follows from Theorem 1.1. □

Remark 2.3. If the system governing y^A is bounded-time controllable by the first player and also *stoppable*, i.e.,

$$\forall x \in \mathbb{R}^m,\ \forall b \in B,\ \exists a \in A\ :\ f_A(x,a,b) = 0,$$

then the variables y^A are asymptotically controllable, because $\widetilde{x}^A$ can be reached from x^A in a time smaller than S and then the first player can keep $y^A(t) = \widetilde{x}^A$ for all later times t. In this case, if $\bar{l} = l(\widetilde{x}^A, \widetilde{x}^B)$, an optimal strategy for the first player amounts to driving the variables y^A to the saddle point $\widetilde{x}^A$ and stopping there, and the strategy of going to $\widetilde{x}^B$ and staying there forever is optimal for the second player. This kind of behavior is called a *turnpike*, see [12, 13].

3 Examples

Example 1. first-order controllability. Assume that for some $\nu > 0$

$$B(0,\nu;m) \subset \overline{\text{conv}}\{f(x,a,b) \mid a \in A\}, \quad \forall x \in \mathbb{R}^m, b \in B, \tag{28}$$

where $B(0,\nu;m)$ denotes the m-dimensional open ball of radius ν centered at the origin and $\overline{\text{conv}}$ the closed convex hull. From the standard theory of differential games (see, for instance, Corollary 3.7 in [32]) it is known that the system is (small-time) controllable by the first player and the time necessary to reach a point $\widetilde{x}$ from x satisfies an estimate of the form

$$t^{\#}(x, \widetilde{x}, \widetilde{\alpha}, b) \leq \frac{C}{\nu}|x - \widetilde{x}|.$$

Therefore, the lower game is uniformly ergodic in this case. Moreover, if $l = l(x)$ it is easy to see that (14) holds, so $\lambda = \min_x l(x)$.

Example 2. higher-order controllability. Consider a system of the form

$$\dot{y}(t) = \sum_{i=1}^{k-1} a_i(t) g^i(y(t)) + a_k(t) g^k(y(t), a(t), b(t)), \tag{29}$$

where the control of the first player $a = (a_1, \dots, a_k)$ varies in a neighborhood of the origin $A \subset \mathbb{R}^k$, and all g^i with $i \le k-1$ are C^∞ vector fields in $\mathbb{R}^m$. Moreover, we suppose the full rank (Hörmander) condition on $g^1, \dots, g^{k-1}$, that is, *the vector fields* $g^1, \dots, g^{k-1}$ *and their commutators of any order span* $\mathbb{R}^m$ *at each point of* $\mathbb{R}^m$. By choosing $a_k \equiv 0$ we obtain a symmetric system independent of the second player. Then the classical Chow's theorem of geometric control theory says that this system is small-time locally controllable at all points of the state space. Moreover, for any small $t > 0$ the reachable set from x in time t is a neighborhood of x, and the same holds for the reachable set backward in time. From this, using the compactness of the torus $\mathbb{T}^m$, it is easy to see that the whole state space is an invariant control set in the terminology of [14]. Then the global bounded-time controllability follows from Lemma 3.2.21 in [14]. In conclusion, the full system (29) is bounded-time controllable by the first player, and therefore the lower game is uniformly ergodic. As in the previous example, if λ is independent of the controls (14) holds and $\lambda = \min_x l(x)$.

Remark 3.1. If (15) holds with $t^\#(x, \widetilde{x}, \widetilde{\alpha}, b) \le \omega(|x - \widetilde{x}|)$ for all $|x - \widetilde{x}| \le \gamma$ and all $b \in \mathcal{B}$, for some modulus ω and $\gamma > 0$, we say that (1) is also *small-time controllable* by the first player. For such systems there exists a continuous solution χ to the additive eigenvalue problem (9), by Proposition 9.2 in [4]. The systems of Examples 1 and 2 are indeed small-time controllable by the first player.

Example 3. (separate controllability). For a system of the form (21) we can assume that the subsystem for the variables y^A either satisfies

$$B(0, \nu; m_A) \subset \overline{\mathrm{conv}}\, f_A(x, A, b) \quad \forall x \in \mathbb{R}^m, b \in B,$$

or it is of the form

$$\dot{y}^A = \sum_{i=1}^{k_A - 1} a_i g_A^i(y^A) + a_{k_A} g_A^{k_A}(y, a, b),$$

where the control of the first player $a = (a_1, \dots, a_{k_A})$ varies in a neighborhood of the origin $A \subset \mathbb{R}^{k_A}$, and the vector fields $g_A^1, \dots, g_A^{k_A - 1}$ are of class C^∞ and satisfy the full rank condition in $\mathbb{R}^{m_A}$. Then the variables y^A are asymptotically controllable because the first player can drive them from x^A to $\widetilde{x}^A$ in bounded time and then stop there by choosing the null control. Similarly, the variables y^B are asymptotically controllable if either f_B verifies

$$B(0, \nu; m_B) \subset \overline{\mathrm{conv}}\, f_B(x, a, B) \quad \forall x \in \mathbb{R}^m, a \in A,$$

or it is of the form

$$\dot{y}^B = \sum_{i=1}^{k_B - 1} b_i g_B^i(y^B) + b_{k_B} g_B^{k_B}(y, a, b),$$

where the control of the second player $b = (b_1, \dots, b_{k_B})$ varies in a neighborhood of the origin $B \subset \mathbb{R}^{k_B}$, and the vector fields $g_B^1, \dots, g_B^{k_B-1}$ are of class C^∞ and satisfy the full rank condition in $\mathbb{R}^{m_B}$. Under these conditions and with $l = l(y^A, y^B)$ satisfying (25), Proposition 2.2 implies the ergodicity of the LTAC game and the formula (26).

4 Ergodicity of Noisy Systems

In this section we study the ergodicity of the lower value for the following class of stochastic differential games. We consider the controlled diffusion process

$$dy(t) = f(y(t), a(t), b(t))dt + \sigma(y(t), a(t), b(t))dW(t), \quad y(0) = x, \quad (30)$$

where W is an r-dimensional Brownian motion, and σ is a continuous map from $\mathbb{R}^m \times A \times B$ to the space of $m \times r$ matrices, Lipschitzian in x uniformly in a, b. The finite horizon cost functional is

$$J(T, x) = J(T, x, a(\cdot), b(\cdot)) := E\left[\frac{1}{T}\int_0^T l(y_x(t), a(t), b(t))\, dt\right],$$

where E denotes the expectation. An admissible control for the first (respectively, second) player is a progressively measurable function of time taking values in A (respectively, B), and we still denote with $\mathcal{A}$ and $\mathcal{B}$ the sets of admissible controls. We also keep the notation Γ and Δ for the set of nonanticipating strategies for the first player and the second player, respectively, and we refer to [19] for the precise definitions in the stochastic setting. The lower value of the finite horizon game is

$$v(T, x) := l - \text{val}\, J(T, x) := \inf_{\alpha\in\Gamma} \sup_{b\in\mathcal{B}} J(T, x, \alpha[b], b)$$

and we say that *the lower game is ergodic* if

$$l - \text{val}\, J(T, \cdot) \to \lambda \quad \text{as } T \to \infty \text{ uniformly in } \mathbb{R}^m.$$

In addition to the periodicity in the state variable of f and l (8), we assume

$$\sigma(y, a, b) = \sigma(y + k, a, b), \quad \forall\, k \in \mathbb{Z}^m,\ y \in \mathbb{R}^m,\ a \in A,\ b \in B.$$

We are going to extend all the results of Section 2 to this setting, and we also present a theorem where the ergodicity is due to the effects of the diffusion without any controllability hypothesis. Analogous results hold for the upper game, which is defined in the obvious way, but we will not state them explicitly.

We begin with the stochastic counterpart of the Abelian–Tauberian-type Theorem 2.1, which is again a consequence of Theorem 4 in [3]. We will use the second-order Hamiltonian

$$H(y, p, X) := \min_{b\in B} \max_{a\in A} \left\{-\frac{1}{2}\text{trace}(\sigma\sigma^T(y, a, b)X) - f(y, a, b)\cdot p - l(y, a, b)\right\}$$

for $y, p \in \mathbb{R}^m$ and X any symmetric $m \times m$ matrix.

Theorem 4.1. *The following statements are equivalent.*

(i) The lower game is ergodic, i.e., $v(T, x) \to \text{const}$ *uniformly in* x *as* $T \to +\infty$.
(ii) $l - \text{val}\, E[\delta \int_0^\infty l(y_x(t), a(t), b(t))\, e^{-\delta t}\, dt] \to \text{const}$ *uniformly in* x *as* $\delta \to 0+$.
(iii) The cell problem

$$\lambda + H(y, \nabla\chi, D^2\chi) = 0 \quad \text{in } \mathbb{R}^m, \quad \chi\ \mathbb{Z}^m\text{-periodic} \tag{31}$$

has the property that

$$\begin{aligned} &\sup\{\lambda \mid \text{there is a viscosity subsolution of (31)}\} \\ &\quad = \inf\{\lambda \mid \text{there is a viscosity supersolution of (31)}\}. \end{aligned} \tag{32}$$

If one of the above assertions is true, then the constants in (i) *and* (ii) *are equal and they coincide with the number defined by* (32).

Proof. The proof is the same as that of Theorem 2.1 after recalling that $w(t, y) := t v(t, y)$ is the viscosity solution of

$$w_t + H(y, D_y w, D^2_{yy} w) = 0 \quad \text{in } (0, +\infty) \times \mathbb{R}^m, \qquad w(0, y) = 0, \tag{33}$$

and $w_\delta(x) := l - \text{val}\, E\left[\int_0^\infty l(y_x(t), a(t), b(t))\, e^{-\delta t}\, dt\right]$ is the viscosity solution of

$$\delta w_\delta + H(y, Dw_\delta, D^2 w_\delta) = 0 \quad \text{in } \mathbb{R}^m, \tag{34}$$

see [19]. □

Proposition 4.1. *Assume the lower game is ergodic.*
Then $\lambda = \lim_{T\to\infty} v(T, x)$ *satisfies*

$$\min_x \max_{b\in B} \min_{a\in A} l(x, a, b) \le \lambda \le \max_x \max_{b\in B} \min_{a\in A} l(x, a, b).$$

If, moreover,

$$\begin{aligned} &\min_{b\in B} \max_{a\in A} \left\{ -\frac{1}{2} trace \left(\sigma\sigma^T(y, a, b)X\right) - f(x, a, b) \cdot p - l(x, a, b) \right\} \\ &\quad \ge \min_{b\in B} \max_{a\in A} \{-l(x, a, b)\} \quad \forall x, p, X \end{aligned}$$

(respectively, $\le$*), then*

$$\lambda = \min_x \max_{b\in B} \min_{a\in A} l(x, a, b) \tag{35}$$

(respectively, $\lambda = \max_x \max_b \min_a l(x, a, b)$*).*

Proof. The proof is the same as that of Proposition 2.1, after observing that

$$-H(y, 0, 0) = \max_{b \in B} \min_{a \in A} l(x, a, b).$$

□

It is well known that nondegenerate diffusion processes are ergodic (in the standard sense). The next result states the ergodicity of games involving a controlled system affected by a nondegenerate diffusion, with no controllability assumptions. It is a consequence of Theorem 7.1 in our paper [4], see also [3]. Earlier related results are due to Evans [15] and Arisawa and Lions [7].

Theorem 4.2. *Assume that the minimal eigenvalue of the matrix $\sigma\sigma^T(y, a, b)$ is positive for all $y \in \mathbb{R}^m$, $a \in A$, $b \in B$. Then the lower game is ergodic.*

Proof. The periodicity in y of σ and the compactness of A and B imply that the minimal eigenvalue of the matrix $\sigma\sigma^T$ is bounded below by a positive constant. Therefore, the second-order partial differential operator associated with the Hamiltonian H is uniformly elliptic. Then Theorem 7.1 in [4] shows that there exists a (unique) constant λ such that the cell problem (31) has a continuous viscosity solution χ. Then $u(t, y) := \lambda t + \chi(y)$ solves

$$u_t + H(y, D_y u, D^2_{yy} u) = 0 \quad \text{in } (0, +\infty) \times \mathbb{R}^m, \qquad u(0, y) = \chi(y).$$

Since $w(t, y) := tv(t, y)$ solves (33), by the comparison principle for viscosity solutions of Cauchy problems we get

$$u(t, y) - \max \chi \le tv(t, y) \le u(t, y) + \min \chi \quad \forall t, y.$$

Therefore, $v(t, \cdot) \to \lambda$ uniformly as $t \to \infty$. □

The next result is an extension to stochastic games of Theorem 2.2. We say that the system (30) is *bounded-time controllable by the first player* if for some $S > 0$ and for all $x, \widetilde{x} \in \mathbb{R}^m$ there is a strategy $\widetilde{\alpha} \in \Gamma$ such that for all admissible control functions $b \in \mathcal{B}$

$$\exists\, t^\# = t^\#(x, \widetilde{x}, \widetilde{\alpha}, b) \le S \text{ such that } y_x(t^\#) - \widetilde{x} \in \mathbb{Z}^m \text{ almost surely,} \tag{36}$$

where $y_x(\cdot)$ is the solution of (30) corresponding to the controls $\widetilde{\alpha}[b]$ and b.

Theorem 4.3. *If the system* (30) *is bounded-time controllable by the first player, then the lower game is ergodic.*

Proof. We follow the argument of the proof of Theorem 2.2. We use the characterization (ii) of ergodicity in Theorem 4.1, and we call w_δ the lower value function of the discounted infinite horizon problem, namely,

$$w_\delta(x) := \inf_{\alpha \in \Gamma} \sup_{b \in \mathcal{B}} E\left[\int_0^\infty l(y_x(t), \alpha[b](t), b(t))\, e^{-\delta t}\, dt\right].$$

The main tool of the proof is the following dynamic programming principle for stochastic games due to Swiech, Corollary 2.6 (iii) of [37],

$$w_\delta(x) = \inf_{\alpha\in\Gamma} \sup_{b\in\mathcal{B}} \inf_{0\le t<\infty} E\left\{\int_0^t l(y_x(s),\alpha[b](s),b(s))e^{-\delta s}\,ds + e^{-\delta t}w_\delta(y_x(t))\right\}. \tag{37}$$

For fixed $x, \widetilde{x}$ we take a strategy $\widetilde{\alpha} \in \Gamma$ such that (36) holds. Then (37) and the periodicity of w_δ give

$$w_\delta(x) \le \sup_{b\in\mathcal{B}} E\left\{\int_0^{t^\#} l(y_x(s),\widetilde{\alpha}[b](s),b(s))e^{-\delta s}\,ds + e^{-\delta t^\#}w_\delta(\widetilde{x})\right\},$$

where $y_x(\cdot)$ is the trajectory of (30) with the controls $\widetilde{\alpha}[b]$ and b. Now the proof of Theorem 2.2 shows that, along a sequence $\delta_k \to 0$, $\delta_k w_{\delta_k} \to \mu$ uniformly.

Finally, the proof that μ does not depend on the sequence δ_k is also the same as in Theorem 2.2, with the new cell problem (31) and using the equation (34) satisfied by w_δ. □

The last result of the section is a stochastic counterpart of Proposition 2.2. We take a controlled diffusion of the form

$$\begin{cases} dy^A(t) = f_A(y(t),a(t),b(t))dt + \sigma_A(y(t),a(t),b(t))dW_A(t), \\ dy^B(t) = f_B(y(t),a(t),b(t))dt + \sigma_B(y(t),a(t),b(t))dW_B(t), \\ y(t) = (y^A(t),y^B(t)),\ \ y^A(0) = x^A \in \mathbb{R}^{m_A},\ \ y^B(0) = x^B \in \mathbb{R}^{m_B}, \end{cases} \tag{38}$$

and we assume that the state variables y^A are *asymptotically controllable* by the first player, and the variables y^B are asymptotically controllable by the second, in the following sense. There exists a function $\eta : [0,\infty) \to [0,\infty)$ with $\lim_{T\to\infty} \eta(T) = 0$, and for all $x^A, \widetilde{x}^A \in \mathbb{R}^{m_A}$, $x^B \in \mathbb{R}^{m_B}$, there is a strategy $\widetilde{\alpha} \in \Gamma$, such that, for $x = (x^A, x^B)$,

$$E\left[\frac{1}{T}\int_0^T \min_{k^A\in\mathbb{Z}^{m_A}} |y_x^A(t) - \widetilde{x}^A - k^A|\,dt\right] \le \eta(T),\ \ \forall b \in \mathcal{B}, \tag{39}$$

whereas for all $x^B, \widetilde{x}^B \in \mathbb{R}^{m_B}$, $x^A \in \mathbb{R}^{m_A}$, there is a strategy $\widetilde{\beta} \in \Delta$ such that

$$E\left[\frac{1}{T}\int_0^T \min_{k^B\in\mathbb{Z}^{m_B}} |y_x^B(t) - \widetilde{x}^B - k^B|\,dt\right] \le \eta(T),\ \ \forall a \in \mathcal{A}.$$

As in Section 2 we assume that the running cost does not depend on the controls and has a saddle point. The Isaacs condition now is

$$\begin{aligned} &\min_{b\in B}\max_{a\in A}\left\{-\frac{1}{2}\mathrm{trace}\left(\sigma\sigma^T(y,a,b)X\right) - f(y,a,b)\cdot p - l(y,a,b)\right\} \\ &\quad = \max_{a\in A}\min_{b\in B}\left\{-\frac{1}{2}\mathrm{trace}\left(\sigma\sigma^T(y,a,b)X\right) - f(y,a,b)\cdot p - l(y,a,b)\right\}. \end{aligned} \tag{40}$$

Proposition 4.2. *Assume the system* (30) *is of the form* (38) *with* y^A *and* y^B *asymptotically controllable, respectively, by the first and by the second player. Suppose also* $l = l(y^A, y^B)$ *satisfies* (25) *and* (40) *holds. Then the lower game is ergodic and its value converges to the value of the static game with payoff* l*, that is,*

$$\lim_{T\to\infty} v(T, x^A, x^B) = \bar{l}, \quad \textit{uniformly in } (x^A, x^B) \in \mathbb{R}^m. \tag{41}$$

Proof. The proof is essentially the same as that of Proposition 2.2. We begin with

$$\limsup_{T\to\infty} v(T, x) \le \min_{y^A\in\mathbb{R}^{m_A}} \max_{y^B\in\mathbb{R}^{m_B}} l(y^A, y^B) \quad \text{uniformly in } x \in \mathbb{R}^m. \tag{42}$$

We fix x^A, $\widetilde{x}^A$, x^B and consider the strategy $\widetilde{\alpha}$ from the asymptotic controllability assumption. If $y_x(\cdot)$ is the corresponding trajectory, then, by the periodicity of l, (39), and the concavity of ω_l,

$$E\left[\frac{1}{T}\int_0^T |l(y_x^A(t), y_x^B(t)) - l(\widetilde{x}^A, y_x^B(t))|\, dt\right] \le \omega_l(\eta(T)).$$

Then

$$\begin{aligned} l - \text{val}\, J(T, x) &\le \sup_{b\in\mathcal{B}} E\left[\frac{1}{T}\int_0^T l(y_x^A(t), y_x^B(t))\, dt\right] \\ &\le \max_{y^B\in\mathbb{R}^{m_B}} l(\widetilde{x}^A, y^B) + \omega_l(\eta(T)). \end{aligned}$$

Now we take the $\limsup_{T\to\infty}$ of both sides and finally the min over $\widetilde{x}^A \in \mathbb{R}^{m_A}$ to get (42). A symmetric proof gives

$$\liminf_{T\to\infty} u - \text{val}\, J(T, x) \ge \max_{y^B\in\mathbb{R}^{m_B}} \min_{y^A\in\mathbb{R}^{m_A}} l(y^A, y^B) \quad \text{uniformly in } x \in \mathbb{R}^m,$$

where $u - \text{val}\, J(T, x)$ denotes the upper value of the finite horizon game. By the Isaacs condition (40), the results of Fleming and Souganidis [19] imply $v(T, x) = l - \text{val}\, J(T, x) = u - \text{val}\, J(T, x)$ for all T, x, so the assumption (25) on l gives (41). □

Remark 4.1. The proof of Proposition 4.2 shows also that the upper game is ergodic and the upper value $u - \text{val}\, J(T, x)$ converges uniformly to the saddle $\bar{l}$ as $T \to \infty$.

Remark 4.2. If the system governing y^A is bounded-time controllable by the first player and also *stoppable*, i.e.,

$$\forall x \in \mathbb{R}^m,\ \forall b \in B,\ \exists a \in A\ :\ f_A(x, a, b) = 0,\ \sigma_A(x, a, b) = 0,$$

then the variables y^A are asymptotically controllable, because $\widetilde{x}^A$ can be reached from x^A in a time smaller than S and then the first player can keep $y^A(t) = \widetilde{x}^A$ for all later times t. As in the deterministic case, an optimal strategy for each player is a turnpike: driving the system to a saddle point and stopping there.

Acknowledgments

Olivier Alvarez was partially supported by the research project ACI-JC 1025 “Dislocation dynamics" of the French Ministry of Education. Martino Bardi was partially supported by the research project “Viscosity, metric, and control theoretic methods for nonlinear partial differential equations” of the Italian M.I.U.R. We thank Wendell Fleming and Bill McEneaney for some remarks on their paper [18], and Pierre Cardaliaguet for useful comments on a preliminary version of this work.

References

[1] O. Alvarez, *Homogenization of Hamilton-Jacobi equations in perforated sets*, J. Differential Equations **159** (1999), 543–577.

[2] O. Alvarez and M. Bardi, *Viscosity solutions methods for singular perturbations in deterministic and stochastic control*, SIAM J. Control Optim. **40** (2001), 1159–1188.

[3] O. Alvarez and M. Bardi, *Singular perturbations of degenerate parabolic PDEs: a general convergence result*, Arch. Rational Mech. Anal. **170** (2003), 17–61.

[4] O. Alvarez and M. Bardi, *Ergodicity, stabilization, and singular perturbations of Bellman-Isaacs equations*, Preprint, University of Padova, to appear.

[5] O. Alvarez, M. Bardi, and C. Marchi, *Multiscale problems and homogenization for second-order Hamilton-Jacobi equations*, Preprint, University of Padova, submitted.

[6] M. Arisawa, *Ergodic problem for the Hamilton-Jacobi-Bellman equation II*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), 1–24.

[7] M. Arisawa and P.-L. Lions, *On ergodic stochastic control*, Comm. Partial Differential Equations **23** (1998), 2187–2217.

[8] Z. Artstein and V. Gaitsgory, *The value function of singularly perturbed control systems*, Appl. Math. Optim. **41** (2000), 425–445.

[9] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, Boston, 1997.

[10] A. Bensoussan, *Perturbation methods in optimal control*, Wiley/Gauthiers-Villars, Chichester, 1988.

[11] R. Buckdahn and N. Ichihara, *Limit theorems for controlled backward SDEs and homogenization of Hamilton-Jacobi-Bellman equations*, Appl. Math. Optim. **51** (2005), 1–33.

[12] D. A. Carlson and A. B. Haurie, *A turnpike theory for infinite horizon open-loop differential games with decoupled controls*, New trends in dynamic games and applications (G.J. Olsder, ed.), Ann. Internat. Soc. Dynam. Games, no. 3, Birkhäuser Boston, 1995, pp. 353–376.

[13] D. A. Carlson, A. B. Haurie, and A. Leizarowitz, *Infinite horizon optimal control: Deterministic and stochastic systems*, Springer-Verlag, Berlin, 1991.

[14] F. Colonius and W. Kliemann, *The dynamics of control*, Birkhäuser, Boston, 2000.

[15] L. Evans, *The perturbed test function method for viscosity solutions of nonlinear P.D.E.*, Proc. Roy. Soc. Edinburgh Sect. A **111** (1989), 359–375.

[16] L. Evans, *Periodic homogenisation of certain fully nonlinear partial differential equations*, Proc. Roy. Soc. Edinburgh Sect. A **120** (1992), 245–265.

[17] L. Evans and P. E. Souganidis, *Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations*, Indiana Univ. Math. J. **33** (1984), 773–797.

[18] W. H. Fleming and W. M. McEneaney, *Risk-sensitive control on an infinite time horizon*, SIAM J. Control Optim. **33** (1995), 1881–1915.

[19] W. H. Fleming and P. E. Souganidis, *On the existence of value functions of two-players, zero-sum stochastic differential games*, Indiana Univ. Math. J. **38** (1989), 293–314.

[20] V. Gaitsgory, *Limit Hamilton-Jacobi-Isaacs equations for singularly perturbed zero-sum differential games*, J. Math. Anal. Appl. **202** (1996), 862–899.

[21] M. K. Ghosh and K. S. M. Rao, *Differential games with ergodic payoff*, SIAM J. Control Optim. **43** (2005), 2020–2035.

[22] L. Grüne, *On the relation between discounted and average optimal value functions*, J. Differential Equations **148** (1998), 65–99.

[23] H. Ishii, *Homogenization of the Cauchy problem for Hamilton-Jacobi equations*, Stochastic analysis, control, optimization and applications. A volume in honor of Wendell H. Fleming (W.M. McEneaney, G. Yin, and Q. Zhang, eds.), Birkhäuser, Boston, 1999, pp. 305–324.

[24] Y. Kabanov and S. Pergamenshchikov, *Optimal control of singularly perturbed linear stochastic systems*, Stochastics Stochastics Rep. **36** (1991), 109–135.

[25] P. V. Kokotović, H. K. Khalil, and J. O'Reilly, *Singular perturbation methods in control: analysis and design*, Academic Press, London, 1986.

[26] H. J. Kushner, *Weak convergence methods and singularly perturbed stochastic control and filtering problems*, Birkhäuser, Boston, 1990.

[27] H. J. Kushner, *Numerical approximations for stochastic differential games: the ergodic case*, SIAM J. Control Optim. **42** (2004), 1911–1933.

[28] P.-L. Lions, *Neumann type boundary conditions for Hamilton-Jacobi equations*, Duke Math. J. **52** (1985), 793–820.

[29] P.-L. Lions, G. Papanicolaou, and S. R. S. Varadhan, *Homogenization of Hamilton-Jacobi equations*, Unpublished, 1986.

[30] C. Marchi, *Homogenization for fully nonllinear parabolic equations*, Nonlinear Anal. **60** (2005), 411–428.

[31] M. Quincampoix and F. Watbled, *Averaging method for discontinuous Mayer's problem of singularly perturbed control systems*, Nonlinear Anal. **54** (2003), 819–837.

[32] P. Soravia, *Pursuit-evasion problems and viscosity solutions of Isaacs equations*, SIAM J. Control Optim. **31** (1993), 604–623.

[33] P. Soravia, *Stability of dynamical systems with competitive controls: the degenerate case*, J. Math. Anal. Appl. **191** (1995), 428–449.

[34] S. Sorin, *New approaches and recent advances in two-person zero-sum repeated games*, Advances in dynamic games (A.S. Nowak and K. Szajowski, eds.), Ann. Internat. Soc. Dynam. Games, no. 7, Birkhäuser, Boston, 2005, pp. 67–93.

[35] N. N. Subbotina, *Asymptotic properties of minimax solutions of Isaacs-Bellman equations in differential games with fast and slow motions*, J. Appl. Math. Mech. **60** (1996), 883–890.

[36] N. N. Subbotina, *Asymptotics for singularly perturbed differential games*, Game theory and applications, Vol. VII, Nova Sci. Publ., Huntington, NY, 2001, pp. 175–196.

[37] A. Swiech, *Another approach to the existence of value functions of stochastic differential games*, J. Math. Anal. Appl. **204** (1996), 884–897.

Subgame Consistent Solutions for a Class of Cooperative Stochastic Differential Games with Nontransferable Payoffs

David W. K. Yeung
Department of Decision Sciences
Hong Kong Baptist University
Kowloon Tong
Kowloon, Hong Kong
and
Centre for Game Theory
St. Petersburg State University
198904 Petrodvorets, St. Petersburg, Russia
wkyeung@hkbu.edu.hk

Leon Petrosyan
Faculty of Applied Mathematics-Control Processes
St. Petersburg State University
198904 Petrodvorets, St. Petersburg, Russia
spbuoasis7@peterlink.ru

Patricia Melody Yeung
Faculty of Arts and Science
University of Toronto
Toronto, ON, M5S 2E4, Canada
yeung_patricia@hotmail.com

Abstract

Subgame consistency is a fundamental element in the solution of cooperative stochastic differential games. In particular, it ensures that the extension of the solution policy to a later starting time and any possible state brought about by prior optimal behavior of the players will remain optimal. Recently, mechanisms for the derivation of subgame consistent solutions in stochastic cooperative differential games with transferable payoffs have been found. In the case when players' payoffs are nontransferable, the derivation of solution candidates is extremely complicated and often intractable. In this chapter, subgame consistent solutions are derived for a class of cooperative stochastic differential games with nontransferable payoffs.

Key words. Cooperative stochastic differential games, subgame consistent solution, nontransferable payoffs, rationality.

1 Introduction

A fundamental element in the theory of cooperative games is the formulation of an optimal behavior for the players satisfying specific optimality principles. In particular, the solution of a cooperative game is produced by a set of optimality principles (for instance, the Nash bargaining solution [8] and the Shapley value [13]). For dynamic games, a stringent condition on their solutions is required: The specific optimality principle must remain optimal at any instant of time throughout the game along the optimal state trajectory chosen at the outset. This condition is known as *dynamic stability or time consistency*. In particular, the dynamic stability of a solution is the property that, when the game proceeds along an optimal trajectory, at each instant of time the players are guided by the same optimality principle, and hence do not have any ground for deviation from the previously adopted optimal behavior throughout the game. The issue of dynamic stability in differential games has been explored rigorously in the past three decades (see Haurie [4], Petrosyan [9]–[11], Petrosyan and Danilov [12], and Yeung and Petrosyan [15]).

In the presence of stochastic elements, a more stringent condition—*subgame consistency*—is required for a credible cooperative solution. A cooperative solution is subgame consistent if an extension of the solution policy to a situation with a later starting time and any feasible state brought about by prior optimal behaviors would remain optimal. In the field of cooperative stochastic differential games, little research has been published to date due to the inherent difficulties in deriving tractable subgame consistent solutions. Haurie et al. [5] derived cooperative equilibria of a stochastic differential game of fishery with the use of a monitoring variable and an associated retaliation scheme. Recently, Yeung and Petrosyan [16] developed a mechanism for the derivation of analytically tractable subgame consistent solutions in stochastic differential games with transferable payoffs. In the case when payoffs are nontransferable, the solution mechanism becomes extremely complicated and intractable. In Yeung and Petrosyan [17] a subgame consistent solution is developed for cooperative stochastic differential games with nontransferable payoffs for the first time.

In this chapter, subgame consistent solutions are derived for a class of cooperative stochastic differential games with nontransferable payoffs. The complexity of stochastic differential games leads to great difficulties in the derivation of their solutions. The stringent requirement of subgame consistency imposes additional hurdles to the derivation of solutions for cooperative stochastic differential games. In the case when players' payoffs are nontransferable, the derivation of solution candidates becomes even more complicated and intractable. The highly intractable subgame consistent solution for such games is rendered tractable in this study.

The chapter is organized as follows. Section 2 presents the basic game formulation. Section 3 derives the Pareto optimal trajectory and individual players' payoff under cooperation. Section 4 examines the notion of subgame consistency. In Section 5, subgame consistent solutions games are presented. Numerical examples are provided in Section 6 and concluding remarks are given in Section 7.

2 Game Formulation

Consider a two-person nonzero-sum stochastic differential game with initial state x_0 and duration $T - t_0$.

2.1 State Dynamics and Payoff Structures

The state space of the game is $X \subset R$, with permissible state trajectories $\{x(s), t_0 \le s \le T\}$. The state dynamics of the game is characterized by the stochastic differential equations

$$dx(s) = [a - bx(s) - u_1(s) - u_2(s)]ds + \sigma x(s)dz(s), \quad x(t_0) = x_0 \in X, \quad (1)$$

where $u_i \in U_i$ is the control vector of player i, for $i \in [1, 2]$, a, b, and σ are positive constants, and $z(s)$ is a Wiener process. Equation (1) could be interpreted as the stock dynamics of a biomass of a renewable resource like forest or fresh water. The state $x(s)$ represents the resource size and $u_i(s)$ the (nonnegative) amount of resource extracted by player i.

At time t_0, the expected payoff of player $i \in [1, 2]$ is

$$J^i(t_0, x_0) = E_{t_0}\left\{\int_{t_0}^{T} [h_i u_i(s) - c_i u_i(s)^2 x(s)^{-1} + k_i x(s)] \exp[-r(s - t_0)]ds \right.$$
$$\left. \exp[-r(T - t_0)]q_i x(T) \,\middle|\, x(t_0) = x_0 \right\}, \quad \text{for } i \in [1, 2], \quad (2)$$

where h_i, c_i, k_i, and q_i are positive parameters.

The term $h_i u_i(s)$ reflects player i's satisfaction level obtained from the consumption of the resource extracted, and $c_i u_i(s)^2 x(s)^{-1}$ measures the (dis)satisfaction level created in the extraction process. $k_i x(s)$ is the satisfaction to player i brought about by the positive environmental effect of the state. The total utility of player i is the aggregate level of satisfaction. Payoffs in the form of utility are not transferable between players. There exists a time discount rate r, and the utility received at time t has to be discounted by the factor $\exp[-r(t - t_0)]$. At time T, player i will receive a termination satisfaction $q_i x(T)^{1/2}$, where q_i is nonnegative.

2.2 Noncooperative Outcome

We use $\Gamma(x_0, T - t_0)$ to denote the game (1)–(2) and $\Gamma(x_\tau, T - \tau)$ to denote an alternative game with state dynamics (1) and payoff structure (2), which starts at time $\tau \in [t_0, T]$ with initial state $x_\tau \in X$. A non-cooperative Nash equilibrium solution of the game $\Gamma(x_\tau, T - \tau)$ can be characterized with the techniques of Isaacs [6], Bellman [1], and Fleming [3] as follows.

Definition 2.1. A set of feedback strategies $\{u_i^{(\tau)*}(t) = \phi_i^{(\tau)*}(t, x)$, for $i \in [1, 2]\}$ provides a Nash equilibrium solution to the game $\Gamma(x_\tau, T-\tau)$, if there exist twice continuously differentiable functions $V^{(\tau)i}(t, x) : [\tau, T] \times R \to R$, $i \in [1, 2]$, satisfying the following partial differential equations:

$$\begin{aligned}
&-V_t^{(\tau)i}(t, x) - \frac{1}{2}\sigma^2 x^2 V_{xx}^{(\tau)i}(t, x) \\
&\quad = \max_{u_i} \left\{ [h_i u_i - c_i u_i^2 x^{-1} + k_i x] \exp[-r(t-\tau)] \right. \\
&\qquad \left. + V_x^{(\tau)i}(t, x)[a - bx - u_i - u_j] \right\}, \text{ and} \\
&V^{(\tau)i}(T, x) = \exp[-r(T-\tau)] q_i x, \text{ for } i \in [1, 2], \ j \in [1, 2], \text{ and } j \neq i. \qquad (3)
\end{aligned}$$

Performing the indicated maximization in Definition 2.1 yields

$$\phi_i^{(\tau)*}(t, x) = \frac{\left[h_i - V_x^{(\tau)i} \exp(r(t-\tau))\right] x}{2c_i}, \quad \text{for } i \in [1, 2] \text{ and } x \in X. \qquad (4)$$

Proposition 2.1. *The value function of player i in the game $\Gamma(x_\tau, T-\tau)$ is*

$$V^{(\tau)i}(t, x) = \exp[-r(t-\tau)][A_i(t)x + B_i(t)], \quad \textit{for } i \in [1, 2] \textit{ and } t \in [\tau, T], \qquad (5)$$

where $A_i(t)$, $B_i(t)$, $A_j(t)$, and $B_j(t)$, for $i \in [1, 2]$ and $j \in [1, 2]$ and $i \neq j$, satisfy

$$\begin{aligned}
\dot{A}_i(t) &= (r+b)A_i(t) - k_i - \frac{[h_i - A_i(t)]^2}{4c_i} + \frac{A_i(t)[h_j - A_j(t)]}{2c_j}, \\
\dot{B}_i(t) &= rB_i(t) - aA_i(t), \\
A_i(T) &= q_i, \quad B_i(T) = 0.
\end{aligned}$$

Proof. Substitution of $\phi_i^{(\tau)*}(t, x)$ from (4) into (3) yields a set of partial differential equations. One can readily verify that (5) is a solution to this set of equations. □

3 Cooperative Scheme

Under cooperation the players negotiate to establish an agreement (optimality principle) on how to play the cooperative game and hence how to distribute the resulting payoff. A set of optimality principles that governs the optimal behavior for the players constitutes a solution. In particular, the solution optimality principle for a cooperative game $\Gamma_c(x_0, T-t_0)$ includes

(i) an agreement on a set of cooperative strategies $[\psi_1^{(t_0)*}(t, x), \psi_2^{(t_0)*}(t, x)]$, for $s \in [t_0, T]$, and
(ii) a mechanism to distribute the total payoff among players.

In nontransferable payoff games, the players' imputations are directly dictated by the agreed-upon cooperative strategies. A necessary condition is that this optimality principle must satisfy group rationality and individual rationality.

3.1 Pareto Optimal Trajectories

To achieve group rationality, the Pareto optimality of outcomes must be validated. Let $\Gamma_c(x_0, T - t_0)$ denote a cooperative game with payoff structure (1) and dynamics (2) starting at time t_0 with initial state x_0. Pareto optimal trajectories for $\Gamma_c(x_0, T - t_0)$ can be identified by choosing a weight $\alpha_1 \in (0, \infty)$ that solves the following stochastic control problem (see Leitmann [7] and Dockner and Jørgensen [2] for analyses in cooperative differential games):

$$\begin{aligned}
&\max_{u_1,u_2}\{J^1(t_0,x_0)+\alpha_1 J^2(t_0,x_0)\}\\
&\quad\equiv \max_{u_1,u_2} E_{t_0}\left\{\int_{t_0}^{T}([h_1u_1(s)-c_1u_1(s)^2x(s)^{-1}+k_1x(s)]\right.\\
&\quad\quad+\alpha_1[h_2u_2(s)-c_2u_2(s)^2x(s)^{-1}+k_2x(s)])\exp[-r(s-t_0)]ds\\
&\quad\quad\left.\exp[-r(T-t_0)][q_1x(T)+q_2x(T)]\,\middle|\, x(t_0)=x_0\right\},
\end{aligned} \tag{6}$$

subject to dynamics (1). Note that when $\alpha_1 = 1/\alpha_2$, the problem $\max_{u_1,u_2}\{J^1(t_0,x_0)+\alpha_1 J^2(t_0,x_0)\}$ is identical to the problem $\max_{u_1,u_2}\{J^2(t_0,x_0)+\alpha_2 J^1(t_0,x_0)\}$ in the sense that $\max_{u_1,u_2}\{J^2(t_0,x_0)+\alpha_2 J^1(t_0,x_0)\} \equiv \max_{u_1,u_2}\{\alpha_2[J^1(t_0,x_0)+\alpha_1 J^2(t_0,x_0)]\}$ yields the same optimal controls as those from $\max_{u_1,u_2}\{J^1(t_0,x_0)+\alpha_1 J^2(t_0,x_0)\}$.

In $\Gamma_c(x_0, T - t_0)$, let α_1^0 be the selected weight according the agreed-upon optimality principle. Invoking the technique developed by Fleming [3], we have the following.

Definition 3.1. A set of controls $\left\{\left[\psi_1^{\alpha_1^0(t_0)}(t,x), \psi_2^{\alpha_1^0(t_0)}(t,x)\right], t \in \left[t_0, T\right]\right\}$ provides an optimal solution to the stochastic control problem $\max_{u_1,u_2}\{J^1(t_0,x_0)+\alpha_1^0 J^2(t_0,x_0)\}$, if there exists the twice continuously differentiable function $W^{\alpha_1^0(t_0)}(t,x) : [t_0,T]\times R \to R$ satisfying the partial differential equation

$$\begin{aligned}
&-W_t^{\alpha_1^0(t_0)}(t,x)-\frac{1}{2}\sigma^2x^2W_{xx}^{\alpha_1^0(t_0)}(t,x)\\
&\quad=\max_{u_1,u_2}\big\{([h_1u_1-c_1u_1^2x^{-1}+k_1x]\\
&\quad\quad+\alpha_1^0[h_2u_2-c_2u_2^2x^{-1}+k_2x])\exp[-r(t-t_0)]\\
&\quad\quad+W_x^{\alpha_1^0(t_0)}(t,x)[a-bx-u_i-u_j]\big\},
\end{aligned}$$

$$W^{\alpha_1^0(t_0)}(T,x)=\exp[-r(T-t_0)][q_1x(T)+\alpha_1^0q_2x(T)]. \tag{7}$$

Performing the indicated maximization in Definition 3.1 yields

$$\psi_1^{\alpha_1^0(t_0)}(t,x)=\frac{\left[h_1-W_x^{\alpha_1^0(t_0)}(t,x)\exp(r(t-t_0))\right]x}{2c_1},$$

and

$$\psi_2^{\alpha_1^0(t_0)}(t,x) = \frac{\left[\alpha_1^0 h_2 - W_x^{\alpha_1^0(t_0)}(t,x)\exp(r(t-t_0))\right]x}{2\alpha_1^0 c_2}, \quad \text{for } t \in [t_0, T]. \quad (8)$$

Proposition 3.1. *The value function of the control problem* $\max_{u_1,u_2}$ $\{J^1(t_0,x_0) + \alpha_1^0 J^2(t_0,x_0)\}$ *is*

$$W^{\alpha_1^0(t_0)}(t,x) = \exp[-r(t-t_0)]\left[A^{\alpha_1^0}(t)x + B^{\alpha_1^0}(t)\right], \quad \textit{for } t \in [t_0,T], \quad (9)$$

where $A^{\alpha_1^0}(t)$ *and* $B^{\alpha_1^0}(t)$ *satisfy*

$$\dot{A}^{\alpha_1^0}(t) = (r+b)A^{\alpha_1^0}(t) - \frac{\left[h_1 - A^{\alpha_1^0}(t)\right]^2}{4c_1} - \frac{\left[\alpha_1^0 h_2 - A^{\alpha_1^0}(t)\right]^2}{4\alpha_1^0 c_2} - k_1 - k_2,$$

$$\dot{B}^{\alpha_1^0}(t) = rB^{\alpha_1^0}(t) - A^{\alpha_1^0}(t)a,$$

$$A^{\alpha_1^0}(T) = q_1 + \alpha_1^0 q_2 \ \textit{and}\ B^{\alpha_1^0}(T) = 0. \quad (10)$$

Proof. Substitution of $\psi_1^{\alpha_1^0(t_0)}(t,x)$ and $\psi_2^{\alpha_1^0(t_0)}(t,x)$ from (8) into (7) yields a partial differential equation. One can readily verify that (9) is a solution to this set of equations. □

Substituting the partial derivatives $W_x^{\alpha_1^0(t_0)}(t,x)$ into $\psi_1^{\alpha_1^0(t_0)}(t,x)$ and $\psi_2^{\alpha_1^0(t_0)}(t,x)$ in (9) yields the optimal controls of the problem $\max_{u_1,u_2}\{J^1(t_0,x_0) + \alpha_1^0 J^2(t_0,x_0)\}$ as

$$\psi_1^{\alpha_1^0(t_0)}(t,x) = \frac{\left[h_1 - A^{\alpha_1^0}(t)\right]x}{2c_1},$$

and

$$\psi_2^{\alpha_1^0(t_0)}(t,x) = \frac{\left[\alpha_1^0 h_2 - A^{\alpha_1^0}(t)\right]x}{2\alpha_1^0 c_2}, \quad \text{for } t \in [t_0,T]. \quad (11)$$

Substituting these controls into (1) yields the dynamics of the Pareto optimal trajectory associated with a weight α_1^0. The Pareto optimal trajectory then can be solved as

$$x^{\alpha_1^0(t_0)}(t) = \left\{\Phi(\alpha_1^0;t,t_0)\left[x_0 + \int_{t_0}^t \Phi^{-1}(\alpha_1^0;s,t_0)a\,ds\right]\right\}^2, \quad (12)$$

where

$$\Phi(\alpha_1^0;t,t_0) = \exp\left[\int_{t_0}^t \left(-b - \frac{h_1 - A^{\alpha_1^0}(s)}{2c_1} - \frac{\alpha_1 h_2 - A^{\alpha_1^0}(s)}{2\alpha_1^0 c_2} - \frac{\sigma^2}{2}\right)ds + \int_{t_0}^t \sigma\,dz(s)\right].$$

We use $X_t^{\alpha_1^0(t_0)}$ to denote the set $x^{\alpha_1^0(t_0)}(t)$ realizable values generated by (12) at $t \in (t_0, T]$.

Now, consider the cooperative game $\Gamma_c(x_\tau, T - \tau)$ with state dynamics (1) and payoff structure (2), which starts at time $\tau \in [t_0, T]$ with initial state $x_\tau \in X_\tau^{\alpha_1(t_0)}$. Let α_1^τ be the selected weight according to the agreed-upon optimality principle.

Following the previous analysis, we can obtain the value function, optimal controls, and optimal trajectory of the control problem $\max_{u_1,u_2}\{J^1(\tau, x_\tau) + \alpha_1^\tau J^2(\tau, x_\tau)\}$.

Remark 3.1. One can readily show that when $\alpha_1^0 = \alpha_1^\tau = \alpha_1^*$, then $\psi_i^{\alpha_1^*(t_0)}(t, x_t) = \psi_i^{\alpha_1^*(\tau)}(t, x_t)$ at the point (t, x_t), for $i \in [1, 2]$, $t_0 \le \tau \le t \le T$, and $x_t \in X_t^{\alpha_1^*(t_0)}$.

3.2 Individual Player's Payoff Under Cooperation

In order to verify individual rationality, we have to derive the players' expected payoffs in the cooperative game $\Gamma_c(x_\tau^*, T - \tau)$. Let α_1^τ be the weight dictated by the solution optimality principle. Following Yeung [14], we substitute

$$\psi_1^{\alpha_1^\tau(\tau)}(t, x) = \frac{\left[h_1 - A^{\alpha_1^\tau}(t)\right]x}{2c_1}$$

and

$$\psi_2^{\alpha_1^\tau(\tau)}(t, x) = \frac{\left[\alpha_1^\tau h_2 - A^{\alpha_1^\tau}(t)\right]x}{2\alpha_1^\tau c_2}$$

into the players' payoff functions and we denote these functions as follows.

Definition 3.2. Given that $x(t) = x_t^{\alpha_1^\tau(\tau)} \in X_t^{\alpha_1^\tau(\tau)}$, for $t \in [\tau, T]$, player 1's expected payoff over the interval $[t, T]$ under the control problem $\max_{u_1,u_2}\{J^1(\tau, x_\tau) + \alpha_1^\tau J^2(\tau, x_\tau)\}$ is given as

$$\begin{aligned}
&\hat{W}^{\tau(\alpha_1^\tau)1}(t, x) \\
&\quad = E_\tau \left\{ \int_t^T \left[\frac{h_1\left[h_1 - A^{\alpha_1^\tau}(s)\right]x(s)}{2c_1} - \frac{\left[h_1 - A^{\alpha_1^\tau}(s)\right]^2 x(s)}{4c_1} + k_1 x(s) \right] \right. \\
&\qquad \left. \times \exp[-r(s - \tau)]ds \ + \exp[-r(T - t_0)]q_1 x(T) \,\middle|\, x(t) = x \right\},
\end{aligned}$$

and the corresponding expected payoff of player 2 over the interval $[t, T]$ as

$$\hat{W}^{\tau(\alpha_1^\tau)2}(t,x)$$
$$= E_\tau \left\{ \int_t^T \left[\frac{h_2\big[\alpha_1^\tau h_2 - A^{\alpha_1^\tau}(s)\big]x(s)}{2\alpha_1^\tau c_2} - \frac{\big[\alpha_1^\tau h_2 - A^{\alpha_1^\tau}(s)\big]^2 x(s)}{4(\alpha_1^\tau)^2 c_2} + k_i x(s) \right] \right.$$
$$\left. \times \exp[-r(s-\tau)]ds \ + \exp[-r(T-\tau)]q_2 x(T) \middle| \ x(t) = x \right\},$$

where

$$dx(s) = \left[a - bx(s) - \frac{\big[h_1 - A^{\alpha_1^\tau}(s)\big]x(s)}{2c_1} - \frac{\big[\alpha_1^\tau h_2 - A^{\alpha_1^\tau}(s)\big]x(s)}{2\alpha_1^\tau c_2} \right] ds$$
$$+ \sigma x(s) dz(s), \quad x(t) = x.$$

Note that when $\Delta t \to 0$, we can express $\hat{W}^{\tau(\alpha_1^\tau)1}(t, x_\tau)$ as

$$\hat{W}^{\tau(\alpha_1^\tau)1}(t,x)$$
$$= E_\tau \left\{ \int_t^{t+\Delta t} \left[\frac{h_1\big[h_1 - A^{\alpha_1^\tau}(s)\big]x(s)}{2c_1} - \frac{c_1\big[h_1 - A^{\alpha_1^\tau}(s)\big]^2 x(s)}{4c_1^2} + k_1 x(s) \right] \right.$$
$$\left. \times \exp[-r(s-\tau)]\, ds \ + \hat{W}^{\tau(\alpha_1^\tau)1}(t+\Delta t, x_t + \Delta x) \middle| \ x(t) = x \right\}. \quad (13)$$

Using Taylor's expansion, we obtain

$$\hat{W}^{\tau(\alpha_1^\tau)1}(t,x)$$
$$= E_\tau \left\{ \left[\frac{h_1\big[h_1 - A^{\alpha_1^\tau}(t)\big]x_t}{2c_1} - \frac{c_1\big[h_1 - A^{\alpha_1^\tau}(t)\big]^2 x_t}{4c_1^2} + k_1 x_t \right] \right.$$
$$\times \exp[-r(t-\tau)]\,\Delta t + \hat{W}^{\tau(\alpha_1^\tau)1}(t,x) + \hat{W}_t^{\tau(\alpha_1^\tau)1}(t,x)\Delta t$$
$$+ \hat{W}_x^{\tau(\alpha_1^\tau)1}(t,x) \left[a - bx - \frac{\big[h_1 - A^{\alpha_1^\tau}(t)\big]x}{2c_1} - \frac{\big[\alpha_1^\tau h_2 - A^{\alpha_1^\tau}(t)\big]x}{2\alpha_1^\tau c_2} \right] \Delta t$$
$$\left. + \hat{W}_x^{\tau(\alpha_1^\tau)1}(t,x)\sigma x \Delta z + \frac{1}{2}\hat{W}_{xx}^{\tau(\alpha_1^\tau)1}(t,x_t)\,\sigma^2 x^2 \Delta t + o(\Delta t) \right\}, \quad (14)$$

where

$$\Delta x = \left[a - bx - \frac{\left[h_1 - A^{\alpha_1^\tau}(t)\right]x}{2c_1} - \frac{\left[\alpha_1^\tau h_2 - A^{\alpha_1^\tau}(t)\right]x}{2\alpha_1^\tau c_2} \right] \Delta t + \sigma x \Delta z + o\,(\Delta t),$$

$$x(t) = x_t \in X, \quad \Delta z = z\,(t + \Delta t) - z(t),$$
$$\text{and } E_\tau\,[o\,(\Delta t)]\,/\Delta t \to 0 \text{ as } \Delta t \to 0. \tag{15}$$

Canceling terms, performing the expectation operator, dividing throughout by Δt, and taking $\Delta t \to 0$ yield

$$\begin{aligned}
&-\hat{W}_t^{\tau(\alpha_1^\tau)1}(t, x_t) - \frac{1}{2}\hat{W}_{x_t x_t}^{\tau(\alpha_1^\tau)1}(t, x_t)\sigma^2 x_t^2 \\
&\quad = \left[\frac{h_1\left[h_1 - A^{\alpha_1^\tau}(t)\right]x_t}{2c_1} - \frac{c_1\left[h_1 - A^{\alpha_1^\tau}(t)\right]^2 x_t}{4c_1^2} + k_1 x_t \right] \exp[-r(t-\tau)] \\
&\qquad + \hat{W}_{x_t}^{\tau(\alpha_1^\tau)1}(t, x_t)\left[a - bx_t - \frac{\left[h_1 - A^{\alpha_1^\tau}(t)\right]x_t}{2c_1} - \frac{\left[\alpha_1^\tau h_2 - A^{\alpha_1^\tau}(t)\right]x_t}{2\alpha_1^\tau c_2} \right].
\end{aligned} \tag{16}$$

Boundary conditions require

$$\hat{W}^{\tau(\alpha_1^\tau)1}\,(T, x) = \exp[-r(T-\tau)]q_1 x. \tag{17}$$

If there exist continuously differentiable functions $\hat{W}^{\tau(\alpha_1^\tau)1}(t, x)$:$[\tau, T] \times R \to R$ satisfying (16) and (17), then player 1's expected payoff in the cooperative game $\Gamma(x_\tau, T-\tau)$ under the cooperation scheme with weight α_1^τ is indeed $\hat{W}^{\tau(\alpha_1^\tau)1}(t, x)$.

Proposition 3.2. *The function $\hat{W}^{\tau(\alpha_1^\tau)1}\,(t, x) : [\tau, T] \times R \to R$ satisfying (16) and (17) can be solved as*

$$\hat{W}^{\tau(\alpha_1^\tau)1}\,(t, x) = \exp[-r(t-\tau)]\left[\hat{A}_1^{\alpha_1^\tau}(t)x + \hat{B}_1^{\alpha_1^\tau}(t)\right], \tag{18}$$

where $\hat{A}_1^{\alpha_1^\tau}(t)$ and $\hat{B}_1^{\alpha_1^\tau}(t)$ satisfy

$$\begin{aligned}
\dot{\hat{A}}_1^{\alpha_1^\tau}(t) &= \left[r + b + \frac{\left[h_1 - A^{\alpha_1^\tau}(t)\right]}{2c_1} + \frac{\left[\alpha_1 h_2 - A^{\alpha_1^\tau}(t)\right]}{2\alpha_1 c_2} \right] \hat{A}_1^{\alpha_1}(t) \\
&\quad - \frac{\left[h_1 - A^{\alpha_1^\tau}(t)\right]\left[h_1 + A^{\alpha_1^\tau}(t)\right]}{4c_1} - k_1, \\
\dot{\hat{B}}_1^{\alpha_1^\tau}(t) &= r\hat{B}_1^{\alpha_1^\tau}(t) - a\hat{A}_1^{\alpha_1^\tau}(t), \quad \hat{A}_1^{\alpha_1^\tau}(T) = q_1 \;\; \textit{and} \;\; \hat{B}_1^{\alpha_1^\tau}(T) = 0.
\end{aligned}$$

Proof. Calculating the derivatives $\hat{W}_t^{\tau(\alpha_1^\tau)1}\,(t, x)$, $\hat{W}_{xx}^{\tau(\alpha_1^\tau)1}\,(t, x)$, and $\hat{W}_x^{\tau(\alpha_1^\tau)1}\,(t, x)$ from (18) and then substituting them into (16) yield Proposition 3.2. □

Following a similar analysis, a continuously differentiable function $\hat{W}^{\tau(\alpha_1^\tau)2} \times (t, x) : [\tau, T] \times R \rightarrow R$ giving player 2's expected payoff under the control problem $\max_{u_1,u_2}\{J^1(\tau, x_\tau) + \alpha_1^\tau J^2(\tau, x_\tau)\}$ can be obtained as follows.

Proposition 3.3. *The function $\hat{W}^{\tau(\alpha_1^\tau)2}\ (t, x) : [\tau, T] \times R \rightarrow R$ can be solved as*

$$\hat{W}^{\alpha_1^\tau(\tau)2}(t, x) = \exp[-r(t-\tau)]\big[\hat{A}_2^{\alpha_1^\tau}(t)x + \hat{B}_2^{\alpha_1^\tau}(t)\big], \tag{19}$$

where $\hat{A}_2^{\alpha_1^\tau}(t)$ and $\hat{B}_2^{\alpha_1^\tau}(t)$ has to satisfy

$$\begin{aligned}\dot{\hat{A}}_2^{\alpha_1^\tau}(t) &= \left[r + b + \frac{\big[h_1 - A^{\alpha_1^\tau}(t)\big]}{2c_1} + \frac{\big[\alpha_1 h_2 - A^{\alpha_1^\tau}(t)\big]}{2\alpha_1 c_2}\right]\hat{A}_2^{\alpha_1^\tau}(t)\\ &\quad - \frac{\big[\alpha_1 h_2 - A^{\alpha_1^\tau}(t)\big]\big[\alpha_1 h_2 + A^{\alpha_1^\tau}(t)\big]}{4\alpha_1^2 c_2} - k_2,\\ \dot{\hat{B}}_2^{\alpha_1^\tau}(t) &= r\hat{B}_2^{\alpha_1^\tau}(t) - a\hat{A}_2^{\alpha_1^\tau}(t), \quad \hat{A}_2^{\alpha_1^\tau}(T) = q_2 \ \ \textit{and} \ \ \hat{B}_2^{\alpha_1^\tau}(T) = 0.\end{aligned}$$

Proof. Follow the proof of Proposition 3.2. □

4 Notion of Subgame Consistency

Under cooperation with nontransferable payoffs, the players negotiate to establish an agreement (optimality principle) on how to play the cooperative game and how to distribute the resulting payoff. In particular, the chosen optimality principle must satisfy group rationality and individual rationality. Subgame consistency requires that the extension of the solution policy to a later starting time and any possible state brought about by prior optimal behavior of the players will remain optimal.

Consider the cooperative game $\Gamma_c(x_0, T - t_0)$ in which the players agree to an optimality principle. In particular, given x_0 at time t_0, according to the solution optimality principle the players will adopt

(i) a weight α_1^0 leading to a set of cooperative controls $\Big\{\Big[\psi_1^{\alpha_1^0(t_0)}(t, x), \psi_2^{\alpha_1^0(t_0)}(t, x)\Big]$, for $t \in [t_0, T]\Big\}$, and

(ii) an imputation $\big[\xi^{(t_0)1}(x_0, T - t_0; \alpha_1^0), \xi^{(t_0)2}(x_0, T - t_0; \alpha_1^0)\big] = \big[\hat{W}^{t_0(\alpha_1^0)1}(t_0, x_0), \hat{W}^{t_0(\alpha_1^0)2}(t_0, x_0)\big]$.

Now consider the game $\Gamma_c(x_\tau, T - \tau)$ where $x_\tau \in X_\tau^{\alpha_1(t_0)}$ and $\tau \in [t_0, T]$, under the same solution optimality principle the players will adopt

(i) a weight α_1^τ leading to a set of cooperative controls $\Big\{\Big[\psi_1^{\alpha_1^\tau(\tau)}(t, x), \psi_2^{\alpha_1^\tau(\tau)}(t, x)\Big]$, for $t \in [\tau, T]\Big\}$, and

(ii) an imputation $[\xi^{(\tau)1}(\tau, T-\tau; \alpha_1^\tau), \xi^{(\tau)2}(\tau, T-\tau; \alpha_1^\tau)] = \big[\hat{W}^{\tau(\alpha_1^\tau)1}(\tau, x_\tau), \hat{W}^{\tau(\alpha_1^\tau)2}(\tau, x_\tau)\big]$.

A formal definition of subgame consistency (Yeung and Petrosyan [17]) can be stated as follows.

Definition 4.1. An optimality principle yielding imputations $\xi^{(\tau)}(x_\tau, T-\tau; \alpha_1^\tau)$, for $\tau \in [t_0, T]$ and $x_\tau \in X_\tau^{\alpha_1^0(t_0)}$, constitutes a subgame consistent solution to the game $\Gamma_c(x_0, T-t_0; \alpha_1^0)$ if the following conditions are satisfied:

(i) $\xi^{(\tau)}(x_\tau, T-\tau; \alpha_1^\tau) = [\xi^{(\tau)1}(x_\tau, T-\tau; \alpha_1^\tau), \xi^{(\tau)2}(x_\tau, T-\tau; \alpha_1^\tau)]$, for $t_0 \le \tau \le t \le T$, is Pareto optimal;

(ii) $\xi^{(\tau)i}(x_\tau, T-\tau; \alpha_1^\tau) \ge V^{(\tau)i}(\tau, x_\tau)$, for $i \in [1,2]$, $\tau \in [t_0, T]$, and $x_\tau \in X_\tau^{\alpha_1^0(t_0)}$; and

(iii) $\xi^{(\tau)i}(x_t, T-t; \alpha_1^\tau)\exp[r(\tau - t)] = \xi^{(t)i}(x_t, T-t; \alpha_1^t)$, for $i \in [1,2]$, $t_0 \le \tau \le t \le T$, and $x_t \in X_t^{\alpha_1^0(t_0)}$.

Part (i) of Definition 4.1 requires that according to the agreed-upon optimality principle Pareto optimality is maintained at every instant of time. Hence group rationality is satisfied throughout the game interval. Part (ii) demands individual rationality to be met throughout the entire game interval. Part (iii) guarantees the consistency of the solution imputations throughout the game interval in the sense that the extension of the solution policy to a situation with a later starting time and any possible state brought about by prior optimal behavior of the players remains optimal.

5 Subgame Consistent Solutions

In this section, we present subgame consistent solutions to the cooperative game $\Gamma_c(x_0, T-t_0)$. First note that group optimality will be maintained only if the solution optimality principle selects the same weight α_1 for all games $\Gamma_c(x_\tau, T-\tau)$, $\tau \in [t_0, T]$, and $x_\tau \in X_\tau^{\alpha_1(t_0)}$. For any chosen α_1 to maintain individual rationality throughout the game interval, the following condition must be satisfied:

$$\xi^{(\tau)i}(x_\tau, T-\tau; \alpha_1) = \hat{W}^{\tau(\alpha_1)i}(\tau, x_\tau) \ge V^{(\tau)i}(\tau, x_\tau),$$
$$\text{for } i \in [1,2],\ \ \tau \in [t_0, T],\ \text{ and } x_\tau \in X_\tau^{\alpha_1(t_0)}. \tag{20}$$

Invoking Propositions 2.1, 3.2, and 3.3, a sufficient condition for (20) to hold is

$$\hat{A}_i^{\alpha_1}(t) \ge A_i(t) \text{ and } \hat{B}_i^{\alpha_1}(t) \ge B_i(t), \ \text{ for } \ i \in [1,2], \quad t \in [t_0, T]. \tag{21}$$

We now introduce the following definitions.

Definition 5.1. We denote the set of α_1 that satisfies

$$\hat{A}_i^{\alpha_1}(t) \geq A_i(t) \text{ and } \hat{B}_i^{\alpha_1}(t) \geq B_i(t), \text{ for } i \in [1, 2],$$

at time $t \in [t_0, T)$ by S_t. We use $\underline{\alpha}_1^t$ to denote the lowest value of α_1 in S_t, and $\overline{\alpha}_1^t$ the highest value. In the case when t tends to T, we use $\overline{\alpha}_1^{T-}$ to stand for $\lim_{t\to T^-} \overline{\alpha}_1^t$, and $\underline{\alpha}_1^{T-}$ for $\lim_{t\to T^-} \underline{\alpha}_1^t$.

Definition 5.2. We define the set $S_\tau^T = \bigcap_{\tau\leq t\leq T} S_t$, for $\tau \in [t_0, T)$.

S_t represents the set of α_1 satisfying individual rationality at time $t \in [t_0, T)$ and S_τ^T represents the set of α_1 satisfying individual rationality throughout the interval $[\tau, T)$. In general $S_\tau^T \neq S_t^T$ for $\tau, t \in [t_0, T)$ where $\tau \neq t$.

5.1 Typical Configurations of S_t

To find typical configurations of the set S_t for $t \in [t_0, T)$ of the game $\Gamma_c(x_0, T - t_0)$, we perform extensive numerical simulations with a wide range of parameter specifications for $a, b, \sigma, h_1, h_2, k_1, k_2, c_1, c_2, q_1, q_2, T, r, x_0$. We calculate the time paths of $A_1(t)$, $B_1(t)$, $A_2(t)$, and $B_2(t)$ in Proposition 2.1 for $t \in [t_0, T]$. Then we select weights α_1 and calculate the time paths of $\hat{A}_1^{\alpha_1}(t)$, $\hat{A}_2^{\alpha_1}(t)$, $\hat{B}_1^{\alpha_1}(t)$, and $\hat{B}_2^{\alpha_1}(t)$ in Propositions 3.2 and 3.3, for $t \in [t_0, T]$. At each time instant $t \in [t_0, T]$, we derive the set of α_1 that yields $\hat{A}_i^{\alpha_1}(t) \geq A_i(t)$ and $\hat{B}_i^{\alpha_1}(t) \geq B_i(t)$, for $i \in [1, 2]$, to derive the set S_t, for $t \in [t_0, T)$.

We denote the locus of the values of $\underline{\alpha}_1^t$ along $t \in [t_0, T)$ as curve $\underline{\alpha}_1$ and the locus of the values of $\overline{\alpha}_1^t$ as curve $\overline{\alpha}_1$. In particular, three typical patterns prevail:

(i) The curves $\underline{\alpha}_1$ and $\overline{\alpha}_1$ are continuous and move in the same direction over the entire game duration: either both increase monotonically or both decrease monotonically (see Figure 1).
(ii) The curves $\underline{\alpha}_1$ and $\overline{\alpha}_1$ are continuous. $\underline{\alpha}_1$ declines and $\overline{\alpha}_1$ rises over the entire game duration (see Figure 2).
(iii) The curves $\underline{\alpha}_1$ and $\overline{\alpha}_1$ are continuous. One of these curves would rise/fall to a peak/trough and then fall/rise (see Figure 3).
(iv) The set $S_{t_0}^T$ can be nonempty or empty.

5.2 Subgame Consistent Cooperative Solutions

In this subsection, we present some subgame consistent solutions to $\Gamma_c(x_0, T - t_0)$.

Solution 5.1. If $S_{t_0}^T \neq \varnothing$ and $\underline{\alpha}_1^{T-} \in S_{t_0}^T$, an optimality principle under which the players agree to choose the weight

$$\alpha_1^* = \underline{\alpha}_1^{T-}, \text{ in } \Gamma_c(x_\tau, T - \tau) \text{ for } \tau \in [t_0, T)$$

yields a subgame consistent solution to the cooperative game $\Gamma_c(x_0, T - t_0)$.

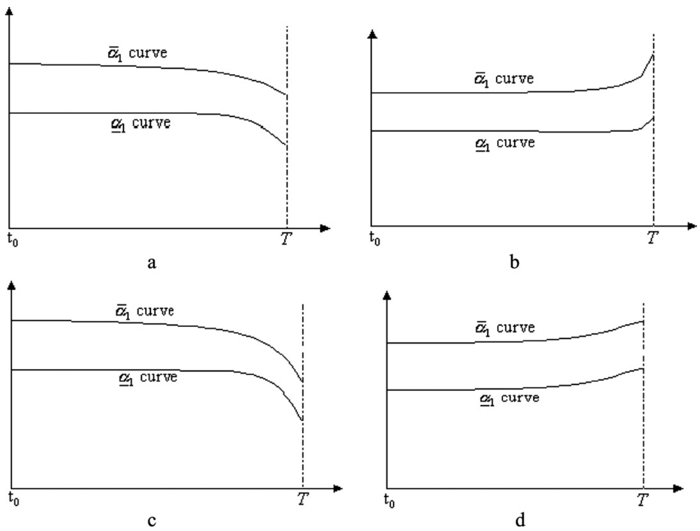


Figure 1: Typical Pattern (i).

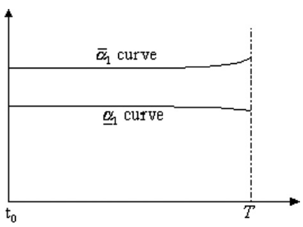


Figure 2: Typical Pattern (ii).

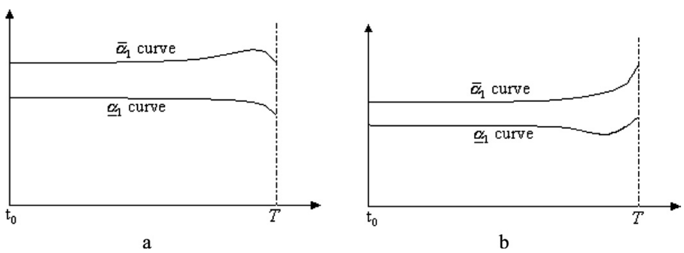


Figure 3: Typical Pattern (iii).

Proof. According to the optimality principle in Solution 5.1, a unique $\alpha_1^* = \underline{\alpha}_1^{T-}$ will be chosen for all the subgames $\Gamma_c(x_\tau, T-\tau)$, for $t_0 \le \tau \le t \le T$ and $x_\tau \in X_\tau^{\alpha_1^*(t_0)}$. The vector $\xi^{(\tau)}(x_\tau, T-\tau; \alpha_1^*) = \left[\hat{W}^{\tau(\alpha_1^*)1}(\tau, x_\tau), \hat{W}^{\tau(\alpha_1^*)2}(\tau, x_\tau)\right]$, for $\tau \in [t_0, T]$, yields a Pareto optimal pair of imputations. Hence part (i) of Definition 4.1 is proved.

From (18) and (19), one can readily verify that $\hat{W}^{\tau(\alpha_1^*)i}(t,x)\exp[r(\tau-t)] = \hat{W}^{t(\alpha_1^*)i}(t,x)$, for $i \in [1,2]$, $t_0 \le \tau \le t \le T$, and $x_t \in X_t^{\alpha_1^0(t_0)}$. Hence part (ii) of Definition 4.1 is satisfied.

Finally, from Definitions 5.1 and 5.2, one can verify that

$$\begin{aligned}\hat{W}^{\tau(\alpha_1^*)i}\,(\tau, x_\tau) &= \exp[-r(t-\tau)]\left[\hat{A}_i^{\alpha_1^*}(t)x^{1/2} + \hat{B}_i^{\alpha_1^*}(t)\right] \ge V^{(\tau)i}\,(\tau, x_\tau)\\ &= \exp[-r(t-\tau)]\left[A_i(t)x^{1/2} + B_i(t)\right],\\ &\quad \text{for } i \in [1,2], \tau \in [t_0, T], \text{ and } x_\tau \in X_\tau^{\alpha_1^*(t_0)}.\end{aligned}$$

Hence part (iii) of Definition 4.1 is fulfilled. □

Solution 5.2. If $S_{t_0}^T \neq \varnothing$ and $\overline{\alpha}_1^{T-} \in S_{t_0}^T$, an optimality principle under which the players agree to choose the weight

$$\alpha_1^* = \overline{\alpha}_1^{T-}, \text{ in } \Gamma_c(x_\tau, T-\tau) \text{ for } \tau \in [t_0, T)$$

yields a subgame consistent solution to the cooperative game $\Gamma_c(x_0, T-t_0)$.

Proof. Follow the proof of Solution 5.1. □

Solution 5.3. If $S_{t_0}^T \neq \varnothing$, $\underline{\alpha}_1^{T-} \notin S_{t_0}^T$, $\overline{\alpha}_1^{T-} \notin S_{t_0}^T$, $(\underline{\alpha}_1^{T-})^{0.5}(\overline{\alpha}_1^{T-})^{0.5} \in S_{t_0}^T$, an optimality principle under which the players agree to choose the weight

$$\alpha_1^* = (\underline{\alpha}_1^{T-})^{0.5}(\overline{\alpha}_1^{T-})^{0.5} \text{ in } \Gamma_c(x_\tau, T-\tau) \text{ for } \tau \in [t_0, T)$$

yields a subgame consistent solution to the cooperative game $\Gamma_c(x_0, T-t_0)$.

Proof. Follow the proof of Solution 5.1. □

Since $\underline{\alpha}_1^{T-}$ and $\overline{\alpha}_1^{T-}$ are tied to the noncooperative equilibrium, the Pareto optimal solutions 5.1–5.3 satisfy the axiom of symmetry in the sense that the switching of the labels of the players (that is, labeling player 1 as player 2, and vice versa) leaves the solution unchanged.

Solution 5.4. If $S_{t_0}^T \neq \varnothing$, $\underline{\alpha}_1^{T-} \notin S_{t_0}^T$, $\overline{\alpha}_1^{T-} \notin S_{t_0}^T$, $(\underline{\alpha}_1^{T-})^b(\overline{\alpha}_1^{T-})^{1-b} \in S_{t_0}^T$, an optimality principle under which the players agree to choose the weight

$$\alpha_1^* = (\underline{\alpha}_1^{T-})^b(\overline{\alpha}_1^{T-})^{1-b} \text{ in } \Gamma_c(x_\tau, T-\tau) \text{ for } \tau \in [t_0, T)$$

yields a subgame consistent solution to the cooperative game $\Gamma_c(x_0, T-t_0)$.

Proof. Follow the proof of Solution 5.1. □

6 Numerical Illustrations

Four numerical examples are provided below.

Example 6.1. Consider the cooperative game $\Gamma_c(x_0, T-t)$ with the following parameter specifications: $a = 10$, $b = 1$, $\sigma = 0.05$, $h_1 = 8$, $h_2 = 7$, $k_1 = 1$, $k_2 = 0.5$, $c_1 = 1$, $c_2 = 1.2$, $q_1 = 0.8$, $q_2 = 0.4$, $T = 6$, $r = 0.02$.

The numerical results are displayed in Figure 4. The curve $\underline{\alpha}_1$ is the locus of the values of $\underline{\alpha}_1^t$ along $t \in [t_0, T)$. The curve $\overline{\alpha}_1$ is the locus of the values of $\overline{\alpha}_1^t$ along $t \in [t_0, T)$. In particular, the set $S_{t_0}^T = \bigcap_{t_0 \le t < T} S_t = [\underline{\alpha}_1^{t_0}, \overline{\alpha}_1^{T-}] = [1.182686, 1.450783]$. Note that $\underline{\alpha}_1^{T-} \in S_{t_0}^T$ and $\overline{\alpha}_1^{T-} \notin S_{t_0}^T$, for $\tau \in [t_0, T)$. According to Solution 5.1, the players would agree to the optimality principle of choosing a weight $\alpha_1^* = \underline{\alpha}_1^{T-} = 1.182686$ throughout the game interval, and a subgame consistent solution to the cooperative game $\Gamma_c(x_0, T-t_0)$ would result.

Example 6.2. Consider the cooperative game $\Gamma_c(x_0, T-t)$ with the following parameter specifications: $a = 6$, $b = 0.8$, $\sigma = 0.04$, $h_1 = 8$, $h_2 = 6$, $k_1 = 1$, $k_2 = 0.5$, $c_1 = 1$, $c_2 = 1.5$, $q_1 = 3$, $q_2 = 2$, $T = 3$, $r = 0.02$.

The numerical results are displayed in Figure 5. In particular, the set $S_{t_0}^T = \bigcap_{t_0 \le t < T} S_t = [\underline{\alpha}_1^{t_0}, \overline{\alpha}_1^{T-}] = [1.246704, 1.443176]$. Note that $\overline{\alpha}_1^{T-} \in S_{t_0}^T$ and $\underline{\alpha}_1^{T-} \notin S_{t_0}^T$, for $\tau \in [t_0, T)$. According to Solution 5.2, the players would agree to the optimality principle of choosing a weight $\alpha_1^* = \overline{\alpha}_1^{T-} = 1.443176$ throughout the game interval, and a subgame consistent solution to the cooperative game $\Gamma_c(x_0, T-t_0)$ would result.

Example 6.3. Consider the cooperative game $\Gamma_c(x_0, T-t)$ with the following parameter specifications: $a = 10$, $b = 1.1$, $\sigma = 0.04$, $h_1 = 8$, $h_2 = 7$, $k_1 = 1$, $k_2 = 0.5$, $c_1 = 1$, $c_2 = 1.2$, $q_1 = 3$, $q_2 = 2$, $T = 3$, $r = 0.02$.

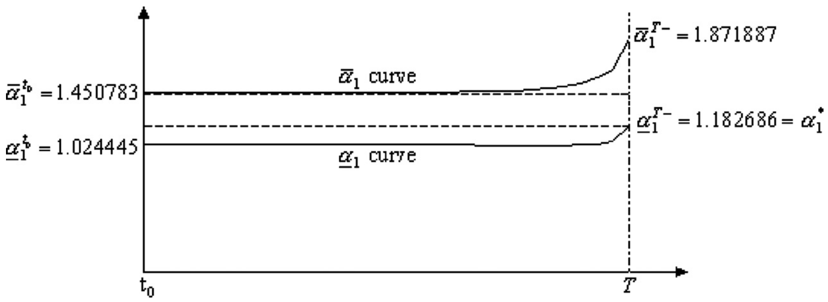


Figure 4: Solution to Example 6.1.

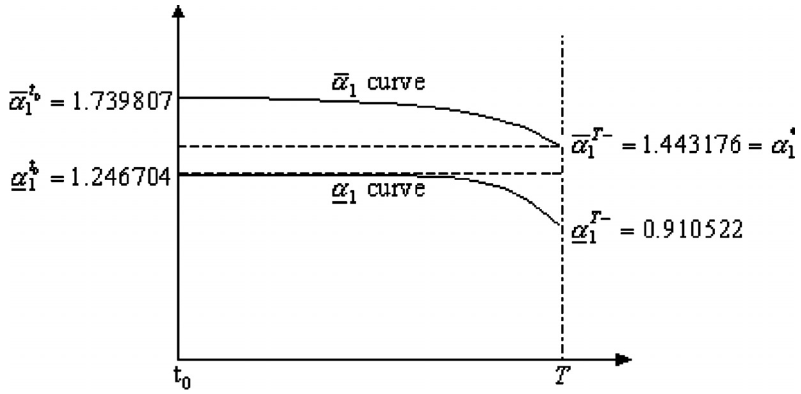


Figure 5: Solution to Example 6.2.

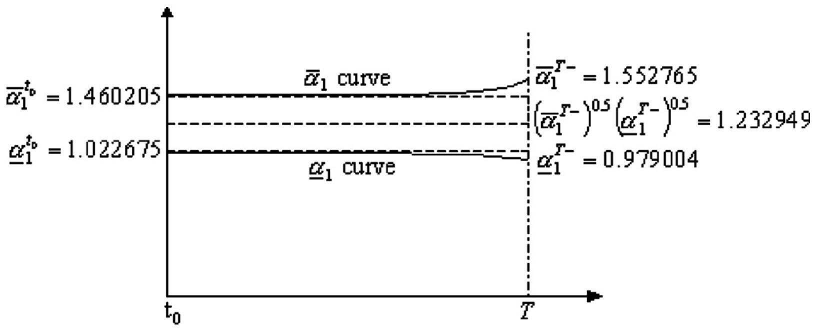


Figure 6: Solution to Example 6.3.

The numerical results are displayed in Figure 6. In particular, the set $S_{t_0}^T = \bigcap_{t_0 \le t < T} S_t = [\underline{\alpha}_1^{t_0}, \overline{\alpha}_1^{T-}] = [1.022675, 1.460205]$. Note that $\underline{\alpha}_1^{T-} \notin S_{t_0}^T$ and $\overline{\alpha}_1^{T-} \notin S_{t_0}^T$, for $\tau \in [t_0, T)$. According to Solution 5.3, the players would agree to the optimality principle of choosing a weight $\alpha_1^* = (\underline{\alpha}_1^{T-})^{0.5}(\overline{\alpha}_1^{T-})^{0.5} = 1.232949$ throughout the game interval, and a subgame consistent solution to the cooperative game $\Gamma_c(x_0, T - t_0)$ would result.

Example 6.4. Consider the cooperative game $\Gamma_c(x_0, T - t)$ with parameters: $a = 6$, $b = 1$, $\sigma = 0.03$, $h_1 = 11$, $h_2 = 6$, $k_1 = 1$, $k_2 = 0.5$, $c_1 = 1$, $c_2 = 1.5$, $q_1 = 3$, $q_2 = 2$, $T = 6$, $r = 0.02$.

The numerical results are displayed in Figure 7. In particular, the set $S_{t_0}^T = \bigcap_{t_0 \le t < T} S_t = \varnothing$. Hence there does not exist any candidate for a subgame consistent solution for the game $\Gamma_c(x_0, T - t_0)$.

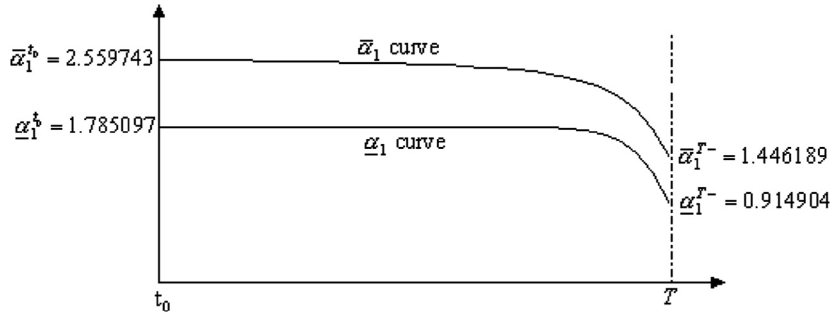


Figure 7: Empty set solution.

7 Concluding Remarks

In this chapter, subgame consistent solutions are derived for a class of cooperative stochastic differential games with nontransferable payoffs. The nontransferability of payoffs together with the stringent requirement of subgame consistency impose great hurdles to the derivation of solutions. The highly intractable subgame consistent solution for a class of such games is rendered tractable in this study. A number of subgame consistent solutions have been presented. Moreover, the analysis can be readily extended to an infinite horizon framework. In particular, one can obtain time invariant values of $\underline{\alpha}_1^t$ and $\overline{\alpha}_1^t$. An optimality principle under which the players agree to choose the weight α_1^* equaling the geometric mean of these two values yields a subgame consistent solution to the game. Since this work is among the first to present analytically tractable subgame consistent solutions of stochastic differential games with nontransferable payoffs, further research along this line is expected.

Acknowledgments

Financial support from the Research Grant Council of Hong Kong (Grants HKBU2103/04H) and Hong Kong Baptist University (Grant FRG/04-05/II-03) is gratefully acknowledged. The authors would like to thank an anonymous referee for invaluable comments and suggestions.

References

[1] Bellman, R., *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1957.

[2] Dockner, E.J., and Jørgensen, S., Cooperative and Non-Cooperative Differential Game Solutions to an Investment and Pricing Problem, *The Journal of the Operational Research Society*, 35, 8, 731–739, 1984.

[3] Fleming, W.H., Optimal Continuous-Parameter Stochastic Control, *SIAM Review*, 11, 470–509, 1969.

[4] Haurie, A., A Note on Nonzero-Sum Differential Games with Bargaining Solution, *Journal of Optimization Theory and Applications*, 18, 31–39, 1976.

[5] Haurie, A., Krawczyk, J. B., and Roche, M., Monitoring Cooperative Equilibria in a Stochastic Differential Game, *Journal of Optimization Theory and Applications*, 81, 79–95, 1994.

[6] Isaacs, R., *Differential Games*, Wiley, New York, NY, 1965.

[7] Leitmann, G., *Cooperative and Non-Cooperative Many Players Differential Games*, Springer-Verlag, New York, NY, 1974.

[8] Nash, J.F., Two-Person Cooperative Games, *Econometrica*, 21, 128–140, 1953.

[9] Petrosyan, L.A., Stable Solutions of Differential Games with Many Participants, *Viestnik of Leningrad University*, 19, 46–52, 1977.

[10] Petrosyan, L.A., Agreeable Solutions in Differential Games, *International Journal of Mathematics, Game Theory and Algebra*, 7, 165–177, 1997.

[11] Petrosyan, L.A., Bargaining in Dynamic Games, *ICM Millennium Lectures on Games*, Edited by L. Petrosyan and D. Yeung, Springer-Verlag, Berlin, Germany, pp. 139–143, 2003.

[12] Petrosyan, L., and Danilov, N., Stability of Solutions in Non-zero Sum Differential Games with Transferable Payoffs, *Viestnik of Leningrad Universtiy*, 1, 52–59, 1979.

[13] Shapley, L.S., *A Value for N-Person Games, Contributions to the Theory of Games*, Edited by H. W. Kuhn and A. W. Tucker, Princeton University Press, Princeton, NJ, vol. 2, pp. 307–317, 1953.

[14] Yeung, D.W.K., Nontransferable Individual Payoff Functions Under Stochastic Dynamic Cooperation, *International Game Theory Review*, 6, 2, 281–289, 2004.

[15] Yeung, D.W.K., and Petrosyan, L.A., Proportional Time-Consistent Solutions in Differential Games, *International Conference on Logic, Game Theory and Social Choice*, St. Petersburg, Russia, pp. 254–256, 2001.

[16] Yeung, D.W.K., and Petrosyan, L.A., Subgame Consistent Cooperative Solutions in Stochastic Differential Games, *Journal of Optimization Theory and Applications*, 120, 3, 651–666, 2004.

[17] Yeung, D.W.K., and Petrosyan, L.A., Subgame Consistent Solution of a Cooperative Stochastic Differential Game with Nontransferable Payoffs, forthcoming in *Journal of Optimization Theory and Applications*, 124, 3, 701–724, 2005.

PART III
Pursuit-Evasion Games

Geometry of Pursuit-Evasion Games on Two-Dimensional Manifolds

Arik Melikyan
Institute for Problems in Mechanics
Russian Academy of Sciences
101-1 Vernadskogo pr.
Moscow 117526, Russia
melik@ipmnet.ru

Abstract

Pursuit-evasion games with simple motion on two-dimensional (2D) manifolds are considered. The analysis embraces the game spaces such as 2D surfaces of revolution (cones and hyperboloids of one and two sheets, ellipsoids), a Euclidean plane with convex bounded obstacle, two-sided Euclidean plane with hole(s), and two-sided plane bounded figures (disc, ellipse, polygon). In a two-sided game space players can change the side at the boundary (through the hole).

In all cases the game space is a 2D surface or a figure in 3D Euclidean space, while the arc length is induced by the Euclidean metric of the 3D space. Due to simple motion, optimal trajectories of the players generally consist of geodesic lines of the game space manifolds. For the game spaces under consideration there may exist two or more geodesic lines with equal lengths, connecting the players. In some cases this gives rise to a singular surface consisting of the trajectories, which are envelopes of a family of geodesics.

In this chapter we investigate the necessary and sufficient conditions for this and some other types of singularity, we specify the game spaces where the optimal pursuit-evasion strategies do not contain singularities and are similar to the case of a Euclidean plane, and we give a short review and analysis of the solutions for the games in several game spaces-manifolds.

In the analysis we use viscosity solutions to the Hamilton–Jacobi–Bellman–Isaacs equation, variation calculus and geometrical methods. We also construct the 3D manifolds in the game phase space representing the positions with two or more geodesic lines with equal lengths, connecting the players.

The investigation of 2D games on the manifolds has several direct applications, and it may also represent an approximate solution for more complicated games as an abstraction.

1 Introduction

Simple motion very often is used in the theory of differential games for modelling of different moving objects [1]–[4]. Differential games with simple motion are very important for modelling in many practical problems, and they are convenient for the first applications of new theoretical methods due to their simplicity. Technically the solution of a game problem with simple motion generally is simpler to achieve than in the problems with complicated dynamics [5], [6]. For instance, the games of convergence and pursuit-evasion in Euclidean space have relatively simple solutions.

Many game models either have complex nonlinear dynamic equations or they have a simple motion dynamic together with a game space of a complex geometry, expressed as Riemannian manifolds possibly having edges or singular points. The nonlinearity properties are represented to some extent by the manifolds' metric tensor, while the character of the dynamic equations allows the effective use of geometric methods. This chapter considers pursuit-evasion differential games with simple motion on two-dimensional (2D) manifolds, the cost function being the (zero-radius) capture time.

One can compute the optimal pursuit-evasion time in the games with simple motion on the Euclidean plane (space) by dividing the initial distance between the players (the length of the shortest geodesic line) by the difference of the speeds. This formula is also true for the games on the surfaces (manifolds) for a sufficiently small initial distance. The optimal trajectories of the players are determined by the shortest geodesic line connecting them. Both players have to move along this common geodesic line. Such a pursuit (evasion) strategy will be called the primary pursuit (evasion) strategy.

The situation changes when the game space is a manifold rather than a Euclidean space, and there exist two or more geodesic lines with the same length, connecting the players. The existence of two equal geodesics for some positions of the players gives rise to what is called the secondary domain where the optimal motion of the players is still a motion along a geodesic, while each player exploits his own geodesic line, which is different from the one connecting them. The singular (equivocal) surface separating the primary and secondary domains represents the optimal trajectories, which are envelopes of a family of geodesics; such surfaces are also called switching envelopes.

As shown in [7], [8], in games on 2D cones in addition to these two types of domains there may exist a domain D of a third type, where the game value depends only on the coordinates of the pursuer. Consequently, that player has a definite optimal behavior, while the evader may exploit an arbitrary control until he reaches the boundary of D. In other words, in D one has an optimal control problem rather than a dynamic game. Such a phenomenon takes place in pursuit-evasion games on a family of 2D surfaces of revolution, including cones and hyperboloids of two sheets. The complete solution of the games on cones are found in [7], [8]; the games on hyperboloids are investigated in [9], [10]. However, if the game space is, for instance, a convex cylinder, the existence of two geodesics does not change the optimality of the primary strategy.

The existence of such a singular surface essentially depends on the compactness and symmetry of the game space. By investigating the geometric properties of a game space we describe a sufficient condition for the primary strategy to be optimal in the whole space.

Based on the solution of the above-mentioned games one can state that all regular paths are of three types: (1) players move along the geodesic line connecting them (primary Euclidean-plane-like strategy); (2) each player exploits his own geodesic line, which is different from the line connecting them (secondary strategy); (3) the pursuer moves along a geodesic, the evader has arbitrary motion (optimal-control-like situation).

2 Game Space and Dynamics

Let two points-players P (pursuer) and E (evader) perform simple motion on a manifold M equipped with nondegenerate Riemannian metrics. In this chapter as a game space M actually considered a 2D surface in 3D Euclidean space will be considered with the metrics on M induced by 3D Euclidean metrics, and subdomains of a Euclidean plane, possibly two-sided.

Let $x, y \in R^2$ be local coordinates describing the positions of the players. Then the dynamic equations for the players, performing simple motions, can be written as

$$\dot{x} = u, \quad \dot{y} = v, \quad x, y, u, v \in R^2 \tag{1}$$

subject to the constraints

$$\sqrt{\langle G(x)u, u\rangle} \leq 1, \quad \sqrt{\langle G(y)v, v\rangle} \leq \nu$$

Here $G(x)$ is a 2×2-matrix, metric tensor of the manifold (surface) specifying a Riemannian metric. This constraints presume that the maximal speed of the pursuer is 1 and that of the evader is $\nu \geq 0$. The constant ν may also be regarded as the speed ratio $\nu = v_E/v_P$. We assume that $\nu \leq 1$.

A pursuit-evasion game is considered for the system (1) with zero capture radius, i.e., the game starting at time $t = 0$ ends for some $t = T$ if the capture condition takes place, $x(T) = y(T)$, and the cost function is T.

For the plane space M the metric tensor is the identity matrix, $G = I$, and the dynamic equations and restrictions take the form

$$\begin{gathered}\dot{x} = u, \quad \dot{y} = v, \qquad x, y, u, v \in R^2 \\ u_1^2 + u_2^2 \leq 1, \quad v_1^2 + v_2^2 \leq \nu^2.\end{gathered}$$

Here $x = (x_1, x_2)$, $y = (y_1, y_2)$ are Cartesian (local) coordinates of the players P and E. For a two-sided game space one also has to specify on which side of the plane, e.g., "front" or "back" side, the player is situated. In addition, a restriction must be imposed on the edges in order to prevent a player from moving outside the space M.

3 Primary Solution

Generally, in a main part $D_1 \subset M \times M$ of the phase space $M \times M$ of the game, called the primary domain, the game value $V(x, y)$ equals what is called the primary time, well-known for the Euclidean space:

$$V_1(x, y) = \frac{L(x, y)}{1 - \nu} \, (x, y) \in D_1, \tag{2}$$

where $L(x, y)$ is the minimal Riemannian distance between P and E defined as a minimum of the geodesic functional:

$$L(x, y) = \min_{\xi(\sigma)} \int_{\sigma_0}^{\sigma_1} \sqrt{\langle G(\xi)\dot{\xi}, \dot{\xi}\rangle} d\sigma, \qquad \xi(\sigma_0) = y, \quad \xi(\sigma_1) = x. \tag{3}$$

Here $\xi(\sigma)$ is a piecewise smooth curve $\xi(\sigma) \in M$, $\sigma_0 \le \sigma \le \sigma_1$.

In the primary domain D_1 the optimal motion of the players is to move along the geodesic line connecting the players. Corresponding optimal strategies have the form [11]

$$\begin{aligned} u(x, y) &= -a(x, y) \quad (a(x, y) = G^{-1}(x)\partial L/\partial x) \\ v(x, y) &= \nu b(x, y) \quad (b(x, y) = G^{-1}(y)\partial L/\partial y), \end{aligned} \tag{4}$$

where a and b are outward unit tangent vectors to this geodesic line at the points P and E, i.e.,

$$|a|_x^2 = \langle G(x)a, a\rangle = 1, \quad |b|_y^2 = \langle G(y)b, b\rangle = 1.$$

The minimum in (3) generally has the form

$$L(x, y) = \min_{\alpha \in A} S(x, y, \alpha) = S(x, y, \alpha), \ \alpha \in A^* \subset A, \tag{5}$$

where $S(x, y, \alpha)$ is the family of local minima of the geodesic functional, and α is a family parameter from some set A. In the considered problems this set usually consists of two elements, $A = \{+, -\}$, and formula (5) takes the form (Figure 1)

$$L(x, y) = \min[L^+(x, y), L^-(x, y)]. \tag{6}$$

One can show that the primary time (2) is guaranteed for player P in the whole space $M \times M$, i.e., the capture takes place not later than at the time (2). Indeed, consider the total time derivative of $S(x, y, \alpha)$ for fixed α:

$$\dot{S} = \langle a, u\rangle + \langle b, v\rangle \quad (a = S_x, b = S_y).$$

The maximal decrease of the distance between the players that player P can achieve is [12]

$$\min_u \dot{L} = \min_u \min_{\alpha \in A^*} \dot{S} = -1 + \min_{\alpha \in A^*} \langle b, v\rangle \le -1 + \nu. \tag{7}$$

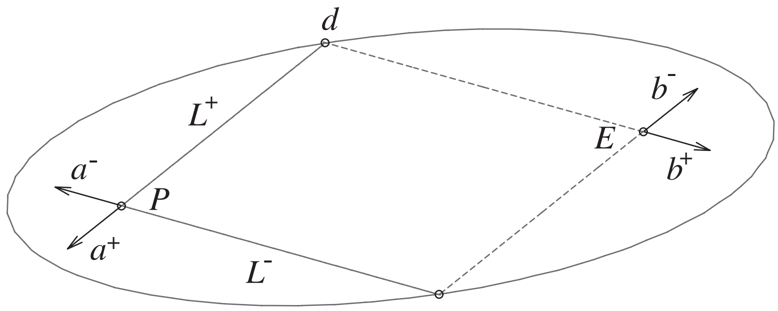


Figure 1.

The set A^* is defined in (5), and for the inequality here, which proves the statement, the restriction for the vector v is used.

A similar statement for the guarantee of the time (2) for player E is not true due to the following maxmin property:

$$\max_v \dot{L} = \max_v \min_{\alpha \in A^*} \dot{S} \leq \min_{\alpha \in A^*} \max_v \dot{S}.$$

In light of the above considerations it is convenient to represent the phase space $M \times M$ as a sum:

$$M \times M = D_1 + D_2,$$

while the optimal capture time in D_1 is equal by definition to (2), and in the second domain D_2 the capture time is strictly less than (2).

4 Nonunique Geodesics, Manifolds Γ and B

For the manifolds under consideration the nonuniqueness of the geodesic line with minimal length is generally violated by the existence of two geodesics with the length denoted L^+ and L^-, see (6). There may be positions with several or infinitely many geodesics. For example, for a two-sided plane ellipse there are four positions, when players are on different sides and in different focal points, with an infinite number of geodesics. For a rectangle one has positions with two or three geodesics (and four in the case of a square).

We denote the manifold with nonunique geodesic line as Γ, $\Gamma \subset M \times M$. An analysis shows that Γ is a 3D submanifold with an edge. The internal points of Γ correspond to the positions with two geodesic lines. In the vicinity of such points there are two local minima of geodesic length denoted as L^+, L^-. Thus for the internal points of Γ one has

$$\Gamma : \quad L^+(x, y) = L^-(x, y). \tag{8}$$

The boundary points of Γ may have one, two or more geodesics with equal length. The construction of such boundaries for some concrete game spaces are given in Section 5.

For many game spaces a part $\Gamma_1 \subset \Gamma$ of the manifold Γ is a dispersal surface [1], i.e., at each point of Γ_1 two optimal (primary) trajectories start with the same capture time (2). Primary pursuit is performed along one of the two geodesics, while the choice of the actual line is upon player P.

Depending on the geometrical characteristics of the space M the set Γ_1 may coincide with the whole Γ, which means that the primary solution is the game solution in the whole space. If Γ_1 and Γ do not coincide, then the existence of two geodesics allows player P to maneuver so that in some (secondary) subspace $D_2 \subset M \times M$ the optimal capture time is strongly less than (2).

Let us derive a necessary condition for the global optimality of the primary solution. Such a condition gives a relation between the geometric characteristics of the space M and the parameter ν.

Relation (8) means that the minimum (6) is attained at both elements and the sets A and A^* in (5) coincide. Then the derivative of the value (6) along the solutions of (1) at the points of the set Γ can be written as

$$\dot{L} = \min[\langle a^+, u\rangle + \langle b^+, v\rangle, \langle a^-, u\rangle + \langle b^-, v\rangle]. \tag{9}$$

Here $a^\pm$ and $b^\pm$ are the unit outward tangent vectors to the geodesics L^+, L^- at the points P and E (Figure 1). Computations are given for the locally Euclidean metric, $G = I$; the general formulas are similar.

The derivation of formula (7) shows that player P can guarantee the time (2) by using the strategy $u = -a^+$ or $u = -a^-$, see (4). Prescribe now to player P for a position from the set Γ the control value

$$u = -(a^+ + a^-)/|a^+ + a^-|, \tag{10}$$

which means that the player does not move along one of the geodesics but chooses a direction between them. Using (10) in (9) one can estimate

$$\dot{L} \le -|a^+ + a^-|/2 + \nu|b^+ + b^-|/2. \tag{11}$$

Here we used the equality

$$|a^+ + a^-|^2 = \langle a^+ + a^-, a^+ + a^-\rangle = 2(1 + \langle a^+, a^-\rangle)$$

and the fact that the maximum

$$\max_v \min[\langle b^+, v\rangle, \langle b^-, v\rangle]$$

is attained at the vector

$$v = \nu(b^+ + b^-)/|b^+ + b^-|.$$

Comparing (7) with (11) one can conclude that the control (10) is preferable for player P if in the considered point of the manifold Γ the following inequality takes place:

$$|a^+ + a^-| - \nu|b^+ + b^-| > 2(1 - \nu). \tag{12}$$

Though the relations (9)–(12) carry instantaneous character (being related to the initial time instant), they lead to the existence of a sufficiently small time interval during which player P gets a faster than (7) decrease of the distance between the players, thus improving the time (2).

The surface Γ_1 lies in the domain D_1, where the time (2) is optimal, and the condition (12) cannot be true there. Thus, for the points of Γ_1 the following relations are true:

$$\Gamma_1 : \quad L^+ = L^-, \quad |a^+ + a^-| - \nu|b^+ + b^-| \le 2(1 - \nu). \tag{13}$$

For the global optimality of the primary solution the set (12) must be empty. In other words, the edge B of the manifold Γ_1

$$B : \quad L^+ = L^-, \quad |a^+ + a^-| - \nu|b^+ + b^-| = 2(1 - \nu) \tag{14}$$

must not have points in the interior of Γ.

5 Examples of Manifolds Γ and B

We give a review of the manifolds Γ, B and their edges $\partial\Gamma$, ∂B constructed for some game spaces.

Surfaces of revolution. Consider the surface of revolution M given in 3D Euclidean space with the coordinates (x_1, x_2, x_3) by the equation

$$x_3 = f\left(\sqrt{x_1^2 + x_2^2}\right), \tag{15}$$

where $f(z)$ is a twice differentiable function for $0 \le z < \infty$, such that

$$f'(0) = 0, \quad f'(z) > 0, \ \ 0 < z < \infty. \tag{16}$$

Under these conditions the surface M has twice differentiable and monotone generatrix and is unbounded. As a particular case of the surface (15) we will consider

$$f(z) = \sqrt{(z^2 + m^2)}/k,$$

which gives the upper nappe of the hyperboloid (the cone) of revolution

$$x_1^2 + x_2^2 + m^2 = k^2 x_3^2, \quad m, k > 0, \tag{17}$$

where m, k are parameters, positive real numbers. For $m = 0$ equation (11) defines a cone of revolution. Taking the term $-m^2$ rather than m^2 one gets a hyperboloid of one sheet. The solution of a pursuit-evasion game on it is given in [9].

Consider as a game space the surface M. One can choose the first two components $x = (x_1, x_2)$, $y = (y_1, y_2)$ as local coordinates, describing the positions of the players. Then the dynamic equations for the players, performing simple motions, can be written in the form (1), where the entries $g_{ij}(x)$ of the metric tensor $G(x)$ can be computed using the quadratic form

$$\langle G(x)dx, dx\rangle = dx_1^2 + dx_2^2 + dx_3^2$$

and excluding dx_3 using (15) and $z = \sqrt{x_1^2 + x_2^2}$:

$$g_{11} = 1 + \frac{f'^2(z)x_1^2}{x_1^2 + x_2^2}, \quad g_{22} = 1 + \frac{f'^2(z)x_2^2}{x_1^2 + x_2^2}, \quad g_{12} = g_{21} = \frac{f'^2(z)x_1x_2}{x_1^2 + x_2^2}. \quad (18)$$

Due to the rotational symmetry of M the manifold Γ also has a symmetry. A point of Γ corresponds to a position of the players situated in a common plane through the x_3-axis. As a local coordinate on Γ one can choose the angle β between that plane and, say, coordinate plane $x_2 = 0$, together with the distances R and r from players P and E to the apex $(0, 0, f(0))$.

We will construct a generatrix section of Γ. The complete manifolds Γ, $\partial\Gamma$, B, ∂B can be obtained by rotation of the curves on Figure 2 about one of the coordinate axes.

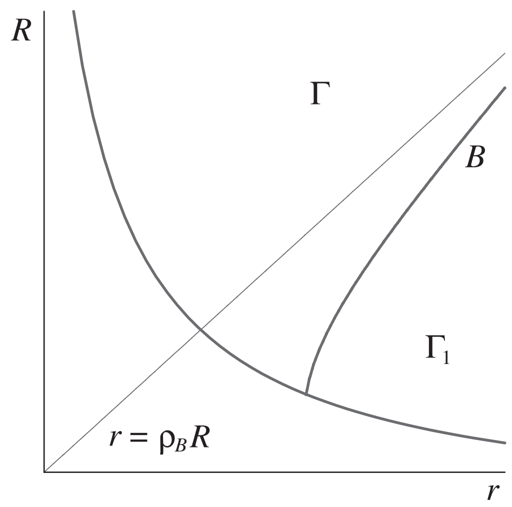


Figure 2.

Numerical constructions show [10] that the points of the manifold Γ on the plane of R, r (for fixed β) lie above the curve $\partial\Gamma$ having asymptotes parallel to coordinate axes, Figure 2. To the points below that curve correspond positions with one geodesic line. One can mention two possible mechanisms of transition from two geodesics to a single one. First, when a point of Γ approaches $\partial\Gamma$ from above, two geodesic lines coincide, so that to the points of $\partial\Gamma$ corresponds only one geodesic. Second, three (or more) geodesics of equal length may arise at $\partial\Gamma$. Such a phenomenon takes place when the game space is a double-sided rectangle, or a surface of revolution with the generatrix of an appropriate shape. For many surfaces, like hyperboloids or paraboloids, numerically investigated in [10], the first mechanism takes place. In this chapter we will consider such surfaces. However, global (numerical) investigations are required to determine which mechanism is actually in order.

Consider first the boundary $\partial\Gamma$. One can suppose (and the considerations in the sequel confirm this) that the two geodesics, L^+ and L^-, coincide at the positions from $\partial\Gamma$. This means that the points P and E are, by one of the definitions, mutually conjugate points with respect to the geodesic functional (3). This phenomenon, numerically investigated in [10], gives an algorithm for direct computation. Thus, the determination of $\partial\Gamma$ means that one has to find all mutually conjugate points P and E. Due to rotational symmetry and the equality $L^+ = L^-$ in (8) these points can be regarded as points of a generatrix of the surface M that lies, say, in the coordinate plane $x_2 = 0$, and according to (15), is defined as

$$x_3 = f(x_1),\ x_1 \geq 0, \quad x_3 = f(-x_1),\ x_1 \leq 0. \tag{19}$$

The two coinciding geodesics PE represent a curve segment of the curve (19), which by virtue of (16) is a twice differentiable curve.

The projection of the curve PE onto the plane $x_3 = const$ is an extremal of the functional (3). The Lagrangian (integrand) in (3) is a homogeneous function in $\dot{\xi}$, which causes a certain degeneracy. The latter could be removed if one introduces x_1 as an independent variable and transforms the functional (3) in terms of one scalar function $x_2 = x_2(x_1)$, or using the classical notation $x_2 = y$, $x_1 = x$, $y = y(x)$, to the following form:

$$L = \int_{x_0}^{x_1} F(x, y, y')dx, \quad F = \sqrt{g_{11}(x, y) + 2g_{12}(x, y)y' + g_{22}(x, y)y'^2}, \tag{20}$$

where g_{ij} are the entries of the symmetric (2×2)-matrix $G(x, y)$.

One can verify that the curve under the investigation, $y = y(x) \equiv 0$, $x_0 \leq x \leq x_1$, is an extremal for (20), i.e., it satisfies the Euler equation

$$F_y - \frac{d}{dx}F_{y'} = 0.$$

The focal point $P(x_1, 0)$ conjugate to $E(x_0, 0)$ with respect to the functional (16) is defined using the solution $h(x)$ of the following initial value problem [17]:

$$\frac{d}{dx}(Ph') + (Q' - R)h = 0, \quad h(x_0) = 0,\ h'(x_0) = 1 \tag{21}$$

$$P(x) = F_{y'y'} = \frac{1}{\sqrt{g_{11}}}, \quad Q(x) = F_{y'y} = \frac{f'^2(x)}{x\sqrt{g_{11}}}$$

$$R(x) = F_{yy} = \frac{f'(x)f''(x)}{x\sqrt{g_{11}}} - \frac{f'^2(x)}{x^2\sqrt{g_{11}}}, \quad g_{11}(x, 0) = 1 + f'^2(x).$$

The value x_1 is the next root of $h(x)$:

$$h(x_1) = 0, \qquad h(x) > 0, \quad x_0 < x < x_1.$$

The above linear ordinary differential equation is known as the Jacobi equation for the functional (20); it is an equation in variations for the Euler equation. For the construction of $\partial\Gamma$ it is sufficient to find corresponding x_1 for all $x_0 > 0$.

One can show that the general solution of the Jacobi equation (21) has the form

$$h(x) = C_1 x + C_2 x\varphi(x), \quad \left(\varphi(x) = \int \sqrt{g_{11}(x, 0)}\, \frac{dx}{x^2}\right).$$

Initial conditions (21) lead to the following values of constants:

$$C_1 = -\frac{x_0\varphi(x_0)}{\sqrt{g_{11}(x, 0)}}, \quad C_2 = \frac{x_0}{\sqrt{g_{11}(x, 0)}}.$$

The requirement $h(x_1) = 0$ leads to the equality

$$\partial\Gamma : \quad \varphi(x_0) - \varphi(x_1) = 0, \tag{22}$$

which specifies the points of $\partial\Gamma$.

The edge ∂B of the manifold B is a submanifold of $\partial\Gamma$ for which the last equality in (14) is fulfilled. Since the tangent vectors a^+ and a^-, b^+ and b^- coincide in (14) and the equality is fulfilled trivially, one needs a limit form of it as a point under consideration tends to $\partial\Gamma$ from the interior of Γ. Using in (14) second-order expansions in terms of small $|a^+ - a^-|$ and $|b^+ - b^-|$, as in [18], one can obtain the following condition for the point ∂B:

$$h'(x_1) = -\sqrt{\nu}\frac{P(x_0)}{P(x_1)} = -\sqrt{\nu}\sqrt{\frac{g_{11}(x_1, 0)}{g_{11}(x_0, 0)}}. \tag{23}$$

Here the connection (21) between $P(x)$ and $g_{11}(x, 0)$ is used. Using the general solution formula one can simplify this equation to the form $x_0 = -\sqrt{\nu}\, x_1$. Thus, one gets for the edge ∂B the following two equations:

$$\partial B : \quad \varphi(x_0) - \varphi(x_1) = 0, \quad x_0 = -\sqrt{\nu}\, x_1. \tag{24}$$

Excluding the variable x_0 one gets a single equation

$$\varphi(x_1) + \varphi(\sqrt{\nu}x_1) = 0$$

with respect to unknown x_1. One can apply Newton's method to solve this equation. The solution of the system (24), a point, defines the manifold ∂B up to revolution about the vertical x_3-axis.

For a paraboloid of revolution, when $f(z) = kz^2/2$ in (15), one has

$$\varphi(x) = \sqrt{1 + k^2x^2}/x - k\ln(kx + \sqrt{1 + k^2x^2}).$$

For the hyperboloid with $f(z) = \sqrt{(z^2 + b)/k}$ in (15) the antiderivative $\varphi(x)$ is expressed in elliptic functions.

As one can see from (24) the numbers x_0 and x_1 have different signs, i.e., the corresponding points E and P are on different sides with respect to the apex A. Thus, in relative coordinates (see Section 7) (R, r, φ) the coordinates of the point ∂B are

$$R_1 = \int_0^{|x_1|} \sqrt{g_{11}(x, 0)}dx, \quad r_1 = \int_0^{|x_0|} \sqrt{g_{11}(x, 0)}dx, \quad \varphi = \pi, \tag{25}$$

where R_1, r_1 are the lengths of a subarcs of the curve (19).

Cone, hyperboloid of revolution. Consider the particular case (17) when M is the (upper) nappe of the hyperboloid of revolution. Numerical constructions show [10] that the curve $\partial\Gamma$ asymptotically approaches coordinate axes (see Figure 2). The manifold B defined in (14) has an edge ∂B lying on the curve $\partial\Gamma$, while the curve B has an asymptote

$$r = \rho_B R, \quad \rho_B = \rho_B(\nu, \alpha)$$

which is the manifold B itself for the cone (Figure 2). The dependence of the coefficient ρ_B on the speed ratio ν and the half-angle of the cone's plane development α can be found explicitly [13].

The cone is approached as $m \to 0$. The curve $\partial\Gamma$ then tends to a right angle formed by positive semiaxes, the curve B tends to its asymptote, and the point ∂B tends to the origin. The boundary $\partial\Gamma$ for the cone is a plane with half-line perpendicular to it. The manifold B again happens to be a cone, so that it is diffeomorphic to the game space M [13].

Two-sided plane angle. Such a game space is equivalent to the case of a cone. There are two deformations of the cone into a plane figure: to fold it along two opposite generatrices into a plane two-sided angle, or to cut along one generatrix and develop in a plane. The value of angle α equals the half-angle of the plane development and $0 < \alpha < \pi$. Both deformations preserve the geodesic lengths and there is one-to-one correspondence of optimal trajectories.

It is convenient to distinguish the acute cones, whose parameters (ν, α) satisfy the relations

$$0 \le \nu \le 1 - \sin\alpha, \quad 0 \le \alpha \le \pi/2 \tag{26}$$

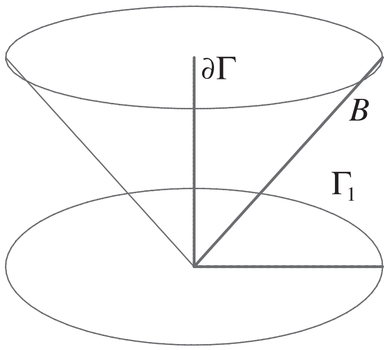


Figure 3.

and the remaining (blunt) cones. The acute cones have nonempty sets D_3 and D_2^0; if the game starts at a point of D_3 the optimal capture takes place at the apex. For the blunt cones the set D_3 is empty, while D_2 is nonempty. The manifolds Γ and B have the structure given in Figure 3.

Two-sided plane with a circular hole. In this game the player can move on both sides of a Euclidean plane and can change the side at the points of the circle [14]. The solution of the problem is similar to the pursuit-evasion problem on a (one-sided) plane with a circular obstacle [15], [16].

Let R and r be distances from the players P and E correspondingly from the center O of the circle with unit radius. The angle between segments OP, OE we denote as φ. The manifold Γ is given by the equality case $\varphi = \pi$, and its dispersal subset is characterized by the relations

$$\varphi = \pi, \quad \sqrt{1 - R^{-2}} - \nu\sqrt{1 - r^{-2}} \le 1 - \nu.$$

The equality case in the right relation corresponds to the set B. The curve B is shown in Figure 4; its asymptote is the line $r = \sqrt{\nu}\, R$.

The right angle with the apex Q together with the curve B gives the generatrix whose revolution about the ordinate axis forms the manifolds Γ and B.

Constructions show that the set B is nonempty for all considered velocity values.

Polygon. When the game space is a polygon, say a triangle, rectangle, etc., one can show that the set B is nonempty, and thus the primary strategy is not globally optimal. Indeed, in a sufficiently small neighborhood of an apex the optimal paths are the same as in the case of the plane angle (cone), when singular equivocal paths always exist. Note that in case of the acute angle at an apex (see condition (26)) the player P can force the capture of E at that apex.

Plane two-sided strip, cylinder. The set B is empty for such a game space since, due to possible permutations of the points P and E, one has for the tangent vectors

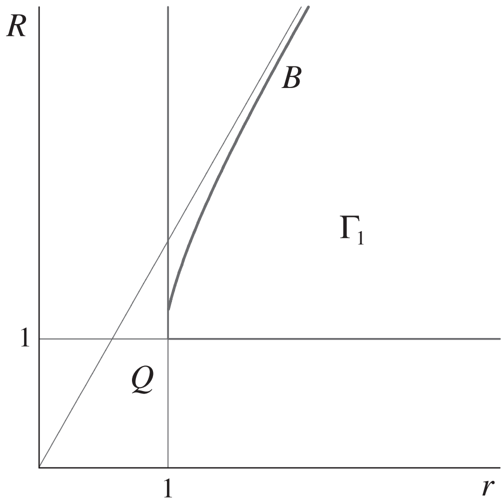


Figure 4.

in the relations (12), (14)

$$|a^+ + a^-| = |b^+ + b^-| < 2.$$

For such vectors the equality (14) is impossible for all values of ν.

Ellipse. For an ellipse the set B may be either empty or not depending upon the parameters ν, ε. Here ε is the eccentricity of the ellipse:

$$\varepsilon = c/a, \quad 0 < \varepsilon < 1, \quad c^2 = a^2 - b^2,$$

where a, b are major and minor semiaxes.

In Figure 5 a curve is shown in the plane (ν, ε), which separates the sets A_1, A_2 corresponding to "long" and "round" ellipses. For an ellipse with the parameters from the region A_1 the set B is empty and the primary strategy is globally optimal. The same is true for the points of the curve itself, for which B is nonempty but the inequality (12) is not fulfilled.

The points (ν, ε) of the set A_1 satisfy the condition

$$A_1: \quad \max_{(P,E)}(|a^+ + a^-| - \nu|b^+ + b^-|) \leq 2(1-\nu), \quad (P, E) \in \Gamma.$$

The boundary of A_1, for which the equality takes place in the above relation, was found numerically.

Figure 6 shows the ellipses with the parameters from the separating curve, i.e., with, maximal velocity ν for a given ε for which the purely primary strategy is still effective.

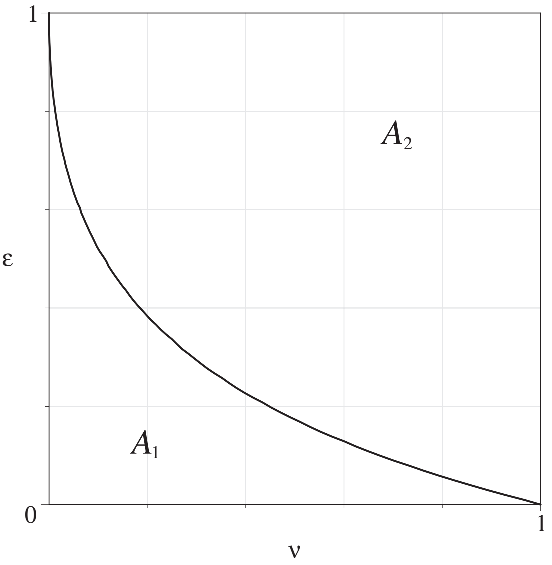


Figure 5.

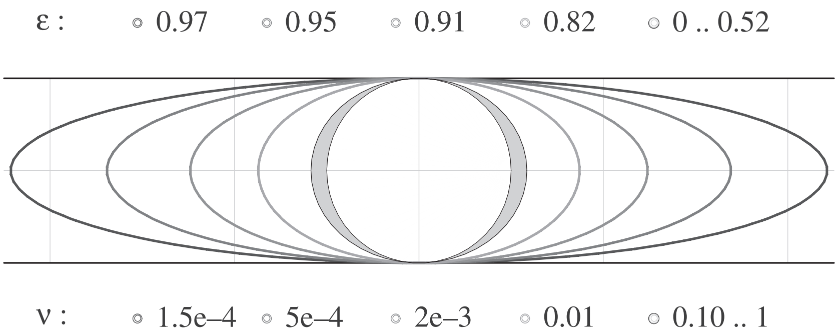


Figure 6.

The manifold Γ for an ellipse does not have a rotational symmetry as do some of the game spaces considered previously. Such a symmetry does appear in the case of a circle. It is useful to discuss first the set Γ for a sphere, where to each position of the point P there exists exactly one position of the player—at the opposite end of the diameter—for which the uniqueness of the geodesic lines with equal length connecting the players is lost. This means that in the appropriate coordinate description Γ is also (is diffeomorphic to) a sphere. Recalling that for 2D game spaces the set Γ is generally 3D, one can see a kind of degeneracy for the sphere, which however can be removed by a deformation of the sphere. In Figure 7 the

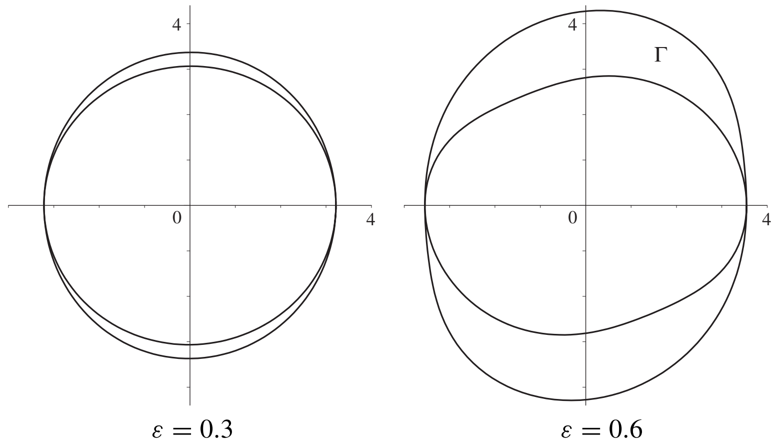


Figure 7.

set Γ is presented for an ellipsoid of revolution with the eccentricity of the ellipse generatrix.

One can prove that two geodesics with equal length exist on the ellipse only when the players are on different sides and on the same confocal (smaller) ellipse. In the case of a disc of radius 1 this means that P and E are on the same circle of some smaller radius r, $0 \leq r \leq 1$. Let us give first the description of Γ for a disc. Let φ, $0 \leq \varphi \leq \pi$ be the angle between the radii OP and OE (O is the center). In the (r, φ)-plane the points of Γ lie in the upper part of the rectangle in Figure 8.

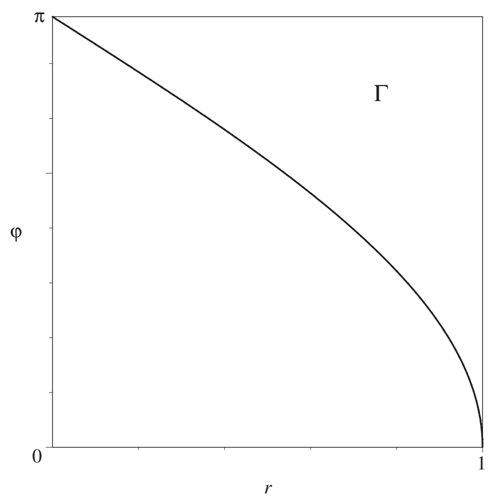


Figure 8.

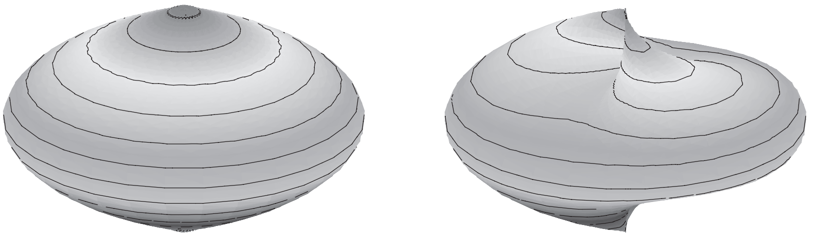


Figure 9. Circle (left) and ellipse, $\varepsilon = 0.8$ (right).

The lower boundary of Γ is represented by the curve $r = \cos(\varphi/2)$. The complete manifold Γ can be obtained by reflection of the picture from the abscissa axis and rotation about the ordinate axis (see Figure 9).

The set Γ for the ellipse has a more complicated structure and is presented on Figure 9 in elliptic coordinates.

6 Secondary Domain, Equivocal Surfaces

In the known examples of the simple pursuit-evasion games, in addition to primary motion, one can meet the following types of optimal paths [7], [13], [15], [18]. The boundary Γ_2 between the sets D_1 and D_2 consists of singular equivocal paths, when each player moves along an envelope of a family of geodesic lines. The region D_2 may consist of two parts, $D_2 = D_2^0 + D_3$, such that in the region D_2^0 each player moves along his own geodesic line, different from the one connecting them. In the region D_3 the pursuer only has to move along a geodesic, while the evader has nonunique optimal choices. Thus, the game value in D_3 is a solution of an optimal control problem for player P only.

In the games on the Euclidean plane with an obstacle there are also parts of the trajectories lying on the obstacle (on the sate constraint) [14], [15], [16].

Depending upon the geometry of the game space the sets D_3 or D_2 may be empty. In the latter case the set D_1 represents the whole game phase space.

Equivocal surface. In the case when the manifold B is nonempty and there exist positions violating conditions (4):

$$L^+ = L^-, \quad |a^+ - a^-|_x - \nu|b^+ - b^-|_y > 2(1-\nu),$$

two branches of a singular equivocal surface (switching envelopes) arise, Γ_2^+ and Γ_2^-. These surfaces can be constructed using the following Cauchy problem in terms of singular characteristics and in inverse time [7]:

$$\dot z = -F_r, \quad \dot r = F_z + \frac{\{\{FF_1\}F\}}{\{\{F_1F\}F_1\}}\left(r - \frac{\partial V_1}{\partial z}(z)\right) \tag{27}$$

$$z(0) = z_0, \quad r(0) = \frac{1}{2}\left(\frac{\partial V_1^+}{\partial z}(z_0) + \frac{\partial V_1^-}{\partial z}(z_0)\right), \quad z_0 \in B.$$

Here $F = F(z, r)$ is the left-hand side of the Hamilton–Jacobi–Bellman–Isaacs (HJBI) equation of the pursuit-evasion game (1) (the Hamiltonian)

$$F(z,r) = -\sqrt{\langle G^{-1}(x)p, p\rangle} + \nu\sqrt{\langle G^{-1}(y)q, q\rangle} + 1 = 0 \tag{28}$$
$$p = \partial V/\partial x, \quad q = \partial V/\partial y, \quad z = (x, y) \in R^4,\ r = (p, q) \in R^4.$$

The function $F_1 = F_1(z, V)$ has the form

$$F_1(z, V) = V - V_1(z).$$

Here and in (27) one has to use the function V_1^+ (V_1^-) for the branch Γ_2^+ (Γ_2^-), where $V_1^\pm$ are given by (2) with $L^\pm$. In (27) one has the Jacobi (Poisson) bracket

$$\{FG\} = \langle F_z + rF_V, G_r\rangle - \langle G_z + rG_V, F_r\rangle$$

for two functions $F(z, V, r)$, $G(z, V, r)$.

Secondary solution. The secondary solution $V = V_2(z)$, $z \in D_2$ is defined as the solution of the HJBI equation (28) with the boundary condition specified on the equivocal surface:

$$V_2(z) = V_1(z), \quad z \in \Gamma_2^+ + B + \Gamma_2^-.$$

The corresponding Cauchy problem is an irregular problem [7] because the characteristics quit the boundary surface Γ_2 tangentially. The set of regular characteristics starting at the manifold B touch both surfaces Γ_2^+ and Γ_2^- and form a *medial surface*. The initial conditions for these characteristics are the same as in (27).

One can show that the optimal trajectories in D_2 for both players are geodesic lines different from the one which connects the points P and E. For more detailed construction algorithm see [7], [11].

7 Optimal Control Subdomain

In [9], [7] it is shown that in the game on a conical surface, in addition to D_1 and D_2^0, there exists a domain D_3 where the game value does not depend on the coordinates of the evader. The possibility of the existence of such a domain can be observed from the analysis of the HJBI equation.

In pursuit-evasion differential games with simple motion the HJBI equation often has a form similar to (28):

$$-H(x, p) + \nu H(y, q) + 1 = 0, \tag{29}$$

where x, y are the coordinate vectors of the pursuer P and the evader E, and p, q are gradients of the game value $V(x, y)$ with respect to x, y, defined in (28).

The function $H(x, p)$ in (29) has the sense of a norm, see (28), and thus is non-negative and vanishes as soon as p and q equal zero:

$$H(x, p) \geq 0,\ H(y, q) \geq 0, \quad H(x, 0) = 0,\ H(y, 0) = 0.$$

Due to these properties in some region D of the x, y-space the following equation may take place:

$$\frac{\partial V}{\partial y} = 0, \quad -H(x, p) + 1 = 0, \quad (x, y) \in D, \tag{30}$$

while the relations

$$\frac{\partial V}{\partial x} = 0, \quad \nu H(x, q) + 1 = 0, \quad (x, y) \in D$$

are not possible. Equation (30) means that the game value in the region D depends only on x and does not depend on y, and consequently, the pursuer has a definite optimal behavior, while the evader may exploit an arbitrary control until he reaches the boundary of D. In other words, in D one has an optimal control problem rather than a dynamic game.

As one can see in the sequel, the relation (30) may have a different analytical form, while the essence is still independent of the game value on the evader's position.

We investigate the solution in D_3 for the game on a surface of revolution. We introduce relative coordinates R, r, φ, where R (respectively r) is the length of the shortest geodesic line connecting the apex $A(0, 0, f(0))$ of the surface M with the point P (point E); φ is the angle between the tangent lines to these two geodesics at the point A in the horizontal plane $x_3 = f(0)$.

Dynamic equations in relative coordinates have the form

$$\dot{R} = u_1, \quad \dot{r} = v_1, \quad \dot{\varphi} = \frac{v_2}{h(r)} - \frac{u_2}{h(R)}$$
$$u_1^2 + u_2^2 \leq 1, \quad v_1^2 + v_2^2 \leq \nu^2$$
$$x = h(R), \quad R = \int_0^x \sqrt{1 + f'^2(\tau)}d\tau,$$

where the function h is defined as an inverse one.

The HJBI equation in relative coordinates could be written as

$$\begin{aligned} F(R, r, \varphi, V_R, V_r, V_\varphi) &= \min_u \max_v [V_R u_1 + V_r v_1 + V_\varphi(v_2/h(r) - u_2/h(R))] + 1 \\ &= -\sqrt{V_R^2 + V_\varphi^2/h^2(R)} + \nu\sqrt{V_r^2 + V_\varphi^2/h^2(r)} + 1 = 0. \end{aligned}$$

In the region D_3 of the relative coordinates the independence of the value on the evader's position means that

$$V_\varphi = 0, \quad V_r = 0,$$

which, under the assumption $V_R > 0$, simplifies the HJBI equation to the differential equation

$$-V_R + 1 = 0$$

with the general solution

$$V = R + C \tag{31}$$

for some constant C.

To compute the constant C one needs to find points where the domains D_1, D_2 and D_3 meet each other. This requires the following geometric analysis.

In the space of relative coordinates R, r, φ one can find such point on the medial surface, introduced in Section 6. In the considered game the medial surface is the plane of symmetry $\varphi = \pi$. Optimal paths lying on this plane obey the regular characteristic system

$$\begin{gathered} \dot{R} = F_{V_R}, \quad \dot{r} = F_{V_r}, \quad \dot{\varphi} = F_{V_\varphi} \\ \dot{V}_R = -F_R, \quad \dot{V}_r = -F_r, \quad \dot{V}_\varphi = -F_\varphi = 0 \end{gathered}$$

with the initial conditions

$$(R, r, \varphi) \in B, \quad V_R = (V_{1R}^+ + V_{1R}^-)/2,\ V_r = (V_{1r}^+ + V_{1r}^-)/2,\ V_\varphi = (V_{1\varphi}^+ + V_{1\varphi}^-)/2.$$

Generally, the optimal phase velocity vector is discontinuous with respect to the initial point at the manifold B. This may happen due to a change of sign of the value V_r in the initial conditions. One can show that on Γ the equality $V_{1r}^+ = V_{1r}^-$ is true, so that $V_r = V_{1r}^+$, and V_r changes sign together with V_{1r}^+. A picture of optimal trajectories on the medial plane $\varphi = \pi$ is shown in Figure 10. At the point $Q = (R_0, r_0)$ one has the critical relation $V_{1r}^+ = 0$.

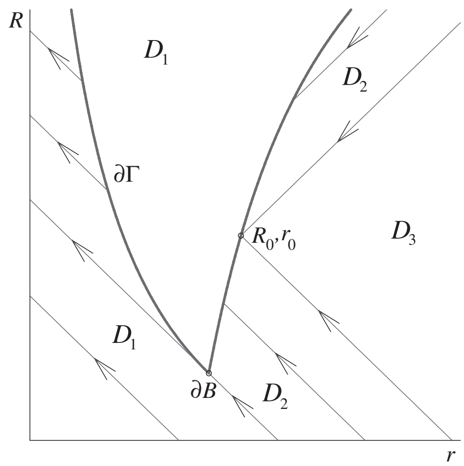


Figure 10.

The optimal paths approaching the set B between the points Q and ∂B obey the system

$$\dot{R} = -1, \quad \dot{r} = \nu, \quad \dot{\varphi} = 0 \quad (V_r \geq 0),$$

and the other paths obey the system

$$\dot{R} = -1, \quad \dot{r} = -\nu, \quad \dot{\varphi} = 0 \quad (V_r \leq 0)$$

so that $\dot{r}$ jumps at Q. Note that the latter system of equations represents an interesting phenomenon: two players move towards each other along a geodesic line, while the "natural" behavior for the evader is to move in the opposite direction.

The point Q is the place where the domains D_1, D_2 and D_3 meet each other. Knowing the coordinates R_0, r_0 of the point Q one can compute the distance between the players $L = R_0 + r_0$. Then equating the expressions (2) and (31) for the game value in the domains D_3 and D_1 at the point Q one gets

$$V = R_0 + C = \frac{R_0 + r_0}{1 - \nu} = V_1, \quad C = \frac{\nu R_0 + r_0}{1 - \nu},$$

which gives the following expression for the game value $V(R)$:

$$V_3 = V(R) = R + \frac{\nu R_0 + r_0}{1 - \nu}. \tag{32}$$

If the point Q tends to infinity the set D_3 is absent. The point Q may coincide with ∂B. In that case the coordinates R_0, r_0 in the above formula must be substituted by the coordinates R_1, r_1 of the point ∂B. The algorithm for construction of R_1, r_1 was given earlier, see (25).

For the point Q one has $b^+ = -b^-$, and (in the notation of Section 5, see (20)) consequently, $y'(x_0) = \infty$. Thus, for the point P one gets: $|a^+ + a^-| = 2(1 - \nu)$. This allows one to get for the computation of R_0, r_0 a two-point problem for the Euler equation with the following conditions:

$$y'(x_0) = \infty, \quad y'^2(x_1) + f'^2(x_1) = \nu(2 - \nu)/(1 - \nu)^2.$$

Thus, generally, the game space consists of three domains: D_1, D_2 and $D = D_3$.

The boundary between the regions D_1 and D_2 is the surface Γ_2 constructed using (27). The boundary between the regions D_1 and D_3 can be constructed by equating the expressions (2) and (32). The boundary between D_2 and D_3 requires more intensive numerical efforts. For the case of a cone ($b = 0$) these surfaces are found in [13], [7].

8 Conclusions

Character of optimal paths. The local analysis of the general case and the solution of a number of particular pursuit-evasion games with simple motion on manifolds shows that optimal trajectories for both players are geodesic lines, either a common line connecting the players, or different geodesics for each player. Singular paths are represented by equivocal trajectories which are envelopes of a family of geodesics. There may also be parts of the trajectories lying on the edge of the game space manifold, like in the game on a plane with an obstacle. In some phase subspace one can meet a nonunique arbitrary optimal motion of the evader, while an optimal control problem for the pursuer arises.

Singular surfaces. The list of singular surfaces arising in the games with simple motion on manifolds includes dispersal, equivocal surfaces, the boundary of the optimal control region and the edges of the game space. An equivocal singular surface (switching envelope) consists of optimal paths, which may also lie on the edge of the manifold. The optimal control region boundary generally does not contain optimal paths, though one can meet such paths in the game on an acute cone, where a degeneracy takes place due to a self-similar solution for the cone.

Primary strategy. The most simple structure has the primary strategy. To obtain it globally one has to check the emptiness of the manifold B. In all considered games, except for the cylinder (plane two-sided strip) and ellipse *that* manifold is nonempty, while an ellipse must not be too "long" for *non emptyness*. The game on a polygon shows *that* for the global optimal primary strategy the boundary of the plane figure must be appropriately smooth.

Open questions. The complete solution of a game on an unbounded surface (of revolution), like a hyperboloid or paraboloid, may exhibit a nondegenerate boundary of the optimal control region and its relation with the primary and secondary domains.

The game on a bounded surface, like a three-axes ellipsoid, may demonstrate another dispersal surface separating motion towards different apexes of the surface.

A relatively simple solution of a simple motion game can be used as an approximation for a more complicated game or serve as an abstraction in the sense of [19].

References

[1] Isaacs, R. *Differential Games.* John Wiley, New York, 1965.

[2] Krasovskii, N.N. and Subbotin, A.I. *Game-Theoretical Control Problems.* Springer-Verlag, New York, 1988.

[3] Subbotin, A.I. *Generalized Solutions of First Order PDEs: the Dynamical Optimization Perspective*. Birkhäuser, Boston, 1995.

[4] Petrosjan, L.A. and Zenkevich, N.A. *Game Theory.* World Scientific Publ., Singapore-London, 1996.

[5] Bernhard, P. Singular Surfaces in Differential Games: An Introduction. In: *Differential Games and Applications*. Springer-Verlag, Berlin, 1977, pp. 1–33.

[6] Patsko, V.S. and Turova, V.L. Level sets of the value function in differential games with the homicidal chauffeur dynamics. Intern. Game Theory Review. 2001, Vol. 3. N 1, pp. 67–112.

[7] Melikyan, A.A. *Generalized Characteristics of First Order PDEs: Applications in Optimal Control and Differential Games.* Birkhäuser, Boston, 1998.

[8] Melikyan, A.A., Hovakimyan, N.V. and Harutunian, L.L. Games of simple pursuit and approach on two dimensional cone. JOTA, 1998, Vol. 98, N 3, pp. 515–543.

[9] Hovakimyan, N.V. and Melikyan, A.A. Geometry of pursuit-evasion on second order rotation surfaces. Dynamics and Control, 2000, Vol. 10, pp. 297–312.

[10] Hovakimyan, N.V. and Harutunian, L.L. Game problems on rotation surfaces. In: Petrosjan, L.A. and Matalov, V.V. (editors), *Game Theory and Applications*, 1998, Vol. IV, pp. 62–74.

[11] Melikyan, A.A. Games of simple pursuit and approach on the manifolds. Dynamics and Control. 1994, Vol. 4. N 4. pp. 395–405.

[12] Demyanov, V.F. *Minimax: Directional Differentiability.* Leningrad State University Press, Leningrad, 1974 (in Russian).

[13] Melikyan, A.A. and Hovakimyan, N.V. Game Problem of Simple Pursuit on Two-Dimensional Cone. J. Appl. Math. Mech. (PMM), 1991, Vol. 55, No 5, pp. 741–751.

[14] Melikyan, A.A. Games of pursuit on two-sided plane with a hole. 8th Intern. Symp. on Dynamic Games and Applications. Maastricht, the Netherlands, 1998. pp. 394–396.

[15] Vishnevetsky, L.S. and Melikyan, A.A. Optimal pursuit on a plane in the presence of an obstacle. J. Appl. Math. Mech. (PMM), 1982, Vol. 46, N 4, pp. 485–490.

[16] Pozharitsky, G.K. Isaacs' problem of island enveloping. J. Appl. Math. Mech. (PMM), 1982, Vol. 46, N 5, pp. 707–713.

[17] Gelfand, I.M. and Fomin, S.V. *Calculus of Variations.* Prentice-Hall, Englewood Cliff, NJ, 1963.

[18] Melikyan, A.A. Some properties of the Bellman-Isaacs equation for the games on the surfaces of revolution. International Game Theory Review, 2001, Vol. 6, N 1, pp. 171–179.

[19] Pappas, G.J. and Simic, S. Consistent Abstractions of Affine Control Systems. IEEE Trans. on Automatic Control, 2002, Vol. 47, N 5, pp. 745–756.

Solution of a Linear Pursuit-Evasion Game with Variable Structure and Uncertain Dynamics

Josef Shinar
Faculty of Aerospace Engineering
Technion, Israel Institute of Technology
Haifa 32000, Israel

Valery Y. Glizer
Faculty of Aerospace Engineering
Technion, Israel Institute of Technology
Haifa 32000, Israel

Vladimir Turetsky
Faculty of Aerospace Engineering
Technion, Israel Institute of Technology
Haifa 32000, Israel

Abstract

A class of pursuit-evasion differential games with bounded controls and a prescribed duration is considered. Two finite sets of possible dynamics of the pursuer and evader, known for both players, are given. The evader chooses his dynamics once before the game starts. This choice is unavailable for the pursuer, which causes a dynamics uncertainty. The pursuer can change his dynamics a finite number of times during the game, yielding a variable structure dynamics. The solution of this game is derived including optimal strategies of the players. The existence of a saddle point is shown. The game value and the shape of the maximal capture zone are obtained. Illustrative examples are presented.

1 Introduction

Control problems with uncertain dynamics are of considerable interest both from theoretical and practical viewpoints. Variable structure systems are not of less importance. There exists a rich bibliography in the open literature devoted to these topics. Most of these works study controlled systems with a single decision maker (see, e.g., [1–7] and references therein), while games with either uncertain or variable structure dynamics are investigated much less often. In [8] a class of N-player Nash linear-quadratic differential games was considered. The equations of motion of each player have a norm bounded uncertainty in the matrix of coefficients for the state vector. The notion of a robust Nash solution of the game was introduced.

Sufficient conditions guaranteeing the existence of this solution were obtained and the solution itself was designed. In [9] a two-player linear-quadratic pursuit-evasion game, as well as more a general Nash equilibrium game, were considered. This work deals with multimodel games. The dynamics of each of these games depends on a parameter, which belongs to a given finite set of numbers. This parameter is constant during the game, however, none of the players has information on its value. Robust (with respect to the parameter) solutions of these games were obtained by extending the minimax control approach [10–12]. Differential games of pursuit with a variable structure dynamics were considered in [13], for the case of a group of pursuers and a single evader. The structure of the dynamics is changed by the evader once during the game. Sufficient conditions for the existence of the game solution were obtained. This research was continued in [14,15].

In this chapter, the combination of two features, variable structure and uncertain dynamics, is considered for a class of pursuit-evasion games, which are the mathematical model of an interception engagement between two moving vehicles, an interceptor P (*pursuer*) and a target E (*evader*). In this class the dynamics of each vehicle is expressed by a first-order transfer function with time constants τ_p and τ_e, respectively. Moreover, it is assumed that the lateral accelerations of the pursuer and the evader are respectively bounded by the constants $a_p^{\max}$ and $a_e^{\max}$. Thus, the dynamics of the players is completely described by the vectors $\omega_p = (a_p^{\max}, \tau_p)$ and $\omega_e = (a_e^{\max}, \tau_e)$. The cost function is the miss distance, the distance of closest approach, nonnegative by definition. The version of this game with prescribed vectors ω_p and ω_e has been studied extensively in the open literature [16–18]. It was shown that this game version has a saddle point solution in feedback strategies. The solution leads to the decomposition of the game space into two regions (singular and regular) of different optimal strategies. The game space decomposition is completely determined by the pair (ω_p, ω_e).

In this chapter, it is assumed that the vectors ω_e and ω_p belong, respectively, to given finite sets $\Omega_e = \{\omega_e^i,\ i = 1, \ldots, N_e\}$ and $\Omega_p = \{\omega_p^j,\ j = 1, \ldots, N_p\}$ known to both players. The evader chooses the vector $\omega_e \in \Omega_e$ once before the game starts and the pursuer does not know the evader's choice. The choice of ω_e by the evader can be considered as its additional control. In the game formulation, it is supposed that the evader chooses ω_e, knowing the initial position of the game. This situation is the worst case for the pursuer. Since the information on ω_e chosen by the evader is unavailable for the pursuer, this game is a game with uncertain dynamics. The pursuer can choose $\omega_p \in \Omega_p$ a finite number of times during the game; therefore, the dynamics of the game has a variable structure. The choice of the pursuer is not known to the evader, and it can be considered as an additional pursuer control.

The chapter is organized as follows. In Section 2 the problem statement is presented in detail. The known results for the case $N_e = N_p = 1$, used in the sequel, are briefly reviewed as well. In Section 3 the solution of the game is obtained separately for the cases (N_e arbitrary, $N_p = 1$), ($N_e = 1$, N_p arbitrary) and (N_e and N_p arbitrary). In Section 4 the application of the game solution to an interception problem is presented. Concluding remarks are presented in Section 5. An auxiliary lemma is proven in the Appendix.

2 Problem Statement

2.1 Engagement Model

The engagement between two moving objects (*the players*)—an interceptor (*pursuer*) and a target (*evader*)—is considered. The mathematical model of this scenario is based on the following assumptions:

- the engagement takes place in a plane;
- both players have constant velocities and bounded lateral accelerations;
- the dynamics of each player is expressed by a first-order transfer function;
- the relative trajectory can be linearized with respect to the nominal collision geometry;
- perfect state information is available.

In Figure 1 the schematic view of the interception geometry is shown. Figure 1a shows the collision triangle.

The x axis of the coordinate system is aligned with the initial line of sight, where P_0 and E_0 are the respective initial positions of the pursuer and evader. The origin is collocated with P_0. The pursuer's collision angle $(\phi_p)_{\text{col}}$, aiming its velocity vector to the collision point C, is determined by the condition

$$V_p \sin{(\phi_p)_{\text{col}}} = V_e \sin \phi_e^0, \tag{1}$$

where V_p and V_e are the constant velocities of the players, and ϕ_e^0 is the initial angle between the evader's velocity vector and the reference line of sight. It is

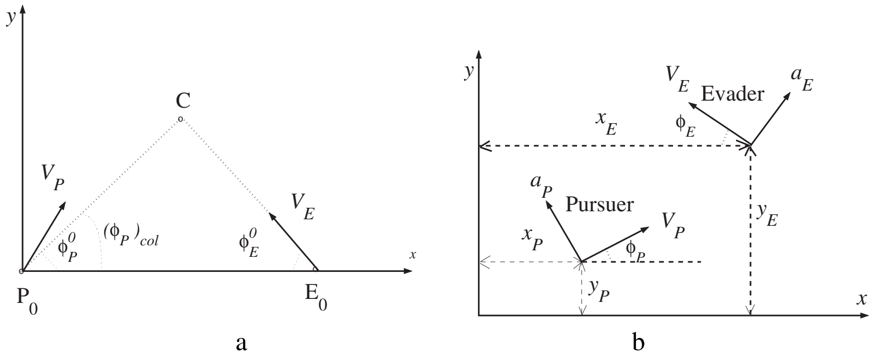


Figure 1: Interception geometry.

assumed that ϕ_p^0 differs only slightly from $(\phi_p)_{\text{col}}$:

$$|\phi_p^0 - (\phi_p)_{\text{col}}| \ll 1. \tag{2}$$

In Figure 1b the current positions are depicted. The points (x_p, y_p), (x_e, y_e) are current coordinates of the players; a_p, a_e are the respective lateral accelerations; ϕ_p, ϕ_e are the respective angles between the velocity vectors and the x-axis (reference line of sight). It is assumed that the actual velocity vector of the pursuer remains close to the collision course:

$$|\phi_p(t) - (\phi_p)_{\text{col}}| \ll 1. \tag{3}$$

It is also assumed that the actual velocity vector of the evader remains close to the initial condition:

$$|\phi_e(t) - \phi_e^0| \ll 1. \tag{4}$$

Based on (3), (4), the trajectories of the pursuer and the evader can be linearized with respect to the nominal collision geometry leading to a constant closing velocity

$$V_c = V_p \cos{(\phi_p)}_{\text{col}} - V_e \cos\phi_e^0. \tag{5}$$

The final interception time t_f can be easily calculated for any given initial range r_0,

$$t_f = r_0/V_c. \tag{6}$$

These assumptions lead to the following linearized model for $0 \le t \le t_f$:

$$\begin{aligned}
&\dot{x}_1 = x_2, && x_1(0) = 0,\\
&\dot{x}_2 = x_3 - x_4, && x_2(0) = x_{20} = V_e \sin\phi_e^0 - V_p \sin\phi_p^0,\\
&\dot{x}_3 = (a_e^{\max} v - x_3)/\tau_e, && x_3(0) = 0,\\
&\dot{x}_4 = (a_p^{\max} u - x_4)/\tau_p, && x_4(0) = 0,
\end{aligned} \tag{7}$$

where $x_1 = y_e - y_p$ is the relative separation normal to the initial line of sight; x_2 is the relative normal velocity; x_3 and x_4 are the lateral accelerations of the evader and the pursuer, respectively, both normal to the initial line of sight; τ_e, τ_p are the respective time constants; $a_e^{\max}$, $a_p^{\max}$ are the respective maximal values of the lateral accelerations.

The normalized lateral acceleration commands of the evader $v(t)$ and the pursuer $u(t)$ satisfy the constraints

$$|v(t)| \le 1, \quad 0 \le t \le t_f, \tag{8}$$

$$|u(t)| \le 1, \quad 0 \le t \le t_f. \tag{9}$$

It is supposed that the functions $v(t)$, $u(t)$ are measurable on $[0, t_f]$ and are bounded according to (8), (9). The system (7) can be rewritten in a matrix form as

$$\dot{x} = Ax + bu + cv, \quad x(0) = x_0, \tag{10}$$

where $x^T = (x_1, x_2, x_3, x_4)$, $x_0^T = (0, x_{20}, 0, 0)$,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1/\tau_e & 0 \\ 0 & 0 & 0 & -1/\tau_p \end{bmatrix}, \tag{11}$$

$$b^T = (0, 0, 0, a_p^{\max}/\tau_p), \quad c^T = (0, 0, a_e^{\max}/\tau_e, 0), \tag{12}$$

and the superscript T denotes the transposition.

Remark 2.1. The engagement model (10) is completely determined by two vectors $\omega_e = (a_e^{\max}, \tau_e)$ and $\omega_p = (a_p^{\max}, \tau_p)$, chosen respectively by the evader and the pursuer from the prescribed finite sets $\Omega_e = \{\omega_e^i, \ \ i = 1, \ldots, N_e\}$ and $\Omega_p = \{\omega_p^i, \ \ i = 1, \ldots, N_p\}$. These sets are known to both players.

For any fixed $\omega_e \in \Omega_e$ and $\omega_p \in \Omega_p$, (10) can be reduced to a scalar form by using the transformation (see [19,20])

$$z(t) = z(t; \omega_e, \omega_p) = d^T \Phi(t_f, t; \tau_e, \tau_p) x(t; \omega_e, \omega_p), \tag{13}$$

where $x(t; \omega_e, \omega_p)$ is the state vector of (10), $\Phi(t_f, t; \tau_e, \tau_p)$ is the transition matrix of the homogeneous system $\dot{x} = Ax$, the final time t_f is given by (6) and

$$d^T = (1, 0, 0, 0). \tag{14}$$

The new state variable is given explicitly by

$$z(t) = x_1(t) + (t_f - t)x_2(t) + \tau_e^2 \Psi((t_f - t)/\tau_e)x_3(t) - \tau_p^2 \Psi((t_f - t)/\tau_p)x_4(t), \tag{15}$$

where

$$\Psi(\xi) \stackrel{\Delta}{=} \exp(-\xi) + \xi - 1 > 0, \quad \xi > 0. \tag{16}$$

Introducing the new independent variable (time-to-go)

$$\vartheta = t_f - t, \tag{17}$$

and using (15) yield the differential equation

$$z' = h(\vartheta, \tau_p, a_p^{\max})u - h(\vartheta, \tau_e, a_e^{\max})v, \tag{18}$$

where derivation with respect to ϑ is denoted by prime,

$$h(\vartheta, \tau, a^{\max}) = \tau a^{\max} \Psi(\vartheta/\tau), \tag{19}$$

and

$$\vartheta_0 = t_f, \quad z(\vartheta_0) = z_0 = \vartheta_0 x_{20}, \tag{20}$$

$$z(0) = x_1(t_f). \tag{21}$$

Due to (17), the controls $u(t)$ and $v(t)$ become $u(t_f - \vartheta)$ and $v(t_f - \vartheta)$, respectively, denoted in the sequel as $u(\vartheta)$ and $v(\vartheta)$. Note that these functions are measurable on $[0, \vartheta_0]$. The constraints (8), (9) become

$$|v(\vartheta)| \le 1, \quad 0 \le \vartheta \le \vartheta_0, \tag{22}$$

$$|u(\vartheta)| \le 1, \quad 0 \le \vartheta \le \vartheta_0. \tag{23}$$

Consider the game with the dynamics (7), constraints (8), (9) and performance index

$$\mathcal{J} = |x_1(t_f)|. \tag{24}$$

Note that (24) is the miss distance. The objective of the pursuer is minimizing and of the evader is maximizing (24), by means of feedback strategies $u(t, x)$ and $v(t, x)$, respectively.

Due to (21), $\mathcal{J} = |z(0)|$. Thus, the transformation (13), (17) reduces the game (7), (8), (9), (24) to the game with the scalar dynamics (18), constraints (22), (23) and performance index

$$J = |z(0)|. \tag{25}$$

It is assumed that the feedback strategies $u(\vartheta, z)$ and $v(\vartheta, z)$ satisfy

$$u(\cdot) \in \mathcal{F}, \quad v(\cdot) \in \mathcal{F}, \tag{26}$$

where $\mathcal{F}$ is the set of functions $f(\vartheta, z)$, $(\vartheta, z) \in [0, \infty) \times R$, which are measurable with respect to Θ for any z, piecewise continuous with respect to z for any ϑ and satisfy $|f(\vartheta, z)| \le 1$.

2.2 The Game with $N_e = N_p = 1$

In [17,18] the game (18), (22), (23), (25) was solved for the case $N_e = N_p = 1$ under the assumption that both players have perfect information on the state z. Due to Remark 2.1, in this case the vectors ω_e and ω_p are known to each player. In the sequel, this game is called a (1,1)-game.

The solution of a (1,1)-game [17,18] leads to the decomposition of the game space (ϑ, z) into two regions of different strategies.

In the first (*singular*) region D_0 the optimal control strategies $u^0(\vartheta, z)$ and $v^0(\vartheta, z)$ are *arbitrary* subject to (22), (23), and the value of the game is constant. In the second (*regular*) region $D_1 = R^2 \backslash D_0$ the optimal strategies have a "bang-bang" structure:

$$u^0(\vartheta, z) = v^0(\vartheta, z) = \operatorname{sign} z(\vartheta), \tag{27}$$

and the value of the game is nonzero, depending on the initial conditions. Note that D_0 and D_1 are symmetrical with respect to the ϑ axis.

In the case

$$a_p^{\max} > a_e^{\max}, \frac{a_p^{\max}}{\tau_p} \geq \frac{a_e^{\max}}{\tau_e}, \tag{28}$$

the singular region is

$$D_0 = D_0(\omega_e, \omega_p) = \{(\vartheta, z) : \vartheta > 0,\ |z| < z^*(\vartheta, \omega_e, \omega_p)\}, \tag{29}$$

where

$$z^*(\vartheta, \omega_e, \omega_p) = \int_0^{\vartheta} H(\xi, \omega_e, \omega_p) d\xi, \tag{30}$$

$$H(\xi, \omega_e, \omega_p) \stackrel{\Delta}{=} h(\xi, \tau_p, a_p^{\max}) - h(\xi, \tau_e, a_e^{\max}) > 0, \xi > 0. \tag{31}$$

An example of the decomposition (for $t_f = 4$ s, $\tau_p = 0.2$ s, $a_p^{\max} = 150$ m/s^2, $\tau_e = 0.16$ s, $a_e^{\max} = 100$ m/s^2) is shown in Figure 2.

For any initial position (ϑ_0, z_0) the value of the game is given by

$$J^* = J^*(\vartheta_0, z_0, \omega_e, \omega_p) = \begin{cases} 0, & (\vartheta_0, z_0) \in D_0(\omega_e, \omega_p), \\ |z_0| + \displaystyle\int_{\vartheta_0}^{0} H(\vartheta, \omega_e, \omega_p) d\vartheta, & (\vartheta_0, z_0) \notin D_0(\omega_e, \omega_p). \end{cases} \tag{32}$$

Thus, if the inequalities (28) are satisfied, the singular region D_0 becomes the maximal *capture zone*, i.e., the maximal set of the initial positions, from which the pursuer can guarantee zero miss distance against any admissible evader strategy.

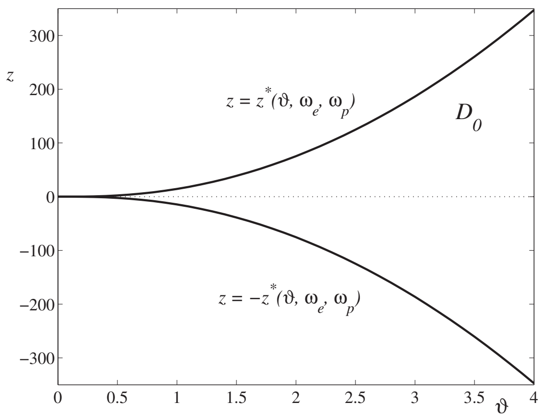


Figure 2: The game space decomposition.

2.3 The Game with Arbitrary N_e, N_p

The new game, called in the sequel the (N_e, N_p)-game, has the following information structure:

- Both players have the perfect information on the current position (ϑ, z).
- The evader chooses his vector $\omega_e \in \Omega_e$, once at the beginning of the game, based on the information on the initial position (ϑ_0, z_0). The pursuer knows the full set Ω_e but has no information on the vector ω_e chosen by the evader.
- The pursuer chooses his vector $\omega_p \in \Omega_p$ a finite number of times during the game, based on the information on the current position (ϑ, z). The evader knows the full set Ω_p but has no information either on the current vector ω_p, or on its eventual changes.

The choice of ω_e can be considered as an additional evader control. Moreover, the evader strategy v depends, in general, on the chosen vector ω_e, i.e., it can be represented as a function of ϑ, z, ω_e. Note that for any $\omega_e \in \Omega_e$, $v(\vartheta, z, \omega_e) \in \mathcal{F}$. The pair

$$V(\cdot) = V(\vartheta_0, z_0, \vartheta, z) \stackrel{\Delta}{=} (\omega_e(\vartheta_0, z_0), v(\vartheta, z, \omega_e)) \in \Omega_e \times \mathcal{F} \qquad (33)$$

is called the *evader extended strategy*.

Let $\mathcal{O}_p$ be the set of functions mapping $[0, \vartheta_0] \times R$ onto Ω_p, such that for each ϑ they are piecewise constant in z and for each z they are piecewise constant in ϑ. The choice of $\omega_p(\cdot) \in \mathcal{O}_p$ is in fact an additional pursuer control. The pursuer strategy u depends on ϑ, z and, in general, on the chosen current vector ω_p. However, since ω_p depends on ϑ, z, the strategy u depends only on ϑ, z. The pair

$$U(\cdot) = U(\vartheta, z) \stackrel{\Delta}{=} (\omega_p(\vartheta, z), u(\vartheta, z)) \in \mathcal{O}_p \times \mathcal{F} \qquad (34)$$

is called the *pursuer extended strategy*.

The objective of the pursuer is minimizing (25) by means of the extended feedback strategy $U(\cdot)$, while the evader wants to maximize (25) by means of the extended strategy $V(\cdot)$. Note that in this (N_e, N_p)-game the performance index (25) becomes $J = J(U(\cdot), V(\cdot))$.

Definition 1. The extended strategy $U^*(\vartheta, z) \in \mathcal{O}_p \times \mathcal{F}$ is called the *optimal pursuer extended strategy* in the (N_e, N_p)-game, if for any $U(\cdot) \in \mathcal{O}_p \times \mathcal{F}$,

$$\max_{V(\cdot)\in\Omega_e\times\mathcal{F}} J(U^*(\cdot), V(\cdot)) \leq \max_{V(\cdot)\in\Omega_e\times\mathcal{F}} J(U(\cdot), V(\cdot)). \qquad (35)$$

Definition 2. For any initial position (ϑ_0, z_0), the extended strategy $V^*(\vartheta_0, z_0, \vartheta, z) \in \Omega_e \times \mathcal{F}$ is called the *optimal evader extended strategy* in the (N_e, N_p)-game, if for any $V(\cdot) \in \Omega_e \times \mathcal{F}$,

$$\min_{U(\cdot)\in\mathcal{O}_p\times\mathcal{F}} J(U(\cdot), V^*(\cdot)) \geq \min_{U(\cdot)\in\mathcal{O}_p\times\mathcal{F}} J(U(\cdot), V(\cdot)). \qquad (36)$$

3 Solution of the Scalar (N_e, N_p)-Game

The (N_e, N_p)-game is solved by its decomposition into two simpler games, namely, the $(N_e,1)$-game and the $(1,N_p)$-game.

3.1 Solution of the Scalar $(N_e, 1)$-Game

In this game $\Omega_p = \{\omega_p\}$.

Remark 3.1. In the $(N_e,1)$-game, the choice of the vector $\omega_e \in \Omega_e$ is equivalent to the choice of the number $i \in I_e \overset{\Delta}{=} \{1, \ldots, N_e\}$. The chosen vector ω_e^i is kept constant during the game and completely determines the game dynamics model (18). Therefore, the pair (ω_e^i, ω_p) can be called the *model* of this game. Due to the information structure of the (N_e, N_p)-game (see Section 2.3), the pursuer knows the set Ω_e but the specific vector $\omega_e^i \in \Omega_e$, chosen by the evader at the beginning of the game, is not known to the pursuer. The evader has the perfect information on the model.

Remark 3.2. In this game, the extended pursuer strategy is reduced to the regular one $u(\vartheta, z) \in \mathcal{F}$. Since the choice of $\omega_e \in \Omega_e$ is equivalent to the choice of the model number $i \in I_e$, the evader extended strategy can be rewritten as $V(\cdot) = (i(\vartheta_0, z_0), v(\vartheta, z, i)) \in I_e \times \mathcal{F}$.

Due to Remarks 3.1 and 3.2, Definitions 1, 2 can be rewritten in the following form.

Definition 1_e. The strategy $u^*(\vartheta, z) \in \mathcal{F}$ is called the *optimal pursuer strategy* in the $(N_e,1)$-game, if for any $u(\cdot) \in \mathcal{F}$,

$$\max_{V(\cdot)\in I_e\times\mathcal{F}} J(u^*(\cdot), V(\cdot)) \le \max_{V(\cdot)\in I_e\times\mathcal{F}} J(u(\cdot), V(\cdot)). \tag{37}$$

Definition 2_e. For any initial position (ϑ_0, z_0), the extended strategy $V^*(\vartheta_0, z_0, \vartheta, z) \in I_e \times \mathcal{F}$ is called the *optimal extended evader strategy* in the $(N_e,1)$-game, if for any $V(\cdot) \in I_e \times \mathcal{F}$,

$$\min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), V^*(\cdot)) \ge \min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), V(\cdot)). \tag{38}$$

The $(N_e, 1)$-game can be partitioned into N_e separate (1,1)-games, denoted by $(1, 1)_i$-game, $i = 1, \ldots, N_e$. Due to Section 2.2, each $(1, 1)_i$-game generates its singular region $D_0(\omega_e^i, \omega_p)$. The optimal strategies in the $(N_e, 1)$-game are designed using the set

$$\mathcal{D}_0^e = \bigcap_{i=1}^{N_e} D_0(\omega_e^i, \omega_p). \tag{39}$$

By (29)–(31), this set can be described analytically as follows:

$$\mathcal{D}_0^e = \{(\vartheta, z) : \vartheta > 0,\ |z| < Z_e(\vartheta)\}, \tag{40}$$

where for each $\vartheta > 0$,

$$\begin{aligned} Z_e(\vartheta) &= \min_{i \in I_e} \int_0^{\vartheta} H(\xi, \omega_e^i, \omega_p) d\xi = \min_{i \in I_e} z^*(\vartheta, \omega_e^i, \omega_p) \\ &= \int_0^{\vartheta} h(\xi, \tau_p, a_p^{\max}) d\xi - \max_{i \in I_e} \int_0^{\vartheta} h(\xi, \tau_{ei}, a_{ei}^{\max}) d\xi, \end{aligned} \tag{41}$$

where $h(\xi, \tau, a^{\max})$ is given by (19).

It follows from Lemma A.1 (see the Appendix) that the curves $z = z^*(\vartheta, \omega_e^i, \omega_p)$, $i \in I_e$, intersect each other a finite number of times. This means that $i(\vartheta) = \arg\min_{i \in I_e} z^*(\vartheta, \omega_e^i, \omega_p)$ changes a finite number of times for $\vartheta > 0$. Therefore, $Z_e(\vartheta)$ is a continuous and piecewise differentiable function for $\vartheta > 0$ with a finite number of the nonsmoothness points.

Based on (39), the following theorem presents the optimal strategies of the $(N_e,1)$-game.

Theorem 3.1. *The optimal pursuer strategy $u^*(\cdot) \in \mathcal{F}$ is given by*

$$u^*(\vartheta, z) = \begin{cases} \text{arbitrary}, & (\vartheta, z) \in \mathcal{D}_0^e, \\ \text{sign}(z), & (\vartheta, z) \notin \mathcal{D}_0^e. \end{cases} \tag{42}$$

For any initial position (ϑ_0, z_0) the optimal extended evader strategy $V^*(\cdot) = \{i^*(\cdot), v^*(\cdot)\} \in I_e \times \mathcal{F}$ is given as follows. If $(\vartheta_0, z_0) \in \mathcal{D}_0^e$, then $i^*(\vartheta_0, z_0) \in I_e$ is arbitrary and

$$v^*(\vartheta, z, i^*) = \begin{cases} \text{arbitrary}, & (\vartheta, z) \in D_0(\omega_e^{i^*}, \omega_p), \\ \text{sign}(z), & (\vartheta, z) \notin D_0(\omega_e^{i^*}, \omega_p). \end{cases} \tag{43}$$

If $(\vartheta_0, z_0) \notin \mathcal{D}_0^e$, then

$$i^*(\vartheta_0, z_0) = \arg\min_{i \in I_e} z^*(\vartheta_0, \omega_e^i, \omega_p), \tag{44}$$

and

$$v^*(\vartheta, z, i^*) = \text{sign}(z). \tag{45}$$

Proof. Note that due to (39),

$$\mathcal{D}_0^e \subseteq D_0(\omega_e^i, \omega_p),\ i \in I_e. \tag{46}$$

Pursuer strategy. Due to Definition 1_e, the optimal pursuer strategy has to satisfy (37). The maximum in (37) for any $u(\cdot) \in \mathcal{F}$ can be rewritten as

$$\max_{V(\cdot)\in I_e\times\mathcal{F}} J(u(\cdot), V(\cdot)) = \max_{i\in I_e} \max_{v(\cdot)\in\mathcal{F}} J\Big(u(\cdot), \{i, v(\cdot)\}\Big). \tag{47}$$

Due to (46), the strategy (42) is an optimal pursuer strategy in the (1,1)-game for each model (ω_e^i, ω_p), $i \in I_e$ (see Section 2.2). Therefore, for any $u(\cdot) \in \mathcal{F}$,

$$\max_{v(\cdot)\in\mathcal{F}} J(u(\cdot), \{i, v(\cdot)\}) \geq \max_{v(\cdot)\in\mathcal{F}} J(u^*(\cdot), \{i, v(\cdot)\}), \quad i \in I_e. \tag{48}$$

Due to (47) and (48), the inequality (37) is valid, which proves the optimality of (42).

Evader strategy. Due to Definition 2_e, the optimal extended evader strategy has to satisfy (38).

Let $(\vartheta_0, z_0) \in \mathcal{D}_0^e$. In this case, due to (43), $v^*(\cdot)$ is an optimal evader strategy in the (1,1)-game for the model $(\omega_e^{i^*}, \omega_p)$ (see Section 2.2). Hence,

$$\min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), V^*(\cdot)) = \min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), \{i^*, v^*(\cdot)\}) = 0. \tag{49}$$

Proceed to the minimum in the right-hand side of (38). Let $V(\cdot) \in I_e \times \mathcal{F}$ be arbitrary. Let $u_i^0(\vartheta, z)$ be an optimal pursuer strategy in the (1,1)-game for the model (ω_e^i, ω_p), $i \in I_e$. Then for any $v(\cdot) \in \mathcal{F}$,

$$J\big(u_i^0(\cdot), \{i, v(\cdot)\}\big) = 0, i \in I_e, \tag{50}$$

and consequently

$$\min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), \{i, v(\cdot)\}) = 0, \quad i \in I_e. \tag{51}$$

Due to (49) and (51), the inequality (38) is valid, which proves the optimality of $V^*(\cdot)$, given by (43), for $(\vartheta_0, z_0) \in \mathcal{D}_0^e$.

Let $(\vartheta_0, z_0) \notin \mathcal{D}_0^e$. Due to (44),

$$z^*(\vartheta_0, \omega_e^{i^*}, \omega_p) \leq z^*(\vartheta_0, \omega_e^i, \omega_p), \quad i \in I_e. \tag{52}$$

Due to (29) and (39), in this case $(\vartheta_0, z_0) \notin D_0(\omega_e^{i^*}, \omega_p)$, which means that $|z_0| \geq z^*(\vartheta_0, \omega_e^{i^*}, \omega_p)$.

By (44) and (45), $V^*(\cdot) = \{i^*, \operatorname{sign} z\}$, which by using (30) and (32), directly yields

$$\min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), V^*(\cdot)) = |z_0| - z^*(\vartheta_0, \omega_e^{i^*}, \omega_p). \tag{53}$$

Proceed to the minimum in the right-hand side of (38). Let $V(\cdot) \in I_e \times \mathcal{F}$ be arbitrary. If $i = i^*$, then for any $v(\cdot) \in \mathcal{F}$,

$$\min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), \{i^*, v(\cdot)\}) \leq |z_0| - z^*(\vartheta_0, \omega_e^{i^*}, \omega_p). \tag{54}$$

If $i \neq i^*$,

$$\min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), \{i, v(\cdot)\}) \leq J^*(\vartheta_0, z_0, \omega_e^i, \omega_p). \tag{55}$$

Due to (32) and (52),

$$J^*(\vartheta_0, z_0, \omega_e^i, \omega_p) < |z_0| - z^*(\vartheta_0, \omega_e^{i^*}, \omega_p), \quad i \in I_e \backslash \{i^*\}. \tag{56}$$

The equation (53) and the inequalities (54)–(56) directly yield the inequality (38). This completes the proof of the theorem. □

Remark 3.3. It is important to note that, due to (41) and (44), the optimal evader number $i^*(\vartheta_0, z_0)$ is independent of ω_p. However, the set $\mathcal{D}_0^e$ depends on ω_p.

Corollary 3.1. *For any fixed initial position* (ϑ_0, z_0)*,*

$$J(u^*(\cdot), V^*(\cdot)) = \begin{cases} 0, & (\vartheta_0, z_0) \in \mathcal{D}_0^e, \\ |z_0| - z^*(\vartheta_0, \omega_e^{i^*}, \omega_p), & (\vartheta_0, z_0) \notin \mathcal{D}_0^e, \end{cases} \tag{57}$$

where $i^* = i^*(\vartheta_0, z_0)$ *is given by (44). Moreover, for any* $u(\cdot) \in \mathcal{F}$*,* $V(\cdot) \in I_e \times \mathcal{F}$*, the saddle point inequality is satisfied:*

$$J(u^*(\cdot), V(\cdot)) \leq J(u^*(\cdot), V^*(\cdot)) \leq J(u(\cdot), V^*(\cdot)), \tag{58}$$

i.e., $(u^*(\cdot), V^*(\cdot))$ *is a saddle point in the* $(N_e, 1)$*-game and (57) is the game value.*

Proof. Since $u^*(\cdot)$ is an optimal strategy in the (1,1)-game for any model (ω_e^i, ω_p), $i \in I_e$, the following equality holds for any $V(\cdot) \in I_e \times \mathcal{F}$:

$$\min_{u(\cdot)\in\mathcal{F}} J(u(\cdot), V(\cdot)) = J(u^*(\cdot), V(\cdot)). \tag{59}$$

Let $(\vartheta_0, z_0) \in \mathcal{D}_0^e$. Then (51) and (59) yield

$$J(u^*(\cdot), V(\cdot)) = 0 \;\; \forall\, V(\cdot) \in I_e \times \mathcal{F}, \tag{60}$$

and consequently

$$J(u^*(\cdot), V^*(\cdot)) = 0. \tag{61}$$

Thus, (57) is proven for the case $(\vartheta_0, z_0) \in \mathcal{D}_0^e$. Furthermore, (60) and (61) yield the left-hand part of (58). The right-hand part of (58) is a direct consequence of (61) and the inequality $J(u(\cdot), V(\cdot)) \geq 0$ for any $u(\cdot) \in \mathcal{F}$, $V(\cdot) \in I_e \times \mathcal{F}$.

Let $(\vartheta_0, z_0) \notin \mathcal{D}_0^e$. In this case (57) directly follows from (53) and (59). The left-hand part of (58) is a consequence of (54)–(56) and (59), while the right-hand part is a consequence of (59). This completes the proof of the corollary. □

Remark 3.4. It can be seen from the proofs of Theorem 3.1 and Corollary 3.1 that if $(\vartheta_0, z_0) \in \mathcal{D}_0^e$, then any $V(\cdot) \in I_e \times \mathcal{F}$ can be taken as an optimal extended

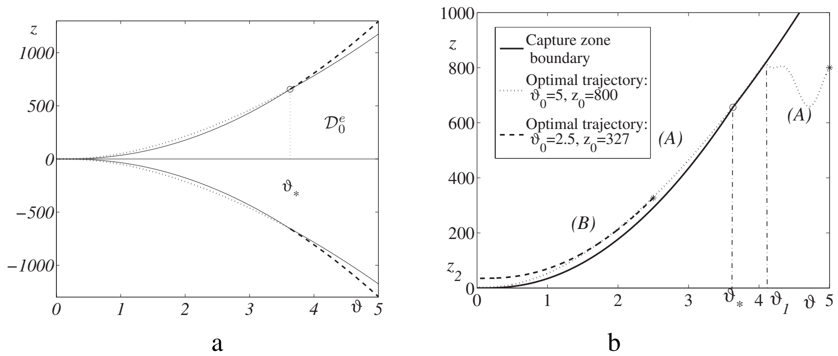


Figure 3: Example of the (2,1)-game solution.

evader strategy $V^*(\cdot)$. This feature reflects the dependence of $V^*(\cdot)$ on the initial position (ϑ_0, z_0).

Remark 3.5. If $(\vartheta_0, z_0) \notin \mathcal{D}_0^e$ and the minimum in (44) is achieved for more than one number $i^* \in I_e$, then an optimal extended evader strategy $V^*(\cdot)$ is not unique.

Remark 3.6. Due to (57), the set $\mathcal{D}_0^e$ is the maximal capture zone in the $(N_e, 1)$-game.

In Figure 3a an example of the capture zone is shown for the (2,1)-game with $\tau_p = 0.2$ s, $a_p^{\max} = 200$ m/s^2, $\tau_{e1} = 0.12$ s, $a_{e1}^{\max} = 85$ m/s^2, $\tau_{e2} = 1$ s, $a_{e2}^{\max} = 133$ m/s^2.

In this example, the boundaries of $D_0(\omega_e^1, \omega_p)$ and $D_0(\omega_e^2, \omega_p)$ (dashed and dotted lines, respectively) intersect at $\Theta = \Theta_* \approx 3.63$ s. The boundary of the capture zone $\mathcal{D}_0^e$ (solid line) coincides with the boundary of $D_0(\omega_e^1, \omega_p)$ for $0 \leq \Theta \leq \Theta_*$ and with the boundary of $D_0(\omega_e^2, \omega_p)$ for $\Theta \geq \Theta_*$. In Figure 3b two optimal trajectories of this game are shown. In both cases $v^* = \sin 10\Theta$ for $(\vartheta, z) \in D_0(\omega_e^{i^*}, \omega_p)$, $u^* = \cos\Theta$ for $(\vartheta, z) \in \mathcal{D}_0^e$. The first trajectory (A) (dotted line) starts at $(\vartheta_0^1, z_0^1) \in \mathcal{D}_0^e$. Due to Theorem 3.1, i^* is arbitrary. In this example, $i^* = 2$. It is seen that at $\vartheta_1 > \vartheta_*$ (A) reaches the boundary of $\mathcal{D}_0^e$, i.e., the boundary of $D_0(\omega_e^2, \omega_p)$. For $\vartheta \leq \vartheta_1$ the system is controlled by $u^* = v^* \equiv 1$. Therefore, the rest of (A) goes alongside the boundary of $D_0(\omega_e^2, \omega_p)$. Thus, for $0 \leq \vartheta \leq \vartheta_*$ this trajectory lies outside $\mathcal{D}_0^e$. However, it ends at the origin. The initial position (ϑ_0^2, z_0^2) of the second trajectory (B) (dashed line) belongs to the first trajectory, such that $\vartheta_0^2 < \vartheta_*$, i.e., $(\vartheta_0^2, z_0^2) \notin \mathcal{D}_0^e$. Due to (44), in this case $i^* = 1$, $v^* = u^* \equiv 1$. Therefore, this trajectory is parallel to the boundary of $D_0(\omega_e^1, \omega_p)$, ending at $(0, z_2)$, where $z_2 > 0$. This example shows that the Bellman Principle cannot be extended to the $(N_e, 1)$-game: the optimal trajectory (B), beginning at the other optimal trajectory (A), does not belong to (A).

3.2 Solution of the Scalar (1, N_p)-Game

In this game $\Omega_e = \{\omega_e\}$.

Remark 3.7. In the $(1, N_p)$-game, the choice of the vector $\omega_p \in \Omega_p$ is equivalent to the choice of the number $j \in I_p \triangleq \{1, \ldots, N_p\}$. Since the pursuer chooses $\omega_p^j \in \Omega_p$ a finite number of times during the game, the dynamics (18) has a variable structure. However, between two successive switch points this dynamics is completely determined by the chosen pair. Therefore, the pair (ω_e, ω_p^j) can be called the *mode* of the dynamics of this game. Due to the information structure of the (N_e, N_p)-game (see Section 2.3), the evader knows the set Ω_p but the specific vector $\omega_p^j \in \Omega_p$, chosen by the pursuer at the current time-to-go instant, is not known to the evader. The pursuer has the perfect information on the mode.

Remark 3.8. In this game, the extended evader strategy is reduced to the regular one $v(\vartheta, z) \in \mathcal{F}$. Since the choice of $\omega_p \in \Omega_p$ is equivalent to the choice of the mode number $j \in I_p$, the pursuer extended strategy can be rewritten as $U(\cdot) = (j(\vartheta, z), u(\vartheta, z)) \in \mathcal{I}_p \times \mathcal{F}$, where $\mathcal{I}_p$ is the set of functions, mapping $[0, \vartheta_0] \times R$ onto I_p, such that for any ϑ they are piecewise constant in z and for any z they are piecewise constant in ϑ. In the sequel, the function $j(\cdot) \in \mathcal{I}_p$ is called a pursuer *schedule*.

Due to Remarks 3.7 and 3.8, Definitions 1, 2 can be rewritten in the following form.

Definition 1_p. The extended strategy $U^*(\cdot) \in \mathcal{I}_p \times \mathcal{F}$ is called the *optimal extended pursuer strategy* in the $(1, N_p)$-game, if for any $U(\cdot) \in \mathcal{I}_p \times \mathcal{F}$,

$$\max_{v(\cdot)\in\mathcal{F}} J(U^*(\cdot), v(\cdot)) \le \max_{v(\cdot)\in\mathcal{F}} J(U(\cdot), v(\cdot)). \tag{62}$$

Definition 2_p. The strategy $v^*(\cdot) \in \mathcal{F}$ is called the *optimal evader strategy* in the $(1, N_p)$-game, if for any $v(\cdot) \in \mathcal{F}$,

$$\min_{U(\cdot)\in\mathcal{I}_p\times\mathcal{F}} J(U(\cdot), v^*(\cdot)) \ge \min_{U(\cdot)\in\mathcal{I}_p\times\mathcal{F}} J(U(\cdot), v(\cdot)). \tag{63}$$

The optimal strategies in the $(1, N_p)$-game are designed using the set

$$\mathcal{D}_0^p = \{(\vartheta, z) : \vartheta > 0, \ |z| < Z_p(\vartheta)\}, \tag{64}$$

where

$$Z_p(\vartheta) = \int_0^{\vartheta} H^*(\xi) d\xi, \tag{65}$$

$$H^*(\xi) \triangleq \max_{j\in I_p} H(\xi, \omega_e, \omega_p^j) \ \text{ for each } \xi > 0. \tag{66}$$

Due to (31),

$$H^*(\xi) = \max_{j \in I_p} \left[h(\xi, \tau_{pj}, a_{pj}^{\max}) \right] - h(\xi, \tau_e, a_e^{\max}), \tag{67}$$

where $h(\vartheta, \tau, a^{\max})$ is given by (19).

From the proof of Lemma A.1 (see the Appendix), it can be seen that the curves $z = H(\vartheta, \omega_e, \omega_p^j)$, $j \in I_p$, intersect each other a finite number of times. This means that $j(\vartheta) = \arg\max_{j \in I_p} H(\vartheta, \omega_e, \omega_p^j)$ changes a finite number of times for $\vartheta > 0$. Therefore, $Z_p(\vartheta)$ is a continuously differentiable function. Moreover, there exists the unique finite set $\Theta_p = \{\vartheta_k \,|k = 1, \ldots, K_p\}$ such that

- $0 < \vartheta_1 < \vartheta_2 < \cdots < \vartheta_{K_p}$;
- for each interval $(\vartheta_k, \vartheta_{k+1})$, $k = 0, \ldots, K_p$, $\vartheta_0 = 0$, $\vartheta_{K_p+1} = \infty$, there exists a number $j_k \in I_p$ such that

$$H^*(\vartheta) = H(\vartheta, \omega_e, \omega_p^{j_k}), \quad \vartheta \in (\vartheta_k, \vartheta_{k+1}), \tag{68}$$

and

$$j_k \neq j_{k+1}, \quad k = 0, \ldots, K_p. \tag{69}$$

Remark 3.9. Due to (29), (30) and (65), the boundary $\partial\mathcal{D}_0^p$ of $\mathcal{D}_0^p$ is parallel to the boundary of $D_0(\omega_e, \omega_p^{j_k})$ for $\vartheta \in (\vartheta_k, \vartheta_{k+1})$, $k = 0, \ldots, K_p$.

Theorem 3.2. *The optimal extended pursuer strategy $U^*(\cdot) \in \mathcal{I}_p \times \mathcal{F}$ is given by*

$$j^*(\vartheta, z) = j_k, \vartheta \in (\vartheta_k, \vartheta_{k+1}], \; k = 0, \ldots, K_p, \tag{70}$$

$$u^*(\vartheta, z) = \begin{cases} \text{arbitrary}, & (\vartheta, z) \in \mathcal{D}_0^p, \\ \text{sign}(z), & (\vartheta, z) \notin \mathcal{D}_0^p. \end{cases} \tag{71}$$

The optimal evader strategy $v^*(\cdot) \in \mathcal{F}$ is

$$v^*(\vartheta, z) = \begin{cases} \text{arbitrary}, & (\vartheta, z) \in \mathcal{D}_0^p, \\ \text{sign}(z), & (\vartheta, z) \notin \mathcal{D}_0^p. \end{cases} \tag{72}$$

Proof. Pursuer strategy. Note that if $U(\cdot) = U^*(\cdot)$ and $v(\cdot) \in \mathcal{F}$ is arbitrary, then

$$\left| h(\vartheta, \tau_{p,j_k}, a_{p,j_k}^{\max}) u^*(\vartheta, z) - h_2(\vartheta, \tau_e, a_e^{\max}) v \right| \geq H(\vartheta, \omega_e, \omega_p^{j_k}), \tag{73}$$

where $(\vartheta, z) \in \partial\mathcal{D}_0^p \cap (\vartheta_k, \vartheta_{k+1})$, $k = 0, \ldots, K_p$. If $(\vartheta_0, z_0) \in \mathcal{D}_0^p$, then, due to (65), the trajectory of (18) cannot leave $\mathcal{D}_0^p \cup \partial\mathcal{D}_0^p$ for $\vartheta < \vartheta_0$. This means that

$$J(U^*(\cdot), v(\cdot)) = 0, \quad (\vartheta_0, z_0) \in \mathcal{D}_0^p, \tag{74}$$

yielding the inequality (62).

Now, let $(\vartheta_0, z_0) \notin \mathcal{D}_0^p$. Solving (18) for any $U(\cdot) = \{j(\cdot), u(\cdot)\} \in \mathcal{I}_p \times \mathcal{F}$ and $v(\cdot) \in \mathcal{F}$ yields

$$z(0; U(\cdot), v(\cdot)) = z_0 - \int_0^{\vartheta_0} \big[\tilde{h}(\vartheta, j(\cdot))u(\vartheta) - h(\vartheta, \tau_e, a_e^{\max})v(\vartheta)\big]d\vartheta, \qquad (75)$$

where

$$\tilde{h}(\vartheta, j(\cdot)) = h\big(\vartheta, \tau_{p,j(\vartheta,z(\vartheta))}, a_{p,j(\vartheta,z(\vartheta))}^{\max}\big). \qquad (76)$$

In the sequel, for the sake of definiteness, it is supposed that $z_0 > 0$. First, consider the left-hand side of (62). Due to (75) and (76), the solution of (18) for $U(\cdot) = U^*(\cdot)$ yields

$$z(0; U^*(\cdot), v(\cdot)) = z_0 - \int_0^{\vartheta_0} \tilde{h}(\vartheta, j^*(\cdot))u^*(\vartheta)d\tau + \int_0^{\vartheta_0} h(\vartheta, \tau_e, a_e^{\max})v(\vartheta)d\vartheta, \quad (77)$$

and

$$\tilde{h}(\vartheta, j^*(\cdot)) = h\big(\vartheta, \tau_{p,j_k}, a_{p,j_k}^{\max}\big), \vartheta \in (\vartheta_k, \vartheta_{k+1}],\ k = 0, \dots, K_p. \qquad (78)$$

Let us show that $z(0) \geq 0$. Indeed, if $z(0) < 0$, then the trajectory of (18) necessarily enters $\mathcal{D}_0^p$. Hence, due to (74), $z(0) = 0$. This contradiction proves that $z(0) \geq 0$. The latter and (77) directly yield

$$J(U^*(\cdot), v(\cdot)) = z(0; U^*(\cdot), v(\cdot)), \qquad (79)$$

and the maximum of $J(U^*(\cdot), v(\cdot))$ is achieved for $v(\vartheta) \equiv 1$. Moreover, for this $v(\cdot)$ the trajectory of (18) does not enter $\mathcal{D}_0^p$ and, consequently, $u^*(\vartheta) \equiv 1$. Therefore,

$$\begin{aligned}\max_{v(\cdot)\in\mathcal{F}} J(U^*(\cdot), v(\cdot)) = z(0; U^*(\cdot), 1) &= z_0 - \int_0^{\vartheta_0} \big[\tilde{h}(\vartheta, j^*(\cdot)) - h(\vartheta, \tau_e, a_e^{\max})\big]d\tau \\ &= z_0 - \int_0^{\vartheta_0} H^*(\vartheta)d\tau. \end{aligned} \qquad (80)$$

Due to (66), (68), for any $U(\cdot) \in \mathcal{I}_p \times \mathcal{F}$ and for all $\vartheta > 0,\ z \in R^1$,

$$H^*(\vartheta) = \tilde{h}(\vartheta, j^*(\cdot)) - h(\vartheta, \tau_e, a_e^{\max}) \geq \tilde{h}(\vartheta, j(\cdot))u(\vartheta) - h(\vartheta, \tau_e, a_e^{\max}). \quad (81)$$

Then

$$z(0; U(\cdot), 1) \geq z(0; U^*(\cdot), 1) \geq 0. \qquad (82)$$

The equation (80) and the inequality (82) immediately imply (62). The case $z_0 < 0$ is treated similarly. Thus, the optimality of $U^*(\cdot)$ is proven.

Evader strategy. If $(\vartheta_0, z_0) \in \mathcal{D}_0^p$, then, due to (74),

$$\min_{U(\cdot)\in\mathcal{I}_p\times\mathcal{F}} J(U(\cdot), v(\cdot)) = 0, \quad \forall\, v(\cdot) \in \mathcal{F}, \tag{83}$$

yielding (63).

Now, let $(\vartheta_0, z_0) \notin \mathcal{D}_0^p$ and, for the sake of definiteness, $z_0 > 0$. Due to (82),

$$\min_{U(\cdot)\in\mathcal{I}_p\times\mathcal{F}} J(U(\cdot), v^*(\cdot)) = J(U^*(\cdot), v^*(\cdot)) = z(0; U^*(\cdot), 1). \tag{84}$$

Due to (77) and (79),

$$z(0; U^*(\cdot), 1) \geq z(0; U^*(\cdot), v(\cdot)) = J(U^*(\cdot), v(\cdot)), \quad \forall\, v(\cdot) \in \mathcal{F}. \tag{85}$$

The equation (84) and the inequality (85) immediately yield (63). The case $z_0 < 0$ is treated similarly. Thus, the optimality of $v^*(\cdot)$ is proven, which completes the proof of the theorem. □

Remark 3.10. It is important to note that, due to (67)–(70), the optimal pursuer schedule $j^*(\cdot)$ is independent of z and ω_e, i.e., it is equally valid for any initial position and against any possible evader. However, the set $\mathcal{D}_0^p$ depends on ω_e.

Corollary 3.2. *For any fixed initial position (ϑ_0, z_0),*

$$J(U^*(\cdot), v^*(\cdot)) = \begin{cases} 0, & (\vartheta_0, z_0) \in \mathcal{D}_0^p, \\ |z_0| - \displaystyle\int_0^{\vartheta_0} H^*(\vartheta)d\vartheta, & (\vartheta_0, z_0) \notin \mathcal{D}_0^p, \end{cases} \tag{86}$$

where $H^*(\vartheta)$ is given by (66). Moreover, for any $U(\cdot) \in \mathcal{I}_p \times \mathcal{F}$, $v(\cdot) \in \mathcal{F}$, the saddle point inequality is satisfied:

$$J(U^*(\cdot), v(\cdot)) \leq J(U^*(\cdot), v^*(\cdot)) \leq J(U(\cdot), v^*(\cdot)), \tag{87}$$

i.e., $(U^*(\cdot), v^*(\cdot))$ is a saddle point in the $(1, N_p)$-game and (86) is the game value.

Proof. If $(\vartheta_0, z_0) \in \mathcal{D}_0^p$, then (86) and (87) are direct consequences of (74).

Now, let $(\vartheta_0, z_0) \notin \mathcal{D}_0^p$ and, for the sake of definiteness, $z_0 > 0$. In this case (86) follows immediately from (80) and (84). The right-hand and the left-hand sides of (87) follow from (82) and (85), respectively. The case of $z_0 < 0$ is treated similarly. This completes the proof of the corollary. □

Remark 3.11. Due to (86), the set $\mathcal{D}_0^p$ is the maximal capture zone in the $(1, N_p)$-game.

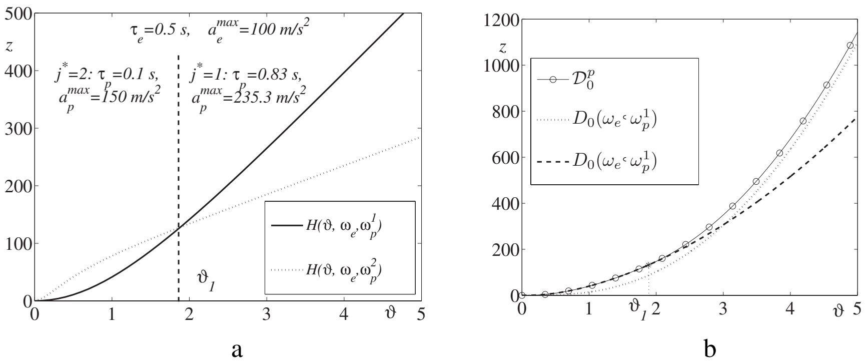


Figure 4: Construction of the (1,2)-game capture zone.

The construction of $\mathcal{D}_0^p$ is illustrated in Figure 4 for the (1,2)-game with $\tau_e = 0.5$ s, $a_e^{\max} = 100$ m/s^2, $\tau_{p1} = 0.83$ s, $a_{p1}^{\max} = 235.3$ m/s^2, $\tau_{p2} = 0.1$ s, $a_{p2}^{\max} = 150$ m/s^2. In Figure 4a the respective graphs of $H(\vartheta, \omega_e, \omega_p)$ are shown.

It is seen that these graphs intersect at the single point $\vartheta = \vartheta_1 \approx 1.88$ s. Thus, in this example, the function H^*, defined by (66), is

$$H^*(\vartheta) = \begin{cases} H(\vartheta, (100, 0.5), (150, 0.1)), & 0 \le \vartheta \le \vartheta_1, \\ H(\vartheta, (100, 0.5), (235.3, 0.83)), & \vartheta > \vartheta_1. \end{cases} \tag{88}$$

Consequently, $K_p = 1$ and $\Theta_p = \{\vartheta_1\}$. Thus,

$$j^*(\vartheta, z) = \begin{cases} 2, & 0 \le \vartheta \le \vartheta_1, \\ 1, & \vartheta > \vartheta_1. \end{cases} \tag{89}$$

Due to (65), (88), the boundary of $\mathcal{D}_0^p$ is

$$Z_p(\vartheta) = \begin{cases} \displaystyle\int_0^{\vartheta} H(\xi, \omega_e, \omega_p^2)d\xi, & 0 \le \vartheta \le \vartheta_1, \\ \displaystyle\int_0^{\vartheta_1} H(\xi, \omega_e, \omega_p^2)d\xi + \int_{\vartheta_1}^{\vartheta} H(\xi, \omega_e, \omega_p^1)d\xi, & \vartheta > \vartheta_1. \end{cases} \tag{90}$$

The upper part of the respective capture zone $\mathcal{D}_0^p$ is shown in Figure 4b.

Note that the set $\mathcal{D}_0^p$ is wider than both capture zones $D_0(\omega_e, \omega_p^1)$ and $D_0(\omega_e, \omega_p^2)$ in the (1,1)-games. This means that an extended strategy $U(\cdot)$ in the $(1, N_p)$-game guarantees the capture from a wider set of the initial positions than the optimal strategy of each (1,1)-game. For example, let $(\vartheta_0, z_0) = (5, 1120)$. This point does not belong to $D_0(\omega_e, \omega_p^1)$ and $D_0(\omega_e, \omega_p^2)$, while it belongs to $\mathcal{D}_0^p$. The game values in the (1,1)-games, starting at this point, are $J^*(5, 1120, \omega_e, \omega_p^1) = 21.22$ m and $J^*(5, 1120, \omega_e, \omega_p^2) = 343.5$ m, while the game value of the (1,2)-game is zero.

3.3 Solution of the Scalar (N_e, N_p)-Game

In this game, the uncertainty arises both in the model and in the mode.

Remark 3.12. In the (N_e,N_p)-game, the choice of the vectors $\omega_e \in \Omega_e$ and $\omega_p \in \Omega_p$ is equivalent to the choice of the indexes $(i, j) \in I_e \times I_p$. Based on Remarks 3.2 and 3.7, the extended evader and pursuer strategies can be rewritten as $V(\cdot) = \Big(i(\vartheta_0, z_0), v(\vartheta, z)\Big) \in I_e \times \mathcal{F}$ and $U(\cdot) = \Big(j(\vartheta, z), u(\vartheta, z)\Big) \in \mathcal{I}_p \times \mathcal{F}$, respectively.

In this game, Definitions 1, 2 can be rewritten in the following form.

Definition 1_{ep}. The extended strategy $U^*(\cdot) \in \mathcal{I}_p \times \mathcal{F}$ is called the *optimal extended pursuer strategy* in the (N_e, N_p)-game, if for any $U(\cdot) \in \mathcal{I}_p \times \mathcal{F}$,

$$\max_{V(\cdot)\in I_e\times\mathcal{F}} J(U^*(\cdot), V(\cdot)) \leq \max_{V(\cdot)\in I_e\times\mathcal{F}} J(U(\cdot), V(\cdot)). \tag{91}$$

Definition 2_{ep}. The extended strategy $V^*(\cdot) \in I_e \times \mathcal{F}$ is called the *optimal extended evader strategy* in the (N_e, N_p)-game, if for any $V(\cdot) \in I_e \times \mathcal{F}$,

$$\min_{U(\cdot)\in\mathcal{I}_p\times\mathcal{F}} J(U(\cdot), V^*(\cdot)) \geq \min_{U(\cdot)\in\mathcal{I}_p\times\mathcal{F}} J(U(\cdot), V(\cdot)). \tag{92}$$

The study of the (N_e, N_p)-game is carried out similarly to the study of the $(N_e, 1)$-game. Namely, the (N_e, N_p)-game is partitioned into N_e separate $(1, N_p)$-games, denoted in the sequel as $(1, N_p)_i$-games, $i = 1, \dots, N_e$. For each $(1, N_p)_i$-game, (64) becomes

$$\mathcal{D}_0^p = \mathcal{D}_0^{pi} = \{(\vartheta, z) :\ \vartheta > 0,\ |z| < Z_{pi}(\vartheta)\},\ i = 1, \dots, N_e, \tag{93}$$

where

$$\begin{aligned} Z_{pi}(\vartheta) &= \int_0^{\vartheta} \max_{j\in I_p} H(\xi, \omega_e^i, \omega_p^j) d\xi \\ &= \int_0^{\vartheta} \max_{j\in I_p} h(\xi, \tau_{pj}, a_{pj}^{\max}) d\xi - \int_0^{\vartheta} h(\xi, \tau_{ei}, a_{ei}^{\max}) d\xi. \end{aligned} \tag{94}$$

Let $U^{*i}(\cdot) = \{j^{*i}(\cdot), u^{*i}(\cdot)\}$, $v^{*i}(\cdot)$ and J^{*i} be the optimal extended pursuer strategy, the optimal evader strategy and the game value for each $(1, N_p)_i$-game. Note that the triplet $\{U^{*i}(\cdot), v^{*i}(\cdot), J^{*i}\}$ is obtained by using (70)–(72) and (86) for $\mathcal{D}_0^p = \mathcal{D}_0^{pi}$.

Due to Remark 3.10, the optimal pursuer schedules $j^{*i}(\cdot)$ are equal to each other, i.e.,

$$j^{*1}(\cdot) = j^{*2}(\cdot) = \cdots = j^{*N_e}(\cdot) = j^*(\cdot). \tag{95}$$

Let

$$\mathcal{D}_0^{ep} \triangleq \bigcap_{i=1}^{N_e} \mathcal{D}_0^{pi}. \tag{96}$$

Due to (39), (65) and (95), this set can be described analytically as follows:

$$\mathcal{D}_0^{ep} = \{(\vartheta, z) :\ \vartheta > 0, |z| < Z_{ep}(\vartheta)\}, \tag{97}$$

where for each $\vartheta > 0$,

$$Z_{ep}(\vartheta) = \min_{i \in I_e} Z_{pi}(\vartheta). \tag{98}$$

Theorem 3.3. *The optimal extended pursuer strategy is $U^*(\cdot) = (j^*(\cdot), u^*(\cdot)) \in \mathcal{I}_p \times \mathcal{F}$, where $j^*(\cdot)$ is given by (95) and*

$$u^*(\vartheta, z) = \begin{cases} \text{arbitrary}, & (\vartheta, z) \in \mathcal{D}_0^{ep}, \\ \text{sign}(z), & (\vartheta, z) \notin \mathcal{D}_0^{ep}. \end{cases} \tag{99}$$

For any initial position (ϑ_0, z_0) the optimal extended evader strategy $V^*(\cdot) = \{i^*(\cdot), v^*(\cdot)\} \in I_e \times \mathcal{F}$ is given as follows. If $(\vartheta_0, z_0) \in \mathcal{D}_0^{ep}$, then $i^*(\vartheta_0, z_0) \in I_e$ is arbitrary and

$$v^*(\vartheta, z, i^*) = \begin{cases} \text{arbitrary}, & (\vartheta, z) \in \mathcal{D}_0^{pi^*}, \\ \text{sign}(z), & (\vartheta, z) \notin \mathcal{D}_0^{pi^*}. \end{cases} \tag{100}$$

If $(\vartheta_0, z_0) \notin \mathcal{D}_0^{ep}$, then

$$i^*(\vartheta_0, z_0) = \arg\min_{i \in I_e} Z_{pi}(\vartheta_0), \tag{101}$$

and

$$v^*(\vartheta, z, i^*) = \text{sign}(z). \tag{102}$$

Proof. The theorem is proven quite similarly to Theorem 3.1. □

Remark 3.13. The optimal strategies in the $(N_e, 1)$-game and the $(1, N_p)$-game, derived in Sections 3.1 and 3.2, are the particular cases of the optimal strategies in the (N_e, N_p)-game for $N_p = 1$ and $N_e = 1$, respectively.

Corollary 3.3. *For any fixed initial position (ϑ_0, z_0),*

$$J(U^*(\cdot), V^*(\cdot)) = \begin{cases} 0, & (\vartheta_0, z_0) \in \mathcal{D}_0^{ep}, \\ |z_0| - Z_{pi^*}(\vartheta_0), & (\vartheta_0, z_0) \notin \mathcal{D}_0^{ep}, \end{cases} \tag{103}$$

where $i^* = i^*(\vartheta_0, z_0)$ is given by (101). Moreover, for any $U(\cdot) \in \mathcal{I}_p \times \mathcal{F}$, $V(\cdot) \in I_e \times \mathcal{F}$, the saddle point inequality is satisfied:

$$J(U^*(\cdot), V(\cdot)) \le J(U^*(\cdot), V^*(\cdot)) \le J(U(\cdot), V^*(\cdot)), \tag{104}$$

i.e., $(U^*(\cdot), V^*(\cdot))$ is a saddle point in the (N_e, N_p)-game and (103) is the game value.

Proof. The corollary is proven similarly to Corollary 3.1. □

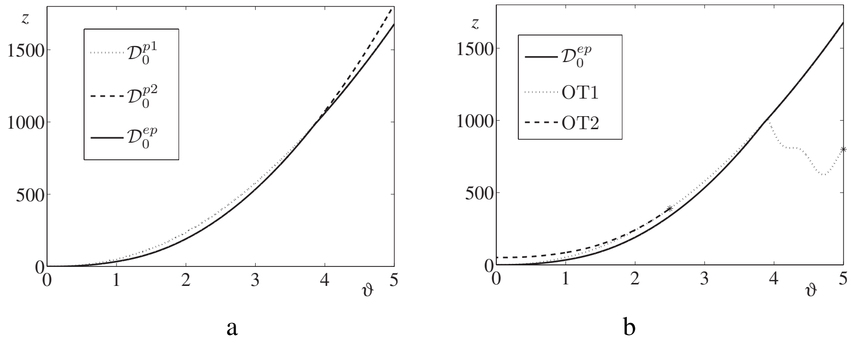


Figure 5: Example of the (2, 2)-game.

Remark 3.14. Due to (103), the set $\mathcal{D}^{ep}$ is the maximal capture zone in the (N_e, N_p)-game.

In Figure 5 an example of the (2, 2)-game, similar to the example of the (2,1)-game in Figure 3, is shown. The game parameters are $\tau_{p1} = 0.2$ s, $\tau_{p2} = 0.6$ s, $a_{p1}^{\max} = 200$ m/s^2, $a_{p2}^{\max} = 300$ m/s^2, $\tau_{e1} = 0.2$ s, $\tau_{e2} = 1.5$ s, $a_{e1}^{\max} = 100$ m/s^2, $a_{e2}^{\max} = 180$ m/s^2. In Figure 5a the upper parts of the capture zone $\mathcal{D}_0^{ep}$ along with the respective sets $\mathcal{D}_0^{p1}$ and $\mathcal{D}_0^{p2}$ are shown.

In Figure 5b two optimal trajectories (denoted as OT1 and OT2) of this game are shown. In both cases $v^* = \sin 10\tau$ for $(\vartheta, z) \in \mathcal{D}_0^{pi^*}$, $u^* = \cos\tau$ for $(\vartheta, z) \in \mathcal{D}_0^{ep}$. In this example, the optimal pursuer schedule is

$$j^*(\vartheta) = \begin{cases} 1, & 0 \le \vartheta \le 1.13, \\ 2, & \vartheta > 1.13. \end{cases} \tag{105}$$

The first trajectory (OT1) starts at $(\vartheta_0^1, z_0^1) \in \mathcal{D}_0^{ep}$. Due to Theorem 3.3, i^* is arbitrary. In this example, $i^* = 2$. This trajectory, similarly to the trajectory (A) in Figure 3b, leaves $\mathcal{D}_0^{ep}$. However, it ends at the origin. The second trajectory (OT2) is similar to the trajectory (B) in Figure 3b. It starts at OT1 outside $\mathcal{D}_0^{ep}$, yielding $i^* = 1$, $v^* = u^* \equiv 1$. Therefore, this trajectory ends at $(0, z_2)$, where $z_2 > 0$. This example shows that the Bellman Principle cannot be extended to the (N_e, N_p)-game either.

4 Application to the Multimodel-Multimode Interception Problem

In this section, the results of Section 3 are applied to the solution of the interception problem with the dynamics (10), control constraints (8), (9) and the sets of parameters Ω_i, $i = e, p$. The objective of the pursuer is to provide a zero miss distance against all admissible evader extended strategies $(i(\cdot), v(\cdot))$ by means of an extended strategy $(j(\cdot), u(\cdot))$. Note that in the case $N_e = N_p = 1$ the solution of the scalar (1,1)-game directly provides the solution of the interception problem. Namely, if $d^T\Phi(t_f, 0; \tau_e, \tau_p)x_0 \in D_0(\omega_e, \omega_p)$, then the strategy $u(t, x) = u^0(t_f - t, d^T\Phi(t_f, t; \tau_e, \tau_p)x)$ solves the interception problem, where $u^0(\vartheta, z)$ is the optimal pursuer strategy in the (1,1)-game.

However, for $N_e > 1$ and/or $N_p > 1$ this relation between the interception problem and the (N_e, N_p)-game is lost. Indeed, in this case,

$$\Phi(t_f, t; \tau_e, \tau_p) = \Phi(t_f, t; \tau_{ei}, \tau_{p,j(t)}), \tag{106}$$

where $i \in I_e$ and

$$j(t) = j_k \in I_p, t \in [t_k, t_{k+1}), k = 1, \dots, K_p. \tag{107}$$

The matrix (106) has the following features:

(1) for each t, it is not single valued, because the index i is unknown to the pursuer;
(2) for a given $i \in I_e$, it is discontinuous at $t = t_k$, causing the jumps of $d^T\Phi\big(t_f, t; \tau_{ei}, \tau_{p,j(t)}\big)x$

$$\Delta z^k = \big(\tau^2_{pj_{k-1}}\Psi((t_f - t_k)/\tau_{pj_{k-1}}) - \tau^2_{pj_k}\Psi((t_f - t_k)/\tau_{pj_k})\big)x_4(t_k). \tag{108}$$

Due to these features of the matrix (106), the pursuer strategy $u(t, x) = u^*(t_f - t, d^T\Phi(t_f, t; \tau_e, \tau_p)x)$, where $u^*(\vartheta, z)$ is the component of the optimal extended pursuer strategy $U^*(\cdot) = (j^*(\cdot), u^*(\cdot))$ in the scalar (N_e, N_p)-game and $\Phi(t_f, t; \tau_e, \tau_p)$ is given by (106), cannot be directly used for the solution of the interception problem.

The following modification of $U^*(\cdot)$ is proposed as an heuristic solution of the interception problem: for any current position $(t, x(t))$

(i) calculate the pursuer index

$$j(t) = j^*(t_f - t); \tag{109}$$

(ii) calculate all possible (from the pursuer viewpoint) zero-effort miss distances for the calculated pursuer index $j(t)$:

$$z^i(t) = d^T\Phi(t_f, t; \tau_{ei}, \tau_{p,j(t)})x(t), i = 1, \dots, N_e; \tag{110}$$

(iii) select the maximal one

$$\tilde{z}(t) = \max_{i \in I_e} |z^i(t)|; \tag{111}$$

(iv) use

$$u(t, x(t)) = \operatorname{sign}(\tilde{z}(t)). \tag{112}$$

Note that

$$\tilde{z}(t_f) = |x_1(t_f)|, \tag{113}$$

i.e., $\tilde{z}(t_f) = 0 \Rightarrow |x_1(t_f)| = 0$.

The pursuer control (112) provides the maximal decreasing rate of $\tilde{z}$ at any instant t. This guarantees the maximal distance between the point $(t_f - t, \tilde{z}(t))$

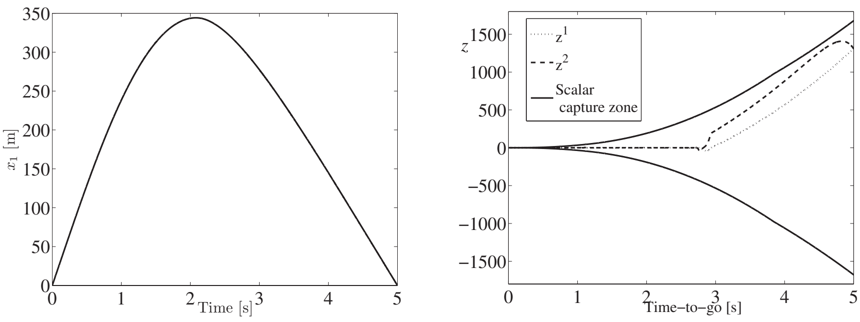
and the boundary of the scalar capture zone $\mathcal{D}_0^{ep}$, thus reducing the influence of the jumps (108) on the result of the interception.

In Figures 6 and 7 an example of the interception is shown for the same parameter sets Ω_e, Ω_p as in Figure 5 and $V_p = 2300$ m/s, $V_e = 2700$ m/s, $\phi_p^0 = -3$ deg, $\phi_e^0 = 3$ deg.

In this example, the evader chooses $i = \hat{i}$ ($\hat{i} = 1$ and $\hat{i} = 2$ in Figure 6 and Figure 7, respectively), while $v = \text{sign}(z^{\hat{i}}(t))$. The optimal pursuer schedule $j^*(\vartheta)$ in the scalar game is given by (105). Due to (105), the schedule $j^*(\vartheta)$ has the single switch point $\vartheta_1 = 1.13$ s, yielding the single switch point $t_1 = 3.87$ s of $j(t) = j^*(t_f - t)$. In Figures 6a and 7a the trajectories $x_1(t)$, generated by (109)–(112), are shown. It is seen that in both cases $x_1(t_f) = 0$. In Figures 6b and 7b the respective z^1 and z^2 are shown as functions of time-to-go. Note that in Figure 6b the true zero-effort miss distance is z^1, while in Figure 7b it is z^2.

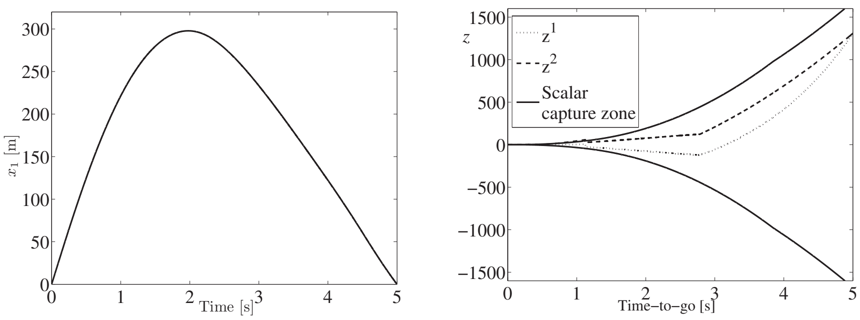
In Figure 8 the jump phenomena (108) are shown for $\hat{i} = 1$.

It is seen that the jump does not push out the true zero-effort miss distance z^1 from the scalar capture zone, while z^2 leaves the capture zone due to this jump.



a b

Figure 6: Original interception problem, $\hat{i} = 1$.



a b

Figure 7: Original interception problem, $\hat{i} = 2$.

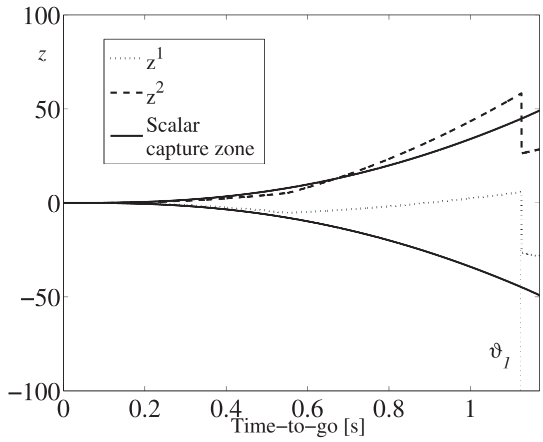


Figure 8: Jump phenomena, $\hat{i} = 1$.

Thus, in this case, the jump does not prevent the pursuer from achieving zero miss distance. Moreover, the alternative zero-effort miss distance z^2, in spite of leaving the capture zone, becomes zero at the final time. In the case $\hat{i} = 2$ both z^1 and z^2 do not leave the scalar capture zone after the jump.

5 Conclusions

In this chapter, a pursuit-evasion differential game with variable structure and uncertain dynamics has been studied. First, the game was reduced to a scalar version. The closed-form solution of this new game is obtained. It was shown that the game has a saddle point. Expressions for the game value and the maximal capture zone of the game were obtained. Based on the solution of the scalar game, a heuristic pursuer feedback control in the zero miss interception problem was proposed. This control was tested by numerical examples.

Appendix Proof of the Lemma

Lemma A.1. *For any given ω_p^1, ω_p^2, ω_e^1, ω_e^2, such that*

$$H(\vartheta, \omega_p^1, \omega_e^1) \not\equiv H(\vartheta, \omega_p^2, \omega_e^2), \ \vartheta > 0, \tag{A.1}$$

the equation

$$\mathcal{G}(\vartheta, \omega_p^1, \omega_e^1, \omega_p^2, \omega_e^2) \stackrel{\Delta}{=} \int_0^{\vartheta} \Big[H(\xi, \omega_p^1, \omega_e^1) - H(\xi, \omega_p^2, \omega_e^2) \Big] d\xi = 0, \tag{A.2}$$

with respect to ϑ, has no more than three different strictly positive roots.

Proof by contradiction. Note that $\vartheta = 0$ is the root of (A.2). Assume that there exist at least four different positive roots of (A.2). Then the derivative with respect to ϑ of the left-hand part of (A.2),

$$\partial \mathcal{G}/\partial \vartheta = H(\vartheta, \omega_p^1, \omega_e^1) - H(\vartheta, \omega_p^2, \omega_e^2) \stackrel{\Delta}{=} H_{12}(\vartheta), \tag{A.3}$$

has at least four different positive roots. Due to (31), $H_{12}(0) = 0$. Hence, the derivative

$$\begin{aligned} H'_{12}(\vartheta) = {} & a_{p1}^{\max}(1 - \exp(-\vartheta/\tau_{p1})) - a_{e1}^{\max}(1 - \exp(-\vartheta/\tau_{e1})) \\ & - a_{p2}^{\max}(1 - \exp(-\vartheta/\tau_{p2})) + a_{e2}^{\max}(1 - \exp(-\vartheta/\tau_{e2})) \end{aligned} \tag{A.4}$$

has at least four different positive roots. Due to (A.4), $H'_{12}(0) = 0$. Hence, the second derivative,

$$\begin{aligned} H''_{12}(\vartheta) = {} & c_{p1} \exp(-\vartheta/\tau_{p1}) - c_{e1} \exp(-\vartheta/\tau_{e1}) \\ & - c_{p2} \exp(-\vartheta/\tau_{p2}) + c_{e2} \exp(-\vartheta/\tau_{e2}), \end{aligned} \tag{A.5}$$

where

$$c_{kl} = \frac{a_{kl}^{\max}}{\tau_{kl}}, \; k = p, e, \quad l = 1, 2, \tag{A.6}$$

has at least four different positive roots.

Multiplying by $\exp(\vartheta/\tau_{p1})$ converts the equation $H''_{12}(\vartheta) = 0$ to

$$\hat{H}_{12}(\vartheta) \stackrel{\Delta}{=} c_{p1} - c_{e1} \exp(\nu_1 \vartheta) - c_{p2} \exp(\nu_2 \vartheta) + c_{e2} \exp(\nu_3 \vartheta) = 0, \tag{A.7}$$

where

$$\nu_1 \stackrel{\Delta}{=} 1/\tau_{p1} - 1/\tau_{e1}, \nu_2 \stackrel{\Delta}{=} 1/\tau_{p1} - 1/\tau_{p2}, \nu_3 \stackrel{\Delta}{=} 1/\tau_{p1} - 1/\tau_{e2}. \tag{A.8}$$

Since $\hat{H}_{12}(\vartheta)$ has at least four different positive roots, the derivative

$$\hat{H}'_{12}(\vartheta) = -\nu_1 c_{e1} \exp(\nu_1 \vartheta) - \nu_2 c_{p2} \exp(\nu_2 \vartheta) + \nu_3 c_{e2} \exp(\nu_3 \vartheta) \tag{A.9}$$

has at least three different positive roots.

Multiplying by $\exp(-\nu_1 \vartheta)$ converts the equation $\hat{H}'_{12}(\vartheta) = 0$ to the following equation:

$$\tilde{H}_{12}(\vartheta) \stackrel{\Delta}{=} -\nu_1 c_{e1} - \nu_2 c_{p2} \exp((\nu_2 - \nu_1)\vartheta) + \nu_3 c_{e2} \exp((\nu_3 - \nu_1)\vartheta) = 0, \tag{A.10}$$

having at least three different positive roots. Hence, the derivative

$$\tilde{H}'_{12}(\vartheta) = -\nu_2 c_{p2}(\nu_2 - \nu_1) \exp((\nu_2 - \nu_1)\vartheta) + \nu_3 c_{e2}(\nu_3 - \nu_1) \exp((\nu_3 - \nu_1)\vartheta) \tag{A.11}$$

has at least two different positive roots.

The following cases can be distinguished.

Case 1. $\nu_2(\nu_2 - \nu_1) \neq 0$ or $\nu_3(\nu_3 - \nu_1) \neq 0$. In this case, by direct solution of the equation $\tilde{H}'_{12}(\vartheta) = 0$, it can be shown that (A.11) has no more than one positive root, which contradicts with the conclusion that (A.11) has at least two different positive roots.

Case 2. $\nu_2(\nu_2 - \nu_1) = \nu_3(\nu_3 - \nu_1) = 0$. This case can be partitioned into the following subcases.

Subcase 2.1. $\nu_1 = \nu_2 = \nu_3 \neq 0$. Due to (A.8), this means that $\tau_{e1} = \tau_{p2} = \tau_{e2}$. The latter along with (A.6) and (A.10) yields

$$-a_{e1}^{\max} - a_{p2}^{\max} + a_{e2}^{\max} = 0. \tag{A.12}$$

Since $a_{p2}^{\max} > a_{e2}^{\max}$, (A.12) is contradiction.

Subcase 2.2. $\nu_1 = \nu_2 = \nu_3 = 0$. In this case, due to (A.8),

$$\tau_{p1} = \tau_{e1} = \tau_{p2} = \tau_{e2}. \tag{A.13}$$

Since (A.4) has at least four different positive roots, (A.13) leads to

$$a_{p1}^{\max} - a_{e1}^{\max} = a_{p2}^{\max} - a_{e2}^{\max}. \tag{A.14}$$

Equations (A.13) and (A.14) yield $H(\vartheta, \omega_p^1, \omega_e^1) \equiv H(\vartheta, \omega_p^2, \omega_e^2)$, which contradicts (A.1).

Subcase 2.3. $\nu_1 = \nu_2 \neq 0$, $\nu_3 = 0$. In this case, $\tau_{e1} = \tau_{p2}$, which, by using (A.6) and (A.10), yields the contradiction $-a_{e1}^{\max} - a_{p2}^{\max} = 0$.

Subcase 2.4. $\nu_1 = \nu_3 \neq 0$, $\nu_2 = 0$. In this case, similarly to Subcase 2.3,

$$\tau_{e1} = \tau_{e2},\ a_{e1}^{\max} = a_{e2}^{\max}. \tag{A.15}$$

Moreover, since $\nu_2 = 0$,

$$\tau_{p1} = \tau_{p2}. \tag{A.16}$$

Using (A.4), (A.15) and (A.16) leads to the same contradiction as in Subcase 2.2.

Subcase 2.5. $\nu_1 \neq 0$, $\nu_2 = \nu_3 = 0$. By using (A.6) and (A.10), this case leads to the contradiction $a_{e1}^{\max} = 0$.

This completes the proof of the lemma. □

References

[1] Bertsekas, D.P. and Rhodes, I.B. (1973), Sufficiently Informative Functions and the Minimax Feedback Control of Uncertain Dynamic Systems, *IEEE Transactions on Automatic Control*, Vol. 18, pp. 117–124.

[2] Corless, M.J. and Leitmann, G. (1981), Continuous State Feedback Guaranteeing Uniform Ultimate Boundedness for Uncertain Dynamic Systems, *IEEE Transactions on Automatic Control*, Vol. 26, pp. 1139–1144.

[3] Lee, Y.I. and Kouvaritakis, B. (2000), Robust Receding Horizon Predictive Control for Systems With Uncertain Dynamics and Input Saturation, *Automatica*, Vol. 36, pp. 1497–1504.

[4] Utkin, V.I. (1983), Variable Structure Systems: Present and Future, *Automation and Remote Control*, Vol. 44, pp. 1105–1120.

[5] Bartolini, G. and Zolezzi, T. (1986), Control of Nonlinear Variable Structure Systems, *Journal of Mathematical Analysis and Applications*, Vol. 118, pp. 42–62.

[6] Chen, X. and Fukuda, T. (1997), Variable Structure System Theory Based on Disturbance Identification and its Applications, *International Journal of Control*, Vol. 68, pp. 373–384.

[7] Choi, H.H. (2004), An Analysis and Design Method for Uncertain Variable Structure Systems with Bounded Controllers, *IEEE Transactions on Automatic Control*, Vol. 49, pp. 602–607.

[8] Amato, F., Mattei, M. and Pironti, A. (1998), Robust Strategies for Nash Linear Quadratic Games Under Uncertain Dynamics, *Proceedings of the 37th IEEE Conference on Decision and Control*, Tampa, Florida, December 1998, pp. 1869–1870.

[9] Poznyak, A.S. and Gallegos, C.J. (2003), Multimodel Prey-Predator LQ Differential Games, *Proceedings of the American Control Conference*, Denver, Colorado, June 4–6, 2003, pp. 5369–5374.

[10] Boltyanski, V.G., and Poznyak, A.S. (1999), Robust Maximum Principle in Minimax Control, *International Journal of Control*, Vol. 72, pp. 305–314.

[11] Boltyanski, V.G., and Poznyak, A.S. (2002), Linear Multi-model Time Optimization, *Optimal Control Applications and Methods*, Vol. 23, pp. 141–161.

[12] Poznyak, A., Duncan, T., Pasik-Duncan, B. and Boltyanski, V. (2002), Robust Maximum Principle for Multimodel LQ Problem, *International Journal of Control*, Vol. 75, pp. 1770–1777.

[13] Grigorenko, N.L. (1985), Differential Games of Pursuit by Several Objects with a Variable Structure, *Some problems in modern mathematics and their applications to problems in mathematical physics*, Moscow Fiz.-Tekhn.Inst., Moscow, pp. 54–61 (in Russian).

[14] Grigorenko, N.L. (1991), Game-Theoretic Problems of Control with Variable Structure, *Moscow University Computational Mathematics and Cybernetics*, No. 4, pp. 4–14.

[15] Grigorenko, N.L. (1998), Dynamic Game of Many Players with Variable Structure, *Proceedings of International Conference Dedicated to the 90th Anniversary of L.S. Pontryagin*, Moscow, August 31–September 6, 1998, pp. 74–77.

[16] Gutman, S. (1979), On Optimal Guidance for Homing Missiles, *Journal of Guidance and Control*, Vol. 3, pp. 296–300.

[17] Shinar, J. (1981), Solution Technics for Realistic Pursuit-Evasion Games, Advances in Control and Dynamic Systems, Edited by C.T. Leondes, Academic Press, New York, NY, Vol. 17, pp. 63–124.

[18] Shima, T. and Shinar, J. (2002), Time-Varying Linear Pursuit-Evasion Game Models with Bounded Controls, *Journal of Guidance, Control and Dynamics*, Vol. 25, pp. 425–432.

[19] Krasovskii, N.N. and Subbotin, A.I. (1988), Game-Theoretical Control Problems, Springer-Verlag, New York, NY.

[20] Bryson, A.E. and Ho, Y.C. (1975), Applied Optimal Control, Hemisphere, New York, NY.

Pursuit-Evasion Games with Impulsive Dynamics

Eva Crück
Laboratoire de Mathématiques
Unité CNRS UMR6205
Université de Bretagne Occidentale
6 avenue Le Gorgeu
29200 Brest, France
cruck@control.ee.ethz.ch

Marc Quincampoix
Laboratoire de Mathématiques
Unité CNRS UMR6205
Université de Bretagne Occidentale
6 avenue Le Gorgeu
29200 Brest, France
marc.quincampoix@univ-brest.fr

Patrick Saint-Pierre
Centre de Recherche Viabilité, Jeux, Contrôle
Université Paris IX-Dauphine
Place du Maréchal de Lattre de Tassigny
F-75775 Paris Cedex 16, France
saint-pierre@viab.dauphine.fr

Abstract

In this chapter, we investigate a two-player zero-sum game with separated impulsive dynamics. We study both qualitative and quantitative games. For the qualitative games, we provide a geometrical characterization of the victory domains. For the quantitative games, we characterize the value functions using the Isaacs partial differential inequalities. As a by-product, we obtain a new result of existence of a value for impulsive differential games. The main tool of our approach is the notion of *impulse discriminating domain*, which is introduced and discussed extensively here.

1 Introduction

We consider a two-player differential game with separated impulsive dynamics. The evolution of the system controlled by the first player—named Ursula—is governed by a pair of equations, a differential equation describing continuous evolution and a discrete equation describing jumps:

$$\begin{cases} y' &= g(y, u) \\ y^+ &= p(y^-, \mu). \end{cases} \tag{1}$$

where u and μ are respectively a continuous and a discrete control chosen by Ursula. The solution of (1) is a discontinuous trajectory $t \mapsto y(t)$ which will be described precisely below.

Similarly, the second player—named Victor—using controls υ and ν, controls a system

$$\begin{cases} z' &= h(z, \upsilon) \\ z^+ &= q(z^-, \nu). \end{cases} \tag{2}$$

The outcome of the game is defined in the following way:

- Ursula has to drive the state $x := (y, z)$ into a set Ω, while keeping it outside a set $\mathcal{T}$.
- Victor has to drive the state in $\mathcal{T}$ while keeping it outside Ω.

We associate with a trajectory $(y(\cdot), z(\cdot))$ of (1)–(2) a payoff

$$\theta_{\mathcal{T}}^{K}(y(\cdot), z(\cdot)) := \inf\{t : (y(t), z(t)) \in \mathcal{T} \quad \text{and} \quad \forall s \le t, \quad (y(s), z(s)) \in K\},$$

where $K := \mathbb{R}^n \setminus \Omega$. We shall investigate the quantitative game in which Victor tries to minimize this payoff and Ursula to maximize it. In the sequel, we assume that K and $\mathcal{T}$ are closed sets.

This kind of game, but without jumps and with $K = \mathbb{R}^n$, has been extensively studied since Isaacs [11] under the name *pursuit-evasion games*. The existence of a value has been proved in the context of nonanticipative strategies introduced by Varaiya [14,15], Roxin [13], and Elliot and Kalton [9]. We refer to [5,6] for the case with state constraints ($K \neq \mathbb{R}^n$).

For the impulsive control case, viewed as a one-player game, we refer to [2], and to [8] for a recent approach allowing one to deal with state constraints.

In the present case of an impulse differential game, we have the difficulty of providing a suitable definition of strategies. Indeed, the actions of the players have to take into account not only the current state of the game, as would feedback strategies, but also the possible past jumps. We propose a definition of strategies which retrieves most of the characteristics of nonanticipative strategies.

Using this context of strategies, we can define a lower value function and an upper value function of the game. We characterize these functions through geometric properties of their epigraph; we also provide an equivalent characterization based on a Hamilton–Jacobi–Isaacs formulation. Then we provide several sufficient conditions for the game to have a value, i.e., for the lower and upper value functions to be equal.

As in [5], we reduce the study of the quantitative game to a qualitative game (*game of kind* in Isaacs' terminology) in which the outcome is "Victor wins" or “Ursula wins.” Therefore, we first investigate qualitative impulse differential games.

The chapter is organized as follows. Section 2 is devoted to the definition of impulsive trajectories. In Section 3 we present the context of strategies and the definition of the value functions. The auxiliary qualitative game is studied in Section 4; the proofs of this section, being quite technical, have been postponed to the Appendix. The characterization of the lower value function is provided in Section 5, as well as necessary conditions for the existence of a value for the pursuit-evasion impulse game.

2 Trajectories of Impulse Dynamical Systems

Let us consider the dynamical system controlled by Ursula. Let us denote by U and M the sets of possible continuous and discrete controls. Moreover, let us assume that Ursula can command jumps only in a prescribed region A_U of her state space $\mathbb{R}^l$. In the sequel, we assume the following.

Assumption 1

- *U and M are compact convex subsets of finite-dimensional spaces;*
- *$g : \mathbb{R}^l \times U \to \mathbb{R}^l$, and $p : A_U \times M \to \mathbb{R}^l$ are Lipschitz continuous with respect to their first variable, and continuous with respect to their second variable;*
- *g has linear growth with respect to the first variable;*
- *A_U is compact and for all $y \in A_U$, $p(y, M) \cap A_U = \emptyset$.*

The last assumption ensures that, after a jump, the trajectory is continuous for some time.

Ursula’s system can be characterized by a pair of set-valued functions (G, P), by changing (1) into

$$\begin{cases} y' \in G(y) := \{g(y, u) : u \in U\} \\ y^+ \in P(y^-) := \{p(y^-, \mu) : \mu \in M\}. \end{cases} \tag{3}$$

The reset map P is defined only on A_U, so that the set from which jumps are allowed coincides with the domain of P, denoted by Dom (P).

The assumptions on g ensure the existence of absolutely continuous solutions defined on $[0, +\infty)$ to the differential inclusion $y' \in G(y)$. Let us denote by $S_G(y_0)$ the set of such solutions. Then we can define the set of trajectories of the impulse dynamical systems using the convenient notion of *run* borrowed from hybrid system literature (see [2] and references therein).

Definition 2.1 (Runs and trajectories). We call a *run* of impulse system (G, P) with initial condition y_0 a finite or infinite sequence $\{\tau_i, y_i, y_i(\cdot)\}_{i\in I}$ of

$(\mathbb{R}^+ \times \mathbb{R}^l \times S_G(\mathbb{R}^l))$ such that for all $i \in I$

$$\begin{cases} y_i'(t) & \in & G(y_i(t)) \\ y_i(0) & = & y_i \end{cases} \quad \text{and} \quad \begin{cases} y_i(\tau_i) & \in & \mathrm{Dom}\,(P) \\ y_{i+1} & \in & P(y_i(\tau_i)). \end{cases}$$

A *trajectory* of impulse system (G, P) is a function $y\colon \mathbb{R} \to \mathbb{R}^l$ associated with a run in the following way:

$$y(t) = \begin{cases} y_0 & \text{if } \ t < 0 \\ y_i(t - t_i) & \text{if } \ t \in [t_i, t_i + \tau_i), \end{cases} \tag{4}$$

where $t_i = \sum_{j<i} \tau_i$. We denote by $S_{G,P}(y_0)$ the set of trajectories with initial condition y_0.

We also assume that the impulse system controlled by Victor satisfies Assumption 1 and can be described similarly by a pair of set-valued maps (H, Q).

3 Context of Strategies and Values of the Game

The aim of this section is to adapt the notion of nonanticipative strategies to differential games with impulsive dynamics. We recall that a map $\beta : \mathcal{U} \longrightarrow \mathcal{V}$, with $\mathcal{U}$ and $\mathcal{V}$ the sets of measurable functions $u : \mathbb{R}^+ \to U$ and $v : \mathbb{R}^+ \to V$ respectively, is a *nonanticipative strategy for Victor* at point x_0 if, for any controls $u(\cdot)$ and $\tilde{u}(\cdot)$ of $\mathcal{U}$ which coincide almost everywhere on $[0, \theta]$, $\beta(u(\cdot))$ and $\beta(\tilde{u}(\cdot))$ coincide almost everywhere on $[0, \theta]$. A nonanticipative strategy for Ursula is defined symmetrically as a map $\alpha : \mathcal{V} \longrightarrow \mathcal{U}$.

The control structure of impulse systems, which has to account for the decision of jumping, the control of the destination of jumps, and the control of the continuous part of the trajectories, makes it difficult to define strategies with respect to control. Following Roxin in [12], we consider that the players choose their trajectories (instead of choosing controls). This provides a straightforward approach to the study of games with separated dynamics: A strategy for Ursula at $x_0 = (y_0, z_0)$ is an application $\mathbf{A} : S_{H,Q}(z_0) \longrightarrow S_{G,P}(y_0)$, while for Victor, it is an application $\mathbf{B} : S_{G,P}(y_0) \longrightarrow S_{H,Q}(z_0)$. These applications must be causal: A player's action at time t cannot depend on his/her opponent's action at time $\theta > t$. Now, using Varaiya's idea in [14], we assume that each player can observe and record his/her opponent's action from the initial time up to the current time t. This leads to the following definition of strategies.

Definition 3.1. We call a Varaiya–Roxin strategy (VR-strategy) for Ursula at initial condition $x_0 = (y_0, z_0)$ a map

$$\mathbf{A} : S_{H,Q}(z_0) \longrightarrow S_{G,P}(y_0)$$

such that for any $\theta > 0$, and for any trajectories $z(\cdot)$ and $\tilde{z}(\cdot)$ of $S_{H,Q}(z_0)$ which coincide on $[0, \theta]$, the trajectories $y(\cdot) = \mathbf{A}(z(\cdot))$ and $\tilde{y}(\cdot) = \mathbf{A}(\tilde{z}(\cdot))$ coincide on $[0, \theta]$.

We denote by $\mathbb{A}(x_0)$ the set of VR-strategies for Ursula at x_0.

A VR-strategy for Victor at initial condition $x_0 = (y_0, z_0)$ is defined symmetrically as a map

$$\mathbf{B} : S_{G,P}(y_0) \longrightarrow S_{H,Q}(z_0)$$

such that for any $\theta > 0$, and for any trajectories $y(\cdot)$ and $\tilde{y}(\cdot)$ of $S_{G,P}(y_0)$ which coincide on $[0, \theta]$, the trajectories $z(\cdot) = \mathbf{B}(y(\cdot))$ and $\tilde{z}(\cdot) = \mathbf{B}(\tilde{y}(\cdot))$ coincide on $[0, \theta]$.

We denote by $\mathbb{B}(x_0)$ the set of VR-strategies for Victor at x_0.

Let us mention that most of the results obtained for differential games with separated dynamics in the context of nonanticipative strategies hold true in the context of nonimpulsive VR-strategies (see Appendix B).

Remark 3.1. VR-strategies encompass feedback strategies. Indeed, feedback control for Ursula is given by $u : \mathbb{R}^l \times \mathbb{R}^m \to U$, Lipschitz continuous with respect to the first variable and continuous with respect to the second, $\tilde{A} \subset \mathrm{Dom}\,(P) \times \mathbb{R}^m$, closed, and $\mu : \tilde{A} \to M$, such that

$$\begin{cases} y'(t) &= g(y(t), u(y(t), z(t))) \quad \text{if } (y(t), z(t)) \notin \tilde{A} \\ y(t) &= p(y(t), \mu(y(t), z(t))) \quad \text{if } (y(t^-), z(t^-)) \in \tilde{A}. \end{cases} \tag{5}$$

For initial conditions (y_0, z_0), let us consider $z(\cdot) \in S_{H,Q}(z_0)$ and let $y(\cdot)$ be a solution to (5). Then $y(\cdot)$ is associated with $z(\cdot)$ in a nonanticipative way.

A drawback of VR-strategies (and of nonanticipative strategies for differential games) is that the game cannot be written in normal form. Let us recall that the outcome is defined by

$$\theta_{\mathcal{T}}^{K}(y(\cdot), z(\cdot)) := \inf\{t : (y(t), z(t)) \in \mathcal{T} \quad \text{and} \quad \forall s \le t, \quad (y(s), z(s)) \in K\}.$$

Then we can define two value functions for the game.

Definition 3.2. The lower value function is given by

$$\vartheta_{\mathcal{T}}^{b}(y_0, z_0) := \inf_{\mathbf{B} \in \mathbb{B}(y_0, z_0)} \sup_{y(\cdot) \in S_{G,P}(y_0)} \theta_{\mathcal{T}}^{K}(y(\cdot), \mathbf{B}(y(\cdot))).$$

The upper value function is given by

$$\vartheta_{\mathcal{T}}^{\#}(y_0, z_0) := \lim_{\varepsilon \to 0^+} \sup_{\mathbf{A} \in \mathbb{A}(y_0, z_0)} \inf_{z(\cdot) \in S_{H,Q}(z_0)} \theta_{\mathcal{T}_\varepsilon}^{K_\varepsilon}(\mathbf{A}(z(\cdot)), z(\cdot)),$$

where $\mathcal{T}_\varepsilon := (\mathcal{T} + \varepsilon\mathcal{B})$ and $K_\varepsilon := (K + \varepsilon\mathcal{B})$.

We say that the game has a value if these two value functions are equal.

Remark 3.2. Let us mention that the limit in the definition of the upper value function exists since

$$\sup_{\mathbf{A}\in\mathbb{A}(y_0,z_0)} \inf_{z(\cdot)\in S_{H,Q}(z_0)} \theta_{\mathcal{T}_\varepsilon}^{K_\varepsilon}(\mathbb{A}(z(\cdot)), z(\cdot))$$

is nondecreasing with respect to ε.

Following [5] for differential games, we shall characterize the lower value function by using properties of its epigraph: We will show that, if we consider an auxiliary qualitative game, $\mathrm{Epi}(\vartheta_{\mathcal{T}}^{b})$ is exactly the domain in which Victor, using a VR-strategy, can win against any trajectory chosen by Ursula. Therefore, proving the existence of a value for the pursuit-evasion game amounts to proving an alternative result for a qualitative game.

4 Victory Domains in a Qualitative Game

Let us recall that Ursula wins if she can lead, in finite time, the full state $x = (y, z)$ in an open domain $\Omega \subset \mathbb{R}^n$, while keeping it out of a closed set $\mathcal{T} \subset \mathbb{R}^n$. Victor has the opposite goal of keeping x inside $K := (\mathbb{R}^n \setminus \Omega)$ until he can lead it to $\mathcal{T}$. In the context of VR-strategies, the victory domains can be defined in the following way.

Definition 4.1. The *victory domain of Victor*, denoted by W_V, is the set of initial conditions $x_0 = (y_0, z_0)$ in K for which Victor can find a VR-strategy $\mathbf{B} \in \mathbb{B}(x_0)$ such that for all trajectories $y(\cdot) \in S_{G,P}(y_0)$ chosen by Ursula, the trajectory $x(\cdot) = (y(\cdot), \mathbf{B}(y(\cdot)))$ stays in K as long as the target $\mathcal{T}$ has not been reached.

The *victory domain of Ursula*, denoted by W_U, is the set of initial conditions $x_0 = (y_0, z_0)$ in K for which Ursula can find a VR-strategy $\mathbf{A} \in \mathbb{A}(x_0)$, a constant $\varepsilon > 0$, and a time $T \geq 0$ such that for all trajectories $z(\cdot) \in S_{H,Q}(z_0)$ chosen by Victor, the trajectory $x(\cdot) = (\mathbf{A}(z(\cdot)), z(\cdot))$ reaches the set

$$\Omega^{-\varepsilon} := \{x \in \Omega : \forall \xi \notin \Omega, \quad \|\xi - x\| \geq \varepsilon\} = \mathbb{R}^n \setminus K_\varepsilon$$

before T while staying out of $\mathcal{T}_\varepsilon$.

Following Cardaliaguet in [3] for a nonimpulsive version of this game, we use geometric conditions to define a class of closed sets, the *impulse discriminating domains*, which will play a key role in the characterization of Victor's victory domain. Then we provide sufficient conditions for the victory domains W_U and W_V to form a partition of K.

4.1 Impulse Discriminating Domains and Kernels

We define, for all $x \in \mathbb{R}^n$ and all $D \subset \mathbb{R}^n$ closed, the functions

$$\mathcal{H}(x, D) := \sup_{\pi\in NP_D(x)} \left\{ \sup_{u\in U} \inf_{v\in V} \langle f(x, u, v), \pi \rangle \right\}, \tag{6}$$

$$\mathcal{L}_V(x, D) := \inf_{\nu \in N} \chi_D(y, q(z, \nu)), \tag{7}$$

$$\mathcal{L}_{UV}(x, D) := \sup_{\mu \in M} \inf \left\{ \chi_D(p(y, \mu), z), \inf_{\nu \in N} \chi_D(p(y, \mu), q(z, \nu)) \right\}, \tag{8}$$

where $f(x, u, v) = f((y, z), u, v) = (g(y, u), h(z, v))$, $\chi_D(\cdot)$ denotes the characteristic function of the set D:

$$\chi_D(x) = \begin{cases} 0 & \text{if } x \in D \\ +\infty & \text{otherwise,} \end{cases}$$

and $NP_D(x)$ denotes the set of proximal normal to D at x:

$$NP_D(x) := \{\pi \in \mathbb{R}^n : \inf_{x' \in D} \left\| (x + \pi) - x' \right\| = \|\pi\|\}.$$

Definition 4.2 (Impulse discriminating domains). A closed subset D of $\mathbb{R}^n$ is an impulse discriminating domain for Victor with target $\mathcal{T}$ if and only if

$$\begin{aligned} &\forall x \in (D \setminus \mathcal{T}), \\ &\quad \max\{\min\{\mathcal{H}(x, D)\ ,\ \mathcal{L}_V(x, D \cup \mathcal{T})\}, \quad \mathcal{L}_{UV}(x, D \cup \mathcal{T})\} \le 0. \end{aligned} \tag{9}$$

Remark 4.1. The condition of Definition 4.2 is a condensed way of writing the following conditions:

(1) If no player can jump, that is, if $x \in \text{Dom}\,(P)^c \times \text{Dom}\,(Q)^c$, then

$$\forall \pi \in NP_D(x), \quad \sup_{u \in U} \inf_{v \in V} \langle f(x, u, v), \pi \rangle \le 0.$$

(2) If only Victor can jump, that is, if $x \in \text{Dom}\,(P)^c \times \text{Dom}\,(Q)$, then

$$\forall \pi \in NP_D(x), \quad \sup_{u \in U} \inf_{v \in V} \langle f(x, u, v), \pi \rangle \le 0 \quad \text{or} \quad (\{y\} \times Q(z)) \cap (D \cup \mathcal{T}) \ne \emptyset.$$

(3) If only Ursula can jump, that is, if $x \in \text{Dom}\,(P) \times \text{Dom}\,(Q)^c$, then

$$\forall \pi \in NP_D(x), \quad \sup_{u \in U} \inf_{v \in V} \langle f(x, u, v), \pi \rangle \le 0 \quad \text{and} \quad (P(y) \times \{z\}) \subset (D \cup \mathcal{T}).$$

(4) If both players can jump, that is, if $x \in \text{Dom}\,(P) \times \text{Dom}\,(Q)$, then

$$\left(\forall \pi \in NP_D(x),\ \sup_{u \in U} \inf_{v \in V} \langle f(x, u, v), \pi \rangle \le 0 \quad \text{or} \quad (\{y\} \times Q(z)) \cap (D \cup \mathcal{T}) \ne \emptyset \right)$$

$$\text{and} \qquad \forall y' \in P(y),\ (\{y'\} \times (\{z\} \cup Q(z))) \cap (D \cup \mathcal{T}) \ne \emptyset.$$

Let us mention that, when no jumps are allowed, condition (9) amounts to

$$\forall x \in (D \setminus \mathcal{T}), \quad \forall \pi \in NP_D(x), \quad \sup_{u \in U} \inf_{v \in V} \langle f(x, u, v), \pi \rangle \le 0,$$

which defines a discriminating domain as in [3] for differential games. Theorem 4.1 states that the interpretation of impulse discriminating domains with VR-strategies is similar to the interpretation of discriminating domains with nonanticipative strategies (Theorem 2.1 in [3]). We need the following assumption, which is equivalent to a local controllability assumption on the boundary of the reset sets.

Assumption 2. *For all time $\theta > 0$, there exist a neighborhood $\mathcal{N}_U$ of* Dom (P) *and a neighborhood $\mathcal{N}_V$ of* Dom (Q) *such that*

$$\begin{array}{ll}\text{(i)} & \forall y_0 \in \mathcal{N}_U, \quad \exists y(\cdot) \in S_G(y_0) \quad \textit{such that } \exists\, t_0 \leq \theta, \quad y(t_0) \in \mathrm{Dom}\,(P) \\ \text{(ii)} & \forall z_0 \in \mathcal{N}_V, \quad \exists z(\cdot) \in S_H(z_0) \quad \textit{such that } \exists\, t_0 \leq \theta, \quad z(t_0) \in \mathrm{Dom}\,(Q)\,.\end{array}$$

Theorem 4.1. *Let D and $\mathcal{T}$ be closed subsets of $\mathbb{R}^n$. Under Assumptions 1 and 2, D is an impulse discriminating domain for Victor with target $\mathcal{T}$ if and only if, for all $x_0 = (y_0, z_0) \in D$, there exists a VR-strategy $\mathbf{B}$ such that, for any $y(\cdot) \in S_{G,P}(y_0)$, the trajectory $(y(\cdot), \mathbf{B}(y(\cdot)))$ stays in D as long as $\mathcal{T}$ has not been reached, namely:*

$$\forall t \leq \inf\{s : (y(s), \mathbf{B}(y(\cdot))(s)) \in \mathcal{T}\}, \qquad (y(t), \mathbf{B}(y(\cdot))(t)) \in D.$$

The proof of this theorem in the nonimpulsive case can be found in Proposition B.1 in Appendix B. The proof in the impulsive case is given in Appendix C.1.

Theorem 4.1 shows the relevance of the notion of impulse discriminating domain for the study of the game from Victor's point of view: If K does not enjoy this property, it is natural to look at the subsets of K which do, and in particular, to the largest of them.

Definition 4.3. We call an *impulse discriminating kernel of K for Victor with target $\mathcal{T}$* the largest closed impulse discriminating domain for Victor with target $\mathcal{T}$ contained in K (when it exists). It is denoted $\mathrm{Disc}_{\mathcal{H},\mathcal{L}_V,\mathcal{L}_{UV}}(K, \mathcal{T})$.

Proposition 4.1 (Existence of the impulse discriminating kernel). *Let K be a closed subset of $\mathbb{R}^n$. Under Assumptions 1 and 2, the impulse discriminating kernel of K for Victor with target $\mathcal{T}$ exists (it may be empty). It contains all the closed impulse discriminating domains for Victor with target $\mathcal{T}$ contained in K.*

See Appendix C.2 for the proof.

4.2 Characterization of Victor's Victory Domain

We are now ready to state the main result of this section, which is proved in Appendix C.3.1.

Theorem 4.2. *Under Assumptions 1 and 2, Victor's victory domain is the impulse discriminating kernel of K with target $\mathcal{T}$ under $(\mathcal{H}, \mathcal{L}_V, \mathcal{L}_{UV})$, that is:*

$$W_V = \mathrm{Disc}_{\mathcal{H},\mathcal{L}_V,\mathcal{L}_{UV}}(K, \mathcal{T}).$$

Theorem 4.2 leads to several sufficient conditions for the alternative to hold true.

Proposition 4.2. *The victory domains W_U and W_V form a partition of K in the following situations:*

(i) If $(\mathrm{Dom}\,(P) \times \mathrm{Dom}\,(Q)) \cap K = \emptyset$*, which happens in particular if only one player enjoys impulsive dynamics,*
(ii) If for all $x \in W_V$,

$$[\mathcal{H}(x, W_V) \leq 0 \quad \textit{and} \quad (P(y) \times Q(z)) \cap W_V = \emptyset] \ \textit{or}$$
$$[(P(y) \times Q(z)) \subset W_V] \, .$$

See Appendix C.3.2 for the proof.

Let us emphasize that the alternative condition does not hold true in general. This is due to the fact that when a player who uses a VR-strategy decides whether to jump or not, he/she knows if his/her opponent has decided to jump or not. This piece of information is crucial as shown in the following example.

Example 4.1. Let us consider the following game with both dynamics in $\mathbb{R}$ defined by

$$\begin{cases} g(y,u) & = 2u \text{ with } u \in [-1,1], \\ p(y) & = y+2 \text{ if } |y| \leq 1, \end{cases} \quad \text{and} \quad \begin{cases} h(z,v) & = v \text{ with } v \in [-1,1], \\ q(z) & = z+3 \text{ if } |z| \leq 1, \end{cases}$$

and with $K := \{(y,z) : |y-z| \geq 1\}$ and $\mathcal{T} = \{(y,z) : |y-z| \geq 2\}$. Then at initial condition $(0, 1)$, Ursula wins if both players jump or if neither of them does, and Victor wins if one, and only one, player jumps. Therefore, the player who uses a VR-strategy can ensure his/her victory.

5 Characterization of the Value Functions

We now come back to the capture time problem stated in the introduction. Using the approach of [5], we use an auxiliary quantitative game played in $(K \times \mathbb{R}^+)$ with target $(\mathcal{T} \times \mathbb{R}^+)$ for Victor, and with dynamics (G, P), and $(H \times \{-1\}, Q \times Id_{\mathbb{R}})$ for Ursula and Victor respectively. The add-on variable $\rho \in \mathbb{R}$ stands for a time counter: Now, a point (x_0, ρ_0) belongs to Victor's victory domain if and only if Victor can find a strategy to drive x into $\mathcal{T}$ without leaving K in a time smaller than or equal to ρ_0.

We shall characterize the epigraph of the lower value function and investigate the existence of a value using the results of the previous section applied to this auxiliary game.

Remark 5.1. Let us mention that $S_{H\times\{-1\},Q\times Id_{\mathbb{R}}}(z_0,\rho_0) = S_{H,Q}(z_0)\times\{\rho_0-\cdot\}$, so that Victor can only choose a trajectory $z(\cdot)$. Therefore, we will consider VR-strategies

$$\mathbf{A} : S_{H,Q}(z_0) \longrightarrow S_{G,P}(y_0) \quad \text{and} \quad \mathbf{B} : S_{G,P}(y_0) \longrightarrow S_{H,Q}(z_0)$$

for Ursula and Victor respectively.

Following the approach of Section 4, we define the functions

$$\begin{aligned}
\widetilde{\mathcal{H}}(x,\rho,D) &:= \sup_{(\pi_x,\pi_\rho)\in NP_D(x,\rho)} \left\{ \sup_{u\in U}\inf_{v\in V} \langle f(x,u,v),\pi_x\rangle - \pi_\rho \right\},\\
\widetilde{\mathcal{L}}_V(x,\rho,D) &:= \inf_{\nu\in N} \chi_D(y,q(z,\nu),\rho),\\
\widetilde{\mathcal{L}}_{UV}(x,\rho,D) &:= \sup_{\mu\in M}\ \inf \left\{ \chi_D(p(y,\mu),z,\rho),\ \inf_{\nu\in N}\chi_D(p(y,\mu),q(z,\nu),\rho) \right\}.
\end{aligned}$$

Then we can state the main result of this section.

Theorem 5.1. *Under Assumptions 1 and 2,*

$$\operatorname{Epi}\left(\vartheta_{\mathcal{T}}^{b}\right) = \operatorname{Disc}_{\widetilde{\mathcal{H}},\widetilde{\mathcal{L}}_V,\widetilde{\mathcal{L}}_{UV}}\left(K\times\mathbb{R}^+,\mathcal{T}\times\mathbb{R}^+\right). \tag{10}$$

Proof. In order to prove that

$$\operatorname{Epi}\left(\vartheta_{\mathcal{T}}^{b}\right) \subset \operatorname{Disc}_{\widetilde{\mathcal{H}},\widetilde{\mathcal{L}}_V,\widetilde{\mathcal{L}}_{UV}}\left(K\times\mathbb{R}^+,\mathcal{T}\times\mathbb{R}^+\right), \tag{11}$$

let $x_0 = (y_0,z_0)$ belong to the domain of $\vartheta_{\mathcal{T}}^{b}$ and let $\rho_0 > \vartheta_{\mathcal{T}}^{b}(x_0)$. Then, by the very definition of $\vartheta_{\mathcal{T}}^{b}$, there exists a VR-strategy $\mathbf{B}\in\mathbb{B}(x_0)$ such that for all $y(\cdot)\in S_{G,P}(y_0)$, the trajectory $x(\cdot)=(y(\cdot),\mathbf{B}(y(\cdot)))$ is such that $\theta_{\mathcal{T}}^{K}(x(\cdot))\le\rho_0$. This means that the trajectory $(y(\cdot),\mathbf{B}(y(\cdot)),\rho(\cdot))$, where $\rho(t)=\rho_0-t$, stays in $(K\times\mathbb{R}^+)$ as long as $(\mathcal{T}\times\mathbb{R}^+)$ has not been reached. Therefore, $(x_0,\rho_0)\in \operatorname{Disc}_{\widetilde{\mathcal{H}},\widetilde{\mathcal{L}}_V,\widetilde{\mathcal{L}}_{UV}}\left(K\times\mathbb{R}^+,\mathcal{T}\times\mathbb{R}^+\right)$. We have proved that

$$\forall\rho_0>\vartheta_{\mathcal{T}}^{b}(x_0),\quad (x_0,\rho_0)\in\operatorname{Disc}_{\widetilde{\mathcal{H}},\widetilde{\mathcal{L}}_V,\widetilde{\mathcal{L}}_{UV}}\left(K\times\mathbb{R}^+,\mathcal{T}\times\mathbb{R}^+\right).$$

Inclusion (11) follows from the fact that the impulse discriminating kernel is closed.

In order to prove the converse inclusion, let (y_0,z_0,ρ_0) belong to the impulse discriminating kernel. Then, Victor can find a VR-strategy $\mathbf{B}$ such that for all $y(\cdot)\in S_{G,P}(y_0)$, the trajectory $(y(\cdot),\mathbf{B}(y(\cdot)),\rho(\cdot))$, where $\rho(t)=\rho_0-t$, stays in $(K\times\mathbb{R}^+)$ as long as $(\mathcal{T}\times\mathbb{R}^+)$ has not been reached. This means that

$$\forall t\le\theta_{\mathcal{T}}(y(\cdot),\mathbf{B}(y(\cdot))),\quad (y(t),\mathbf{B}(y(\cdot))(t))\in K \quad\text{and}\quad \rho_0-t\ge 0.$$

So we have

$$\vartheta_{\mathcal{T}}^{b}(x_0)\le\theta_{\mathcal{T}}(y(\cdot),z(\cdot))\le\rho_0.$$

Therefore, $(x_0,\rho_0)\in\operatorname{Epi}\left(\vartheta_{\mathcal{T}}^{b}\right)$, which completes the proof. □

Corollary 5.1. *The lower value function $\vartheta^b_{\mathcal{T}}$ is the lowest positive lower semicontinuous function such that $\vartheta^b_{\mathcal{T}}(x) = 0$ if $x \in \mathcal{T}$ and*

$$\begin{aligned}\min\Bigg\{ &\max\Big\{ \inf_{\pi\in\partial_-\vartheta^b_{\mathcal{T}}(x)} \inf_{u\in U}\sup_{v\in V} \langle f(x,u,v),\pi\rangle + 1,\\ &\qquad \sup_{\nu\in N}\left(\vartheta^b_{\mathcal{T}}(y,q(z,\nu)) - \vartheta^b_{\mathcal{T}}(y,z)\right)\Big\},\\ &\inf_{\mu\in M}\sup\Big\{\left(\vartheta^b_{\mathcal{T}}(p(y,\mu),z) - \vartheta^b_{\mathcal{T}}(y,z)\right),\\ &\qquad \sup_{\nu\in N}\left(\vartheta^b_{\mathcal{T}}(p(y,\mu),q(z,\nu)) - \vartheta^b_{\mathcal{T}}(y,z)\right)\Big\}\Bigg\} \ \geq\ 0,\end{aligned} \tag{12}$$

where $\partial_-\vartheta(x)$ denotes the subdifferential of ϑ at x.

Proof. The lower semicontinuity of $\vartheta^b_{\mathcal{T}}$ can be deduced from the fact that the epigraph of $\vartheta^b_{\mathcal{T}}$ is an impulse discriminating kernel, which is closed.

Let us prove that $\vartheta^b_{\mathcal{T}}$ satisfies (12). From (10) and the definition of the impulse discriminating domains, we have

$$\begin{aligned}\max\Big\{&\min\Big\{\widetilde{\mathcal{H}}\left(x,\vartheta^b_{\mathcal{T}}(x),\ \mathrm{Epi}\left(\vartheta^b_{\mathcal{T}}\right)\right), \widetilde{\mathcal{L}}_V\left(x,\vartheta^b_{\mathcal{T}}(x),\mathrm{Epi}\left(\vartheta^b_{\mathcal{T}}\right)\right)\Big\},\\ &\widetilde{\mathcal{L}}_{UV}\left(x,\vartheta^b_{\mathcal{T}}(x),\mathrm{Epi}\left(\vartheta^b_{\mathcal{T}}\right)\right)\Big\} \leq\ 0.\end{aligned} \tag{13}$$

We recall that the statement $\widetilde{\mathcal{H}}\left(x,\vartheta^b_{\mathcal{T}}(x),\mathrm{Epi}\left(\vartheta^b_{\mathcal{T}}\right)\right) \leq 0$ is equivalent to

$$\max_{(\pi_x,\pi_\rho)\in NP_{\mathrm{Epi}\left(\vartheta^b_{\mathcal{T}}\right)}(x)} \quad \sup_{u\in U}\inf_{v\in V}\langle f(x,u,v),\pi_x\rangle - \pi_\rho \leq 0.$$

Therefore, from [10], it is also equivalent to

$$\forall\pi\in\partial_-\vartheta^b_{\mathcal{T}}(x),\quad \inf_{u\in U}\sup_{v\in V}\langle f(x,u,v),\pi\rangle + 1 \geq 0.$$

Moreover, we can write

$$\widetilde{\mathcal{L}}_V\left(x,\vartheta^b_{\mathcal{T}}(x),\mathrm{Epi}\left(\vartheta^b_{\mathcal{T}}\right)\right) \leq 0 \ \Leftrightarrow\ \sup_{\nu\in N}\left(\vartheta^b_{\mathcal{T}}(y,q(z,\nu)) - \vartheta^b_{\mathcal{T}}(y,z)\right) \geq 0,$$

and similarly, the statement $\widetilde{\mathcal{L}}_{UV}\left(x,\vartheta^b_{\mathcal{T}}(x),\mathrm{Epi}\left(\vartheta^b_{\mathcal{T}}\right)\right) \leq 0$ is equivalent to

$$\begin{aligned}\inf_{\mu\in M}\sup\Big\{&\left(\vartheta^b_{\mathcal{T}}(p(y,\mu),z) - \vartheta^b_{\mathcal{T}}(y,z)\right),\\ &\sup_{\nu\in N}\left(\vartheta^b_{\mathcal{T}}(p(y,\mu),q(z,\nu)) - \vartheta^b_{\mathcal{T}}(y,z)\right)\Big\} \geq 0.\end{aligned}$$

We have proved that $\vartheta^b_{\mathcal{T}}$ is a lower semicontinuous function which satisfies (12). It is the smallest because its epigraph is the largest closed set such that (13) is satisfied. □

Theorem 5.2. *If one of the following assumptions holds true:*

(i) If $(\mathrm{Dom}\,(P) \times \mathrm{Dom}\,(Q)) \cap K = \emptyset$*, which happens in particular if only one player enjoys impulsive dynamics;*
(ii) For all x*, such that* $\vartheta_T^b(x) < +\infty$*, the two following conditions hold true:*

$$\begin{array}{ll} (a) & \forall \pi \in \partial_- \vartheta_T^b(x), \quad \inf_{v \in V} \sup_{u \in U} \langle f(x,u,v), \pi \rangle - 1 \leq 0 \\ & \quad \textit{and} \quad \forall \mu \in M,\ \forall \nu \in N, \quad \vartheta_T^b(p(y,\mu), q(z,\nu)) > \vartheta_T^b(x); \\ (b) & \forall \mu \in M,\ \forall \nu \in N, \quad \vartheta_T^b(p(y,\mu), q(z,\nu)) \leq \vartheta_T^b(x); \end{array}$$

then the game has a value, that is $\vartheta_T^b = \vartheta_T^\#$.

Proof. First, let us emphasize the fact that the assumptions of Theorem 5.2 yield the assumptions of Proposition 4.2 for the game with dynamics (G, P) and $(H \times \{-1\}, Q \times Id)$. Therefore, using Theorem 5.1, Proposition 4.2, and Definition 4.1, we know that Epi $\left(\vartheta_T^b\right)$ is exactly the set of points (x_0, ρ_0) such that for all $\varepsilon > 0$ and for all $T \geq 0$,

$$\forall \mathbf{A} \in \mathbb{A}(x_0),\ \exists z(\cdot) \in S_{H,Q}(z_0), \quad (\mathbf{A}(z(\cdot))(t), z(t), \rho_0 - t) \in K_\varepsilon \times [-\varepsilon, +\infty),$$

if $t \leq \min\{T, \inf\{s\colon (\mathbf{A}(z(\cdot))(s), z(s), \rho_0 - s) \in \mathcal{T}_\varepsilon \times [-\varepsilon, +\infty)\}\}$. This is equivalent to

$$\forall \varepsilon > 0, \quad \sup_{\mathbf{A} \in \mathbb{A}(x_0)} \inf_{z(\cdot) \in S_{H,Q}(z_0)} \theta_{\mathcal{T}_\varepsilon}^{K_\varepsilon}(\mathbf{A}(z(\cdot)), z(\cdot)) \leq \rho_0 + \varepsilon.$$

We can take the limit when ε tends to 0 in the previous expression, so we have proved the equality between Epi $\left(\vartheta_T^b\right)$ and Epi $\left(\vartheta_T^\#\right)$. □

6 Concluding Remarks

In this chapter, we have introduced the notion of impulse discriminating kernels which can be used to characterize victory domains of qualitative games and epigraphs of value functions for quantitative games in the case of separated impulsive dynamics.

These results can be extended to hybrid games with separated dynamics. Indeed, most hybrid dynamics can be written as impulse dynamics (see [2]).

As was the case for (nonimpulse) discriminating kernels (see [5]) and for impulse viability kernels (see [7]), the geometrical characterization of impulse discriminating kernels should lead to a numerical approximation algorithm. This is work in progress.

Appendix A Cardaliaguet–Plaskacz Lemma

We state here a result that is often used in the proofs of the following appendices. In the sequel, we call it the *Cardaliaguet–Plaskacz Lemma* because it is a consequence of Lemma 2.7 in [4] (see also Lemma 4.1 of [3]).

Definition A.1. A set-valued map $\mathcal{B} : S_{G,P}(y_0) \rightsquigarrow S_{H,Q}(z_0)$ is said to be nonanticipative if for any $s \geq 0$, for any pair of trajectories $y_1(\cdot)$ and $y_2(\cdot)$ of $S_{G,P}(y_0)$ which coincide over $[0, s]$, and for any $z_1(\cdot) \in \mathcal{B}(y_1(\cdot))$, there exists $z_2(\cdot) \in \mathcal{B}(y_2(\cdot))$ which coincides with $z_1(\cdot)$ over $[0, s]$.

Lemma A.1. *Let $\mathcal{B} : S_{G,P}(y_0) \rightsquigarrow S_{H,Q}(z_0)$ be a nonanticipative set-valued map with nonempty closed values. Then there exists a non-anticipative selection of $\mathcal{B}$ which is a VR-strategy.*

Appendix B VR-Strategies for Differential Games

In this appendix, we consider a nonimpulsive version of the qualitative game of Section 4 with dynamics

$$y'(t) = g(y(t), u(t)), \quad u(t) \in U \tag{14}$$

$$z'(t) = h(z(t), v(t)), \quad v(t) \in V. \tag{15}$$

An alternative theorem has been proved in [3] in the context of non-anticipative strategies. We prove here that this alternative result holds true in the context of VR-strategies:

$$\widetilde{\mathbf{A}} : S_H(z_0) \longrightarrow S_G(y_0) \quad \text{and} \quad \widetilde{\mathbf{B}} : S_G(y_0) \longrightarrow S_H(z_0)$$

with the same victory domains. This result is the counterpart of Theorem 2.1 and Theorem 2.3 from [3] using VR-strategies.

Proposition B.1. *We assume that g and h satisfy the conditions of Assumption 1 and we set $f(x, u, v) := (g(y, u), h(z, v))$. Let D and $\mathcal{T}$ be closed subsets of $\mathbb{R}^n$; then the two following propositions are equivalent:*

(i) D is a discriminating domain for Victor with target $\mathcal{T}$, namely,

$$\forall x \in (D \setminus \mathcal{T}), \qquad \forall \pi \in NP_D(x), \quad \sup_{u \in U} \inf_{v \in V} \langle f(x, u, v), \pi \rangle \leq 0.$$

(ii) for all $x_0 = (y_0, z_0) \in D$, there exists a VR-strategy $\widetilde{\mathbf{B}}$ such that for any $\tilde{y}(\cdot) \in S_G(y_0)$, the trajectory $(\tilde{y}(\cdot), \widetilde{\mathbf{B}}(\tilde{y}(\cdot)))$ stays in D as long as $\mathcal{T}$ has not been reached.

And the two following propositions are equivalent:

(iii) D is a leadership domain for Victor with target $\mathcal{T}$, namely,

$$\forall x \in (D \setminus \mathcal{T}), \qquad \forall \pi \in NP_D(x), \quad \inf_{v \in V} \sup_{u \in U} \langle f(x, u, v), \pi \rangle \leq 0.$$

(iv) for all $x_0 = (y_0, z_0) \in D$, for all VR-strategy $\widetilde{\mathbf{A}}$, for all time $T \geq 0$ and for all $\varepsilon > 0$, there exists $\tilde{z}(\cdot) \in S_H(z_0)$ such that the trajectory $(\widetilde{\mathbf{A}}(\tilde{z}(\cdot)), \tilde{z}(\cdot))$ stays in D_ε as long as $\mathcal{T}_\varepsilon$ has not been reached or up to time T.

Remark B.1. Since $f(x,u,v) := (g(y,u), h(z,v))$, the Isaacs condition holds true, that is,

$$\sup_{u\in U}\inf_{v\in V}\langle f(x,u,v),\pi\rangle = \inf_{v\in V}\sup_{u\in U}\langle f(x,u,v),\pi\rangle\,,$$

and discriminating domains are leadership domains (the converse is always true).

Proof of Proposition B.1

(i) ⇒ (ii): Let us set $x_0 = (y_0, z_0) \in D$. From Theorem 2.1 in [3], Victor can find a nonanticipative strategy β such that for any control $u(\cdot)$, played by Ursula, the trajectory $x[x_0, u(\cdot), \beta(u(\cdot))]$ stays in D as long as $\mathcal{T}$ has not been reached. Let us define the set-valued map

$$\begin{array}{rcl}\mathcal{B}_{(y_0,z_0)} : S_G(y_0) & \rightsquigarrow & S_H(z_0)\\ y(\cdot) & \mapsto & \overline{cl}\,\{z[z_0,\beta(u(\cdot))] : y[y_0,u(\cdot)] = y(\cdot)\}\,.\end{array}$$

It has obviously nonempty values and is nonanticipative by construction, so using the Cardaliaguet–Plaskacz Lemma, we can prove that it has a single-valued selection which is a VR-strategy.

(ii) ⇒ (i): Let us set $u \in U$ and let us assume that Ursula plays trajectories $y'(t) = g(y(t), u)$. Victor can find a winning strategy only if

$$\begin{array}{l}\forall x_0 \in D,\quad \exists x(\cdot) \in S_{f(\cdot,u,V)}(x_0) \text{ such that}\\ \qquad\qquad\qquad \forall t \le \inf\{s : x(s) \in \mathcal{T}\},\quad x(t) \in D.\end{array} \tag{16}$$

From the Viability Theorem (see Theorem 3.3.5 in [1]), (16) yields

$$\forall x_0 \in D\setminus\mathcal{T},\quad \forall \pi \in NP_D(x_0),\quad \inf_{v\in V}\langle f(x_0,u,v),\pi\rangle \le 0. \tag{17}$$

Since (17) holds true for all $u \in U$, our claim (i) is allowed.

(iii) ⇒ (iv): We proceed by contradiction. For this purpose, we assume that there exists $x_0 = (y_0, z_0) \in D$, a time $T \ge 0$, $\varepsilon > 0$, and a VR-strategy $\widetilde{\mathbf{A}}$ such that, for all trajectories $z(\cdot)$ chosen by Victor, the trajectory $(\widetilde{\mathbf{A}}(\tilde{z}(\cdot)), \tilde{z}(\cdot))$ leaves D_ε before T without entering in $\mathcal{T}_\varepsilon$. Let us define the set-valued map

$$\begin{array}{rcl}\mathcal{A}_{(y_0,z_0)} : \mathcal{V} & \rightsquigarrow & \mathcal{U}\\ v(\cdot) & \mapsto & \left\{u(\cdot) : y[y_0,u(\cdot)] = \widetilde{\mathbf{A}}(z[z_0,v(\cdot)])\right\}\end{array}$$

We can prove that $\mathcal{A}_{(y_0,z_0)}$ has a single-valued nonanticipative selection using the Cardaliaguet–Plaskacz Lemma; we denote it α. Now, let us set $v(\cdot) \in \mathcal{V}$. Then by construction, the trajectory $x[x_0, \alpha(v(\cdot)), v(\cdot)]$ leaves D_ε before T without entering in $\mathcal{T}_\varepsilon$. From Theorem 2.3 in [3], this contradicts the fact that D is a leadership domain for Victor.

(iv) ⇒ (iii): We will prove that for all $x_0 \in (D \setminus \mathcal{T})$

$$\forall y(\cdot) \in S_G(y_0), \exists \tau > 0, \quad \exists z(\cdot) \in S_H(z_0), \quad \forall t \leq \tau, \quad (y(t), z(t)) \in D. \tag{18}$$

From [1], this yields that D is a discriminating domain.

Let us set $x_0 \in (D \setminus \mathcal{T})$, $T > 0$, $\overline{y}(\cdot) \in S_G(y_0)$, and let us define the constant VR-strategy $\widetilde{\mathbf{A}}(z(\cdot)) = \overline{y}(\cdot)$. Then, for all $k \in \mathbb{N}$, there exists $z_k(\cdot) \in S_H(z_0)$ such that the trajectory

$$(\widetilde{\mathbf{A}}(z_k(\cdot)), z_k(\cdot)) = (\overline{y}(\cdot), z_k(\cdot))$$

stays in $D_{\frac{1}{k}}$ as long as $\mathcal{T}_{\frac{1}{k}}$ has not been reached or up to time T. Then the sequence $z_k(\cdot)$ converges to $z(\cdot) \in S_H(z_0)$ up to a subsequence and we have

$$\forall t \leq \min\{T, \inf\{s : (y(t), z(t)) \in \mathcal{T}\}\}, \ (y(t), z(t)) \in D.$$

Since $x_0 \notin \mathcal{T}$, we have $\inf\{s : (y(t), z(t)) \in \mathcal{T}\} > 0$, so the proof is complete. □

Appendix C Proofs of Section 4

C.1 Proof of the Interpretation Theorem 4.1

Part 1: Let us assume that D is an impulse discriminating domain and let us set $x_0 = (y_0, z_0) \in D$. We have to find a VR-strategy $\mathbf{B}$ such that for all $y(\cdot) \in S_{G,P}(y_0)$, the trajectory $(y(\cdot), \mathbf{B}(y(\cdot)))$ stays in D until reaching $\mathcal{T}$. For this purpose, let us define the set

$$E := \{x \in D : \mathcal{L}_V(x, D) \leq 0\}.$$

Then we can check that E is closed, and that D is a (nonimpulse) discriminating domain for Victor with target $(\mathcal{T} \cup E)$. Indeed, from the very definition of an impulse discriminating domain, we have

$$\forall x \in (D \setminus (E \cup \mathcal{T})), \quad \forall \pi \in NP_D(x), \ \mathcal{H}(x, \pi) \leq 0. \tag{19}$$

$$\begin{aligned} &\forall (y, z) \in ((D \setminus \mathcal{T}) \cap (\mathrm{Dom}\,(P) \times \mathbb{R}^m)), \\ &\quad \forall y' \in P(y), \quad (y', z) \in (D \cup \mathcal{T} \cup E) \end{aligned} \tag{20}$$

From Proposition B.1, (19) yields that, for all $x = (y, z) \in (D \setminus (E \cup \mathcal{T}))$, there exists a nonimpulsive VR-strategy $\widetilde{\mathbf{B}}_x$ such that, for all $\tilde{y}(\cdot) \in S_G(y)$, the trajectory $(\tilde{y}(\cdot), \widetilde{\mathbf{B}}(\tilde{y}(\cdot)))$ stays in D until reaching $(E \cup \mathcal{T})$. We use this result to construct an impulse VR-strategy for Victor to win. For this purpose, let us define a feedback map

$$\begin{array}{rlcl} b : & E & \longrightarrow & N \\ & x = (y, z) & \mapsto & \nu \text{ such that } (y, q(z, \nu)) \in (D \cup \mathcal{T}). \end{array}$$

Then we can map each trajectory $y(\cdot) \in S_{G,P}(y_0)$ to the unique trajectory $z(\cdot) = \mathbf{B}(y(\cdot))$ such that the run $(\tau_i, (y_i, z_i), (y_i(\cdot), z_i(\cdot)))_{i\in I}$ associated with the trajectory $(y(\cdot), \mathbf{B}(y(\cdot)))$ satisfies

$$\forall i \in I, \quad \begin{cases} z_i(\cdot) &= \widetilde{\mathbf{B}}_{(y_i,z_i)}(y_i(\cdot)) \\ z_{i+1} &= \begin{cases} z_i(\tau_i) & \text{if } (y_i, z_i(\tau_i)) \notin E, \\ b(y_i, z_i(\tau_i)) & \text{otherwise.} \end{cases} \end{cases}$$

Then $\mathbf{B}$ is obviously a VR-strategy. Moreover, we can prove by iteration that for all $y(\cdot) \in S_{G,P}(y_0)$ the trajectory $x(\cdot) = (y(\cdot), \mathbf{B}(y(\cdot)))$ stays in D until reaching $\mathcal{T}$.

Part 2: Conversely, let us choose $x_0 = (y_0, z_0) \in (D \setminus \mathcal{T})$ at which there exists a VR-strategy $\mathbf{B}$ such that for all $y(\cdot) \in S_{G,P}(y_0)$, the trajectory $(y(\cdot), \mathbf{B}(y(\cdot)))$ stays in D until reaching $\mathcal{T}$.

Case 1. Let us set $u \in U$ and let us assume that Ursula plays a trajectory without jumps such that $y'(t) = g(y(t), u)$. Then we have

$$(y(\cdot), \mathbf{B}(y(\cdot))) \in S_{f(\cdot,u,V),R_V}(x_0), \quad \text{with } R_V(y, z) = y \times Q(z),$$

and the trajectory $(y(\cdot), \mathbf{B}(y(\cdot)))$ stays in D as long as $\mathcal{T}$ has not been reached. From [8], this yields

$$\left(\forall \pi \in NP_D(x_0),\ \inf_{v\in V} \langle f(x_0, u, v), \pi\rangle \leq 0\right) \text{ or } R_V(x_0) \cap (D \cup \mathcal{T}) \neq \emptyset. \tag{21}$$

Case 2. We assume that $y_0 \in \mathrm{Dom}\,(P)$. Let us set $\mu \in M$ and let us assume that Ursula plays a trajectory such that

$$\begin{cases} y'(t) &= g(y(t), u) \quad \text{if } y(t) \notin \mathrm{Dom}\,(P)\,, \\ y(t) &= p(y(t), \mu) \quad \text{if } y(t^-) \in \mathrm{Dom}\,(P)\,. \end{cases}$$

Since the trajectory $(y(\cdot), \mathbf{B}(y(\cdot)))$ stays in D until reaching $\mathcal{T}$, we have

$$\begin{array}{ll} \text{if } z_0 \notin \mathrm{Dom}\,(Q)\,, & (p(y_0, \mu), z_0) \in (D \cup \mathcal{T}) \\ \text{if } x_0 \in \mathrm{Dom}\,(Q)\,, & [\{p(y_0, \mu)\} \times (\{z_0\} \cup Q(z_0))] \cap (D \cup \mathcal{T}) \neq \emptyset. \end{array} \tag{22}$$

Since (21) holds true for all $u \in U$ and (22) holds true for all $\mu \in M$, we have proved that D is an impulse discriminating domain for Victor with target $\mathcal{T}$. □

C.2 Proof of the Existence of the Discriminating Kernel (Proposition 4.1)

Let us denote by D^* the closure of the union of all the closed impulse discriminating domains for Victor with target $\mathcal{T}$ contained in K. We have to prove that D^* is an impulse discriminating domain for Victor with target $\mathcal{T}$ (by definition, D^* is closed and contained in K). For this purpose, let us set $x \in D^*$. Then by definition of D^*, there exists a sequence $\{x^k\}_{k\in\mathbb{N}}$ such that

$$\forall k \in \mathbb{N},\ x_k \in D_k, \quad \text{and} \quad \lim_{k\to+\infty} x^k = x,$$

where the $D^k \subset K$ are closed impulse discriminating domains for Victor with target $\mathcal{T}$.Therefore, we have

$$\forall k \in \mathbb{N},\ \max\{\min\{\mathcal{H}(x^k, D^k)\,,\ \mathcal{L}_V(x^k, D^k \cup \mathcal{T})\}\,,\ \mathcal{L}_{UV}(x^k, D^k \cup \mathcal{T})\} \le 0.$$

Proof of $\mathcal{L}_{UV}(x, D^* \cup \mathcal{T}) \le 0$.

We proceed by contradiction: Let us assume that $\mathcal{L}_{UV}(x, D^* \cup \mathcal{T}) > 0$. Then we have $y \in \mathrm{Dom}\,(P)$, and we can find μ such that

$$(p(y,\mu), z) \notin (D^* \cup \mathcal{T}) \ \text{ and } \ \forall \nu \in N,\ (p(y,\mu), q(z,\nu)) \notin (D^* \cup \mathcal{T}).$$

By continuity of the jump functions, we can find $\eta > 0$ and $\varepsilon > 0$ such that for all $\tilde{y} \in (y + \eta\mathcal{B}) \cap \mathrm{Dom}\,(P)$,

$$\forall \tilde{z} \in (z + \eta\mathcal{B}), \qquad (p(\tilde{y},\mu), z) \notin (D^* \cup \mathcal{T}) + \varepsilon\mathcal{B}, \tag{23}$$

$$\forall \tilde{z} \in (z + \eta\mathcal{B}) \cap \mathrm{Dom}\,(Q)\,, \quad \forall \nu \in N, (p(\tilde{y},\mu), q(\tilde{z},\nu)) \notin (D^* \cup \mathcal{T}) + \varepsilon\mathcal{B}. \tag{24}$$

Since y^k tends to y, with $\mathcal{L}_{UV}(x^k, D^k \cup \mathcal{T}) \le 0$, we can assume without loss of generality that for all k, $y_k \notin \mathrm{Dom}\,(P)$. Then, using Assumption 2, we can find sequences $y_0^k(\cdot) \in S_G(y^k)$ and $\theta_k > 0$ such that

$$\forall k \in \mathbb{N}, \quad y_0^k(\theta_k) \in (y + \eta\mathcal{B}) \cap \mathrm{Dom}\,(P)\,, \quad \text{and} \quad \lim_{k\to\infty} \theta_k = 0.$$

Therefore, we can find a sequence $y^k(\cdot) \in S_{G,P}(y^k)$ such that

$$y^k(t) = \begin{cases} y_0^k(t) & \text{if } \ t \in [0, \theta_k) \\ p(y_0^k(\theta_k) & \text{if } \ t = \theta_k. \end{cases}$$

Now, the sets D^k are discriminating domains. Therefore, from Theorem 4.1, Victor can find a sequence of VR-strategy $\mathbf{B}_k$ such that for any $\tilde{y}^k(\cdot) \in S_{G,P}(y^k)$,

$$\forall t \le \inf\{s : (\tilde{y}^k(s), \mathbf{B}_k(\tilde{y}^k(\cdot))(s)) \in \mathcal{T}\}, \quad (\tilde{y}^k(t), \mathbf{B}_k(\tilde{y}^k(\cdot))(t)) \in D^k.$$

Let us set $z^k(\cdot) = \mathbf{B}_k(y^k(\cdot))$ for all k and let us denote by $(\tau_j^k, z_j^k, z_j^k(\cdot))_{j\in J_k}$ the associated runs. Then (23) yields that for k large enough, we have $\tau_0^k \le \theta_k$. Moreover, (24) yields that for k large enough, we have $\tau_0^k < \theta_k$, and from the last item of Assumption 1, we can assume that $\theta_k \in (\tau_0^k, \tau_1^k)$.

Now, the sequence θ_k tends to 0. Therefore, we have

$$\lim_{k\to\infty} \left\| (y^k(\theta_k)\,,\ z^k(\theta_k)) - (p(y^k(\tau_0^k), \mu), z_1^k) \right\| = 0$$

and

$$\lim_{k\to\infty} \left\| (y^k(\tau_0^{k-})\,,\ z^k(\tau_0^{k-}))\ -\ x \right\| = 0.$$

But the fact that $(y^k(\theta_k)\ ,\ z^k(\theta_k)) \in D^k$ contradicts (24). Therefore, we have proved that $\mathcal{L}_{UV}(x, D^* \cup \mathcal{T}) \leq 0$.
Proof of $\min\{\mathcal{H}(x_0, D^*)\ ,\ \mathcal{L}_V(x_0, D^* \cup \mathcal{T})\} \leq 0$

Case 1. There exists $k_1 \in \mathbb{N}$ such that

$$\forall k \geq k_1, \quad \mathcal{L}_V(x^k, D^k \cup \mathcal{T}) \leq 0.$$

By continuity of the jump function q, this yields $\mathcal{L}_V(x_0, D^* \cup \mathcal{T}) \leq 0$.

Case 2. There exists $k_2 \in \mathbb{N}$ such that

$$\forall k \geq k_2, \quad \mathcal{H}(x^k, D^k) \leq 0.$$

Let $\pi \in NP_{D^*}(x_0)$. We can assume without loss of generality that x_0 is the only projection of $(x_0 + \pi)$ onto D^*. For all $k \subset \mathbb{N}$, let us denote by $\tilde{x}^k$ the projection of $(x_0 + \pi)$ onto D^k. Then

$$\lim_{k \to +\infty} \tilde{x}^k = x_0.$$

And for all $k \in \mathbb{N}$, $(x_0 + \pi - \tilde{x}^k)$ belongs to $NP_{D^k}(\tilde{x}_k)$. Therefore, we have

$$\forall k \geq k_2, \quad \inf_{v \in V} \sup_{u \in U} \left\langle f(x, u, v), x_0 + \pi - \tilde{x}^k \right\rangle \leq 0.$$

By continuity of f and of the scalar product, we obtain $\mathcal{H}(x_0, D) \leq 0$.

Since either Case 1 or Case 2 holds true up to a subsequence, we have proved that D^* is an impulse discriminating domain for Victor with target $\mathcal{T}$.

C.3 Proofs of the Characterization Results

Lemma C.1. *Under Assumptions 1 and 2 for* (G, P)*, we have:*

- *For all* $\varepsilon > 0$ *and all time* T *there exists* $\eta > 0$*, such that for any initial condition* $\overline{y}$ *and any trajectory* $\overline{y}(\cdot) \in S_{G,P}(\overline{y})$*, we have*

  $$\forall y \in (\overline{y} + \eta\mathcal{B}),\ \exists y(\cdot) \in S_{G,P}(y) \text{ such that } \forall t \leq T,\ y(t) \in \{\overline{y}(s)\}_{s \leq t} + \varepsilon\mathcal{B}.$$

- *For any initial condition* $\overline{y}$ *and any trajectory* $\overline{y}(\cdot) \in S_{G,P}(\overline{y})$*, for any sequence of initial conditions* y^k *which tends to* $\overline{y}$*, there exists a sequence of trajectories* $y^k(\cdot) \in S_{G,P}(y^k)$ *such that*

  $$\forall t \geq 0, \quad \lim_{k \to +\infty} y^k(t) = \overline{y}(t).$$

Proof. The second item is a consequence of the first, so we only prove the first item. For this purpose, let us set $\varepsilon > 0$ and $T > 0$. From Filippov's theorem, there exists $\hat{\eta} \leq \varepsilon$ such that for all $\overline{y}$ and all $y \in (\overline{y} + \hat{\eta}\mathcal{B})$,

$$\forall \overline{y}(\cdot) \in S_G(\overline{y}), \quad \exists y(\cdot) \in S_F(y), \quad \forall t \leq T, \quad \|\overline{y}(t) - y(t)\| \leq \varepsilon. \tag{25}$$

Let us denote by k^* the maximal number of jumps on the interval $[0, T]$. Using (25) and Assumption 2, we define

$$0 < \eta_0 \leq \varepsilon_0 \leq \eta_1 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_{k^*} \leq \eta_{k^*+1} = \hat{\eta}$$

such that for all $k < k^*$, $\varepsilon_k \leq \frac{\eta_{k+1}}{2\ell_P}$ and

$$\begin{aligned}
&\forall\, \overline{y}_k \in \mathbb{R}^n, \quad \forall\, \overline{y}_k(\cdot) \in S_G(\overline{y}_k), \quad \forall y_k \in (\overline{y}_k + \eta_k \mathcal{B}), && (26)\\
&\quad \exists y_k(\cdot) \in S_G(y_k), \quad \text{such that } \forall t \leq T,\ \left\| y_k(t) - \overline{y}_k(t) \right\| \leq \varepsilon_k \mathcal{B},\\
&\forall \tilde{y}_{k-1} \in (\mathrm{Dom}\,(P) + \varepsilon_{k-1}\mathcal{B}), && (27)\\
&\quad \exists \tilde{y}_k(\cdot) \in S_G(y), \quad \exists \theta_{k-1} \leq \frac{\eta_k}{2M\ell_P}, \quad \tilde{y}_k(\theta_i) \in \mathrm{Dom}\,(P)\,.
\end{aligned}$$

Let us set $\overline{y} \in \mathbb{R}^n$, $\overline{y}(\cdot) \in S_{G,P}(\overline{y})$, and $y \in (\overline{y} + \eta_0 \mathcal{B})$. Let $(\overline{\tau}_i, \overline{y}_i, \overline{y}_i(\cdot))_{i \in I} \equiv \overline{y}(\cdot)$. Then for all $i \leq k^*$, and all $y_i \in (\overline{y}_i + \eta_i \mathcal{B})$,

$$\exists y_i(\cdot) \in S_G(y_i), \qquad \begin{cases} \forall t \leq \overline{\tau}_i, \quad \left\| y_i(t) - \overline{y}_i(t) \right\| \leq \varepsilon_i \ \leq \dfrac{\eta_{i+1}}{2\ell_P}. \\ \exists \theta_i \leq \frac{\eta_{i+1}}{2M\ell_P}, \quad y_i(\overline{\tau}_i + \theta_i) \in (\overline{y}_i(\overline{\tau}_i) + \varepsilon_i \mathcal{B}) \cap \mathrm{Dom}\,(P)\,. \end{cases}$$

So the proof is completed. □

C.3.1 Proof for Victor's Victory Domain (Theorem 4.2)

The inclusion $\mathrm{Disc}_{\mathcal{H},\mathcal{L}_V,\mathcal{L}_{UV}}(K, \mathcal{T}) \subset W_V$ stems from the very definition of the impulse discriminating kernel. In order to prove the converse inclusion, we will prove that W_V is a closed impulse discriminating kernel for Victor with target $\mathcal{T}$.

Let $x_0 \in W_V$ and let $\mathbf{B}$ be a VR-strategy at x_0 which ensures Victor's victory. Let $y(\cdot) \in S_{G,P}(y_0)$ and set $z(\cdot) = \mathbf{B}(y(\cdot))$. Then by the very definition of the strategy $\mathbf{B}$, the trajectory $x(\cdot) = (y(\cdot), z(\cdot))$ stays in K as long as the target $\mathcal{T}$ has not been reached. We claim that it actually stays in W_V. Indeed, for $T \leq \theta^K_{\mathcal{T}}(x(\cdot))$, let us denote by $y_T(\cdot) \in S_{G,P}(y(T))$ and $z_T(\cdot) \in S_{H,Q}(z(T))$ the trajectories which coincide with $y(\cdot - T)$ and $z(\cdot - T)$ respectively on $[0, +\infty)$. Then $z_T(\cdot)$ is associated with $y_T(\cdot)$ in a nonanticipative way. Furthermore,

$$\forall t \leq (\hat{T} - T), \quad (y_T(t), z_T(t)) \in K,$$

which means that $(y(T), z(T)) \in W_V$. Therefore, we have proved that

$$\forall y(\cdot) \in S_{G,P}(y_0), \quad \forall T \leq \theta^K_{\mathcal{T}}(y(\cdot), \mathbf{B}(y(\cdot))), \quad (y(T), \mathbf{B}(y(\cdot))(T)) \in W_V.$$

It remains to prove that W_V is closed.

Let us consider a sequence $\{x^k\}_{k \in \mathbb{N}} = \{(y^k, z^k)\}_{k \in \mathbb{N}}$ of W_V which tends to $x = (y, z)$. We have to exhibit a VR-strategy $\mathbf{B}$ such that for all trajectories

$y(\cdot) \in S_{G,P}(y)$, the trajectory $(y(\cdot), \mathbf{B}(y(\cdot)))$ stays in K as long as the target $\mathcal{T}$ has not been reached. For this purpose, we define the set-valued map

$$\begin{array}{lll} \mathcal{B} : S_{G,P}(y_0) & \rightsquigarrow & S_{H,Q}(z_0) \\ y(\cdot) & \mapsto & \{z(\cdot) : \forall t \le \{\inf\{s : (y(s), z(s)) \in \mathcal{T}\},\ (y(t), z(t)) \in K\}\,. \end{array}$$

Then $\mathcal{B}$ has closed values. In order to use the Cardaliaguet–Plaskacz Lemma, we have to prove that it has nonempty values and that it is nonanticipative.

$\mathcal{B}$ has nonempty values:

Let us set $y(\cdot) \in S_{G,P}(y)$. By Lemma C.1, there exists a sequence $y^k(\cdot) \in S_{G,P}(y^k)$ such that

$$\forall t \ge 0, \qquad \lim_{k\to+\infty} y^k(t) = y(t).$$

By the definition of W_V, for all $k \in \mathbb{N}$, Victor can find a VR-strategy $\mathbf{B}^k$ such that the trajectory $(y^k(\cdot), \mathbf{B}^k(y^k(\cdot)))$ satisfies

$$\forall t \le \inf\{s : (y^k(s), z^k(s)) \in \mathcal{T}\}, \quad (y^k(t), z^k(t)) \in K. \tag{28}$$

Let us set $z^k(\cdot) = \mathbf{B}^k(y^k(\cdot))$ for all $k \in K$. Then, by Proposition 2 in [8], there exists a subsequence of trajectories again denoted $\{z^k(\cdot)\}_{k\in\mathbb{N}}$ which converges to some $z(\cdot) \in S_{H,Q}(z)$. Since $\mathcal{T}$ and K are closed, passing to the limit in (28) provides

$$\forall t \le \inf\{s : (y(s), z(s) \in \mathcal{T}\}, \quad (y(t), z(t)) \in K.$$

$\mathcal{B}$ is nonanticipative:

Let us set $s \ge 0$ and let $y_1(\cdot)$ and $y_2(\cdot)$ of $S_{G,P}(y)$ coincide over $[0, s]$. Let us set $\overline{z}_1(\cdot) \in \mathcal{B}(y_1(\cdot))$. Then by the very definition of $\mathcal{B}$, $(y_1(s), \overline{z}_1(s)) \in W_V$. Therefore, there exists $\mathbf{B}^* \in \mathbb{B}_{(y_1(s),\overline{z}_1(s))}$ such that for all $y^*(\cdot) \in S_{G,P}(y_1(s), \overline{z}_1(s))$,

$$\forall t \le \inf\{s : (y^*(s), \mathbf{B}^*(y^*(\cdot))(s)) \in \mathcal{T}\}, \quad (y^*(t), \mathbf{B}^*(y^*(\cdot))(t)) \in W_V.$$

Let us set

$$\overline{z}_2(t) = \begin{cases} z_1(t) & \text{if } t \le s \\ \mathbf{B}^*(y_2(\cdot - s))(t) & \text{otherwise.} \end{cases}$$

Then $\overline{z}_2(\cdot) \in \mathcal{B}(y_2(\cdot))$ and it coincides with $\overline{z}_1(\dot{})$ over [0,s].

We have proved that $\mathcal{B}$ has a nonanticipative selection which is a VR-strategy, so $x \in W_V$.

C.3.2 Proof of Proposition 4.2

We shall first prove that W_U is open and that $(K \setminus W_U)$ is a discriminating domain for Victor, which yields that $(K \setminus W_U) \subset W_V$. Then, we shall prove that $W_V \cap W_U = \emptyset$.

W_U is open: Let us set $\overline{x} \in W_U$. We will prove that there exists $\varepsilon > 0$, $\eta > 0$, and a time $T > 0$ such that for any $(y, z) \in (\overline{x} + \eta\mathcal{B})$, there exists a VR-strategy $\mathbf{A}_{(y,z)} \in \mathbb{A}(y, z)$ such that for any $z(\cdot) \in S_{H,Q}(z)$, there exists $\theta \le T$ such that

$$(\mathbf{A}_{(y,z)}(z(\cdot))(\theta), z(\theta)) \notin K_\varepsilon \quad \text{and} \quad \forall t \le \theta, \quad (\mathbf{A}_{(y,z)}(z(\cdot))(t), z(t)) \notin \mathcal{T}_\varepsilon.$$

Since $\overline{x} = (\overline{y}, \overline{z}) \in W_U$, there exists $\varepsilon > 0$, $T > 0$, and a VR-strategy $\mathbf{A}_{\overline{x}}$ such that for any $z(\cdot) \in S_{H,Q}(\overline{z})$, there exists $\theta \leq T$ such that

$$(\mathbf{A}_{\overline{x}}(z(\cdot))(\theta), z(\theta)) \notin K_{\frac{\varepsilon}{2}} \quad \text{and} \quad \forall t \leq \theta, \quad (\mathbf{A}_{\overline{x}}(z(\cdot))(t), z(t)) \notin \mathcal{T}_{\frac{\varepsilon}{2}}.$$

Let us define the set-valued map

$$\begin{array}{rcl} \mathcal{A}_{(y,z)} : S_{H,Q}(z) & \rightsquigarrow & S_{G,P}(y) \\ z(\cdot) & \mapsto & \left\{ y(\cdot) : \exists \overline{z}(\cdot) \in S_{H,Q}(\overline{z}), \quad \forall t \leq T, \right. \\ & & \left. \quad \|(y(t), z(t)) - (\mathbf{A}_{\overline{x}}(\overline{z}(\cdot))(t) - \overline{z}(t))\| \leq \frac{\varepsilon}{2} \right\}. \end{array}$$

From Lemma C.1, this map has nonempty values in a neighborhood of $\overline{x}$; it is nonanticipative by construction. Using the Cardaliaguet–Plaskacz Lemma, we can prove that in this neighborhood there exists a single-valued selection of $\mathcal{A}_{(y,z)}$ which is a VR-strategy.

$(K \setminus W_U)$ is an impulse discriminating domain: We proceed by contradiction: we assume that $(K \setminus W_U)$ is not an impulse discriminating domain for Victor. We have to find a point $x_0 \in (K \setminus W_U)$, a VR-strategy for Ursula $\mathbf{A} \in \mathbb{A}(x_0)$, a time $T \geq 0$, and $\varepsilon > 0$ such that for all $z(\cdot) \in S_{Q,H}(z_0)$,

$$\exists t \leq \min\{T\ ,\ \inf\{s : (\mathbf{A}(z(\cdot))(s), z(s)) \in \mathcal{T}_\varepsilon\}\}, \quad (\mathbf{A}(z(\cdot))(t), z(t)) \in O_\varepsilon,$$

where $O_\varepsilon := \{x \in W_U : \forall x' \notin W_U,\ \left\|x - x'\right\| \geq \varepsilon\}$.

Since $(K \setminus W_U)$ is not an impulse discriminating domain for Victor, there exists $x_0 \in (K \setminus W_U)$ such that one of the two following propositions holds true:

$$\exists \mu \in M, \quad (p(y_0, \mu), z_0) \in W_U \quad \text{and} \quad \forall \nu \in N,\ (p(y_0, \mu), z_0) \in W_U \tag{29}$$

$$\exists \pi \in NP_D(x_0),\ \sup_{u \in U} \inf_{v \in V} \langle f(x_0, u, v), \pi \rangle > 0 \quad \text{and}$$
$$\forall \nu \in N,\ (y_0, q(z_0, \nu)) \in W_U. \tag{30}$$

We will construct a VR-strategy for Ursula in each case.

Case 1. If (29) holds true, then Ursula's best move is to jump from her initial position. Indeed, from Assumption 1, there exists $\varepsilon > 0$ such that

$$(p(y_0, \mu), z_0) \in O_\varepsilon \quad \text{and} \quad \forall \nu \in N, \quad (p(y_0, \mu), q(z_0, \nu)) \in O_\varepsilon. \tag{31}$$

Case 2. If (30) holds true, then Ursula's best choice is to avoid jumping. Indeed, there exist $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that

$$\forall (y, z) \in (y_0, z_0) + \varepsilon_1 \mathcal{B}, \quad \forall \nu \in N,\ (y, q(z, \nu)) \in O_{\varepsilon_2}.$$

On the other hand, there exists a nonimpulsive VR-strategy $\tilde{\mathbf{A}}$, a time $T \geq 0$, and $\varepsilon \leq \min\{\varepsilon_1, \varepsilon_2\}$ such that for all $\tilde{z}(\cdot) \in S_G(y_0)$,

$$\exists \theta \leq T, \quad (\tilde{\mathbf{A}}(\tilde{z}(\cdot))(\theta), \tilde{z}(\cdot)(\theta)) \in O_\varepsilon \ \text{and} \ \forall t \leq \theta,\ (\tilde{\mathbf{A}}(\tilde{z}(\cdot))(t), \tilde{z}(\cdot)(t)) \notin \mathcal{T}_\varepsilon.$$

Therefore, Ursula's strategy will be a nonanticipative selection[1] of the set-valued application with nonempty values

$$\begin{array}{llll} \mathcal{A}: & S_{H,Q}(z_0) & \rightsquigarrow & S_{G,P}(y_0) \\ & z(\cdot) \equiv (\tau_j^z, z_j, z_j(\cdot))_{j\in J} & \mapsto & \left\{ y(\cdot) \equiv (\tau_i^y, z_i, z_i(\cdot))_{i\in I} : y_0 = \widetilde{\mathbf{A}}(z_0(\cdot)) \text{ and } \tau_0^y \geq \tau_0^z \right\}. \end{array}$$

Both cases contradict the fact that $x_0 \notin W_U$, so the proof is complete.

$\boldsymbol{W_V \cap W_U = \emptyset}$**:** Let us set $x_0 \in W_V$, $\mathbf{A} \in \mathbb{A}(x_0)$, $T \geq 0$, and $\varepsilon > 0$. We have to prove that there exists $z(\cdot) \in S_{H,Q}(z_0)$ such that the trajectory $(\mathbf{A}(z(\cdot)), z(\cdot))$ stays in K_ε up to time T or as long as $\mathcal{T}_\varepsilon$ has not been reached.

We recall from the last item of Assumption 1, that a trajectory $y(\cdot) \in S_{G,P}(y_0)$ or $z(\cdot) \in S_{H,Q}(z_0)$ can only have a limited number of jumps on the interval $[0, T]$. Let us denote this number by k^*. We shall construct a run $(\tau_j^z, z_j, z_j(\cdot))_{j\leq j^*}$, with $j^* \leq k^*$, for which there exists a family $\{\varepsilon_j\}_{j\in J}$, with

$$0 < \varepsilon_0 \leq \cdots \leq \varepsilon_{j^*} = \varepsilon,$$

such that $\sum_{j<j*} \tau_j \leq T$ and for all $j \leq j^*$,

$$\forall t \leq \min\{\tau_j, \inf\{s : (y^j(s), z_j(s)) \in \mathcal{T}_{\varepsilon_j}\}\}, \qquad (y^j(s), z_j(s)) \in K_{\varepsilon_j}, \tag{32}$$

where $y^j(\cdot)$ coincides with $\mathbf{A}(z(\cdot))$ over $[\sum_{j'<j} \tau_j, \sum_{j'\leq j} \tau_j)$.

For this purpose, we will need the following result (proved later), which states sufficient conditions for Victor to be able to keep the state close to W_V, without jumping, whatever Ursula plays: Let us set

$$\widehat{C} := \{(y, z) \in W_V : \exists \nu \in N \text{ such that } (\{y\} \cup P(y)) \times \{q(z, \nu)\} \subset W_V\}.$$

Lemma C.2. *For all $\varepsilon > 0$, there exists $\eta > 0$ such that for all $(y, z) \in (W_V + \eta\mathcal{B})$, there exists $\tilde{z}(\cdot) \in S_H(z)$ such that for all $\tilde{y}(\cdot) \in S_G(y)$,*

$$\forall t \leq \min\{\, T\,,\ \inf\{s : (\tilde{y}(s), \tilde{z}(s)) \in (\mathcal{T}_\varepsilon \cup C_\varepsilon)\,\}\,\}, \quad (\tilde{y}(t), \tilde{z}(t)) \in (W_V + \varepsilon\mathcal{B}),$$

with $C_\varepsilon := \{(y, z) \in \widehat{C} + \varepsilon\mathcal{B} : z \in \mathrm{Dom}\,(Q)\}$.

Using Lemma C.2, the fact that the reset map p is Lipchitz continuous with respect to its first variable, and the fact that Ursula can jump only k^* times on any interval $[0, s]$ with $s \leq T$, we can prove that for all $\tilde{\varepsilon} > 0$, there exists $\eta > 0$ such that for all $(y, z) \in (W_V + \eta\mathcal{B})$, there exists $\tilde{z}(\cdot) \in S_H(z)$ such that for all $y(\cdot) \in S_{G,P}(y)$,

$$\forall t \leq \min\{\, T\,,\ \inf\{s : (\tilde{y}(s), \tilde{z}(s)) \in (\mathcal{T}_{\tilde{\varepsilon}} \cup C_{\tilde{\varepsilon}})\,\}\,\}, \quad (y(t), \tilde{z}(t)) \in (W_V + \tilde{\varepsilon}\mathcal{B}).$$

[1]We can prove that nonanticipative single-valued selections exist using the Cardaliaguet–Plaskacz Lemma.

Therefore, using the Lipschitz continuity of the jump functions, we can construct a finite nondecreasing sequence

$$\{\ (0 = \varepsilon_0)\ , \varepsilon_1, \ldots,\ (\varepsilon_{2k^*+1} = \varepsilon)\ \}$$

such that, for all $k \le k^*$,

$$\forall (y_k, z_k) \in ((W_V + \varepsilon_{2k}\mathcal{B}) \setminus \mathcal{T}_\varepsilon), \quad \exists \tilde{z}_k(\cdot) \in S_H(z_k) \text{ such that } \forall \tilde{y}_k(\cdot) \in S_G(y_k),$$
$$\forall t \le \min\Big\{\ T\ ,\ \inf\{s : (\tilde{y}_k(s), \tilde{z}_k(s)) \in (\mathcal{T}_\varepsilon \cup C_{\varepsilon_{2k+1}})\ \}\ \Big\},$$

$$(\tilde{y}_k(t), \tilde{z}_k(t)) \in (W_V + \varepsilon_{2k+1}\mathcal{B})$$

$$\forall \xi_k \in (W_V + \varepsilon_{2k+1}\mathcal{B}) \setminus \mathcal{T}_\varepsilon, \quad \forall \mu \in M, (p(y,\mu), z) \in ((W_V + \varepsilon_{2k+2}\mathcal{B}) \cup \mathcal{T}_\varepsilon)$$
$$\forall \xi_k \in C_{\varepsilon_{2k+1}} \setminus \mathcal{T}_{\hat{\varepsilon}},\ \exists \nu_k \in N,$$
$$\forall \mu \in M, (p(y,\mu), q(z,\nu_k)) \in ((W_V + \varepsilon_{2k+2}\mathcal{B}) \cup \mathcal{T}_\varepsilon).$$

Now, if we assume[2] that $x_0 \notin \widehat{C}$ and that $y_0 \notin \mathrm{Dom}\,(P)$, we can construct the run which satisfies (32) by iteration.

Proof of Lemma C.2 From Assumption 2, there exists a constant $\tilde{\varepsilon} \le \frac{\varepsilon}{2}$ such that

$$\forall z_0 \in (\mathrm{Dom}\,(Q) + \tilde{\varepsilon}\mathcal{B}),\ \exists z(\cdot) \in S_H(z_0),\ \exists \theta \le \frac{\varepsilon}{2M} \text{ such that } z(\theta) \in \mathrm{Dom}\,(Q)\,, \tag{33}$$

where M denotes a bound of H on $\mathrm{Dom}\,(Q) + \varepsilon\mathcal{B}$. Let us denote by ℓ a Lipschitz constant of f, and let $\eta > 0$ be such that

$$\eta\ \mathrm{e}^{\ell T} \le \frac{\tilde{\varepsilon}}{2}$$

and let us set $\overline{x} \in (D + \eta\mathcal{B})$. We assume that $\overline{x} \notin \mathcal{T}_\varepsilon$ (the result is obvious otherwise). Let us denote by (y_0, z_0) a projection of $\overline{x}$ onto D.

Assumption 1 yields that $\widehat{C}$ is closed. The fact that W_V is an impulse discriminating domain and the conditions of Proposition 4.2 yield

$$\forall x \in W_V \setminus (\mathcal{T} \cup \widehat{C}), \quad \forall \pi \in NP_{W_V}(x), \quad \sup_{u \in U} \inf_{v \in V} \langle f(x,u,v), \pi \rangle \le 0. \tag{34}$$

From Proposition B.1, there exists $\tilde{z}(\cdot) \in S_H(z_0)$ such that for all $\tilde{y}(\cdot) \in S_G(y_0)$,

$$\forall t \le \min\{\ T\ ,\ \inf\{s : (\tilde{y}(s), \tilde{z}(s)) \in (\mathcal{T}_{\frac{\tilde{\varepsilon}}{2}} \cup C_{\frac{\tilde{\varepsilon}}{2}})\}\ \},\ (\tilde{y}(t), \tilde{z}(t)) \in (W_V + \frac{\tilde{\varepsilon}}{2}\mathcal{B}).$$

□

[2]If $x_0 \in \mathrm{Dom}\,(R_U)$, we have $R_U(x_0) \subset (D \cup \mathcal{T})$, so we can use the same proof with $\hat{x}_0 \in R_U(x_0)$. If $x_0 \in C$, we can find $\nu \in N$ such that $R_{VU}(x_0, \nu) \subset (D \cup \mathcal{T})$, so we can use the same proof with $\hat{x}_0 \in R_{VU}(x_0, \nu)$.

Now let $\overline{y}(\cdot) \in S_G(\overline{y})$. From Filippov's theorem,

$$\begin{aligned} &\exists \tilde{z}(\cdot) \in S_H(\overline{z}), && \forall t \geq 0, \quad \|\tilde{z}(t) - \overline{z}(t)\| \leq \|z_0 - \overline{z}\| \, \mathrm{e}^{\ell t}, \\ &\exists \tilde{y}(\cdot) \in S_G(y_0), && \forall t \geq 0, \quad \|\tilde{y}(t) - \overline{y}(t)\| \leq \|y_0 - \overline{z}\| \, \mathrm{e}^{\ell t}. \end{aligned}$$

Then, for t small enough, we have

$$\begin{aligned} d_{W_V}\,(\overline{y}(t), \overline{z}(t)) &\leq \|(\overline{y}(t), \overline{z}(t)) - (\tilde{y}(t), \tilde{z}(t))\| && + d_{W_V}\,(\tilde{y}(t), \tilde{z}(t)) \\ &\leq \|(y_0, z_0) - (\overline{y}, \overline{z})\| \, \mathrm{e}^{\ell t} && + \frac{\tilde{\varepsilon}}{2} \\ &\leq \frac{\tilde{\varepsilon}}{2} && + \frac{\tilde{\varepsilon}}{2} \end{aligned}$$

So we have proved that one of the two following propositions holds true:

$$\forall t \leq \min\{T, \inf\{s\colon (\overline{y}(s), \overline{z}(s)) \in \mathcal{T}_{\tilde{\varepsilon}}\}\}, \quad (\overline{y}(t), \overline{z}(t)) \in (W_V + \tilde{\varepsilon}\mathcal{B}) \tag{35}$$

$$\forall t \leq \min\{T, \inf\{s\colon (\overline{y}(s), \overline{z}(s)) \in C_{\tilde{\varepsilon}}\}\}, \quad (\overline{y}(t), \overline{z}(t)) \in (W_V + \tilde{\varepsilon}\mathcal{B}). \tag{36}$$

If (35) holds true, the proof is complete since $\tilde{\varepsilon} \leq \varepsilon$. If not, we use (33) to conclude. Indeed, let us assume that

$$\tilde{T} := \inf\{s\colon (\overline{y}(s), \overline{z}(s)) \in C_{\tilde{\varepsilon}}\} < T.$$

Then we have

$$\overline{z}(\tilde{T}) \in C_{\tilde{\varepsilon}} \quad \subset \mathbb{R}^l \times (\mathrm{Dom}\,(Q) + \tilde{\varepsilon}\mathcal{B}).$$

Then

$$\exists z_{\tilde{T}}(\cdot) \in S_H(\overline{z}(T)), \quad \theta \leq \frac{\varepsilon}{2M}, \quad z(\theta) \in \mathrm{Dom}\,(Q)\,.$$

Let us set

$$\hat{z}(t) = \begin{cases} \overline{z}(t) & \text{if } t \leq \tilde{T}, \\ z(t - \tilde{T}) & \text{if } t > \tilde{T}. \end{cases}$$

Then

$$\begin{aligned} \forall t \leq \theta, \quad d_{\widehat{C}}\left(\overline{y}(\tilde{T} + t), \hat{z}(\tilde{T} + t)\right) &\leq d_{\widehat{C}}\left(\overline{y}(\tilde{T}), \hat{z}(\tilde{T})\right) + Mt \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}, \end{aligned}$$

which concludes the proof. □

References

[1] J.-P. Aubin. *Viability Theory*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, 1991.

[2] J.-P. Aubin, J. Lygeros, M. Quincampoix, S. Sastry, and N. Seube. Impulse Differential Inclusions: a Viability Approach to Hybrid Systems. *IEEE Transactions on Automatic Control*, 47(1):2–20, 2002.

[3] P. Cardaliaguet. A Differential Game with Two Players and One Target. *SIAM Journal of Control and Optimization*, 34(4):1441–1460, 1996.

[4] P. Cardaliaguet and S. Plaskacz. Invariant Solutions of Differential Games and Hamilton–Jacobi Isaacs Equations for Time-Measurable Hamiltonians. *SIAM Journal of Control and Optimization*, 38(5):1501–1520, 2000.

[5] P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre. Numerical Methods for Differential Games. In M. Bardi, T.E.S. Raghavan, and T. Parthasarathy, editors, *Stochastic and Differential Games: Theory and Numerical Methods*, Annals of the International Society of Dynamic Games, pages 177–247. Birkhäuser, 1999.

[6] P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre. Pursuit differential games with state constraints. *SIAM Journal of Control and Optimization*, 39(5):1615–1632, 2001.

[7] E. Crück. Target Problems under State Constraints for Nonlinear Controlled Impulsive Systems. *Journal of Mathematical Analysis and Applications*, 270(2):636–656, 2002.

[8] E. Crück, and P. Saint-Pierre. Nonlinear impulse target problems under state constraint: A Numerical analysis based on viability theory. *Set-Valued Analysis*, 12:383–416, 2004.

[9] R.J. Elliott and N.J. Kalton. The Existence of Value in Differential Games of Pursuit and Evasion. *Journal of Differential Equations*, (12):504–523, 1972.

[10] H. Frankowska. Lower Semicontinuous Solutions of Hamilton-Jacobi-Bellman Equations. *SIAM Journal of Control and Optimization*, 31(1): 257–272, 1993.

[11] R. Isaacs. *Differential Games*. John Wiley & Sons, New York, 1965.

[12] E. Roxin. Stability in General Control Systems. *Journal of Differential Equations*, 1:115–150, 1965.

[13] E. Roxin. Axiomatic Approach in Differential Games. *Journal of Optimization Theory and Applications*, 3(3):153–163, 1969.

[14] P. Varaiya. On the Existence of Solutions to a Differential Game. *SIAM Journal of Control*, 5(1):153–162, 1967.

[15] P. Varaiya. Differential Games with Dynamical Systems. In H.W. Kuhn and G.P. Szeg, editors, *Differential Games and Related Topics*, pages 129–144. North-Holland Publishing Company, Amsterdam, 1971.

Approaching Coalitions of Evaders on the Average

Igor Shevchenko
Pacific Fisheries Research Center
4 Shevchenko Alley
Vladivostok, 690950, Russia
igor@tinro.ru
and
Informatics Department
Far Eastern State University
8 Sukhanova Street
Vladivostok, 690950, Russia
shevch@dvgu.ru

Abstract

In the game Φ_N with simple motions, the pursuer P and the coalition $E^N = \{E_1, E_2, \ldots, E_N\}$ of evaders move in a plane with constant speeds 1, $\beta_1, \beta_2, \ldots, \beta_N$. The average distance from a point to a set of points is defined as a weighted sum of the corresponding Euclidean distances with given positive constant weights. P strives to minimize the distance to E^N and terminates the game when the distance shortening is not guaranteed.

First, we describe several conditions that are met by the states on the terminal manifold M_{Φ_N} of Φ_N depending on the index of evaders caught there. Then, we study Φ_2 in detail. This game is a game of alternative pursuit since there are three different terminal sub-manifolds: P catches E_1 (E_2) on $\mathsf{M}^{\{1\}}_{\Phi_2}$ ($\mathsf{M}^{\{2\}}_{\Phi_2}$) and all players are apart on $\mathsf{M}^{\emptyset}_{\Phi_2}$. We set up and study associated games $\Phi_2^{\{1\}}$ ($\Phi_2^{\{2\}}$) and $\Phi_2^{\emptyset}$ with the payoffs equal to the average distance to E^2 at instants when the state reaches $\mathsf{M}^{\{1\}}_{\Phi_2}$ ($\mathsf{M}^{\{2\}}_{\Phi_2}$) and $\mathsf{M}^{\emptyset}_{\Phi_2}$ correspondingly. It is shown that Φ_2 is strategically equivalent to the associated game with the minimal value.

1 Introduction

In this chapter we describe and study a simple game Φ_N where the pursuer P approaches the coalition of several evaders $E^N = \{E_1, E_2, \ldots, E_N\}$ as a whole and the payoff is equal to the average distance to E^N at some instant chosen by P. In the literature, several similar games were investigated for the various kinematics equations and restrictions imposed on the control variables, see, e.g., [1], [2], [8], [10], [11], [12], [13], [14]. In most of these works, the average distance is represented as a sum of the corresponding Euclidean distances to evaders with given constant weights. Those weights may reflect the relative importance of particular evaders for the pursuer or probabilities for them to be a true target [12]. In [12] and [13], the case when the weights are inversely related to the current Euclidean distances to evaders is also mentioned. In [11], arbitrary positive functions of time are allowed as the weights.

The chapter is organized as follows. First, we give a description of conditions satisfied on the terminal manifold of Φ_N. These conditions are specified for $N = 2$. If the average velocity of the coalition is less than the difference of weights, P terminates the game Φ_2 by a point capture of the most valuable evader. Otherwise, there are several sub-manifolds including those where a point capture of one evader occurs. Then, we set up and study associated games corresponding to the fixed terminal sub-manifolds and treat Φ_2 as a game of alternative pursuit. It is shown that Φ_2 is strategically equivalent to the associated game with the minimal value.

2 Setup of Φ-Games

Let $N \geq 2$ and B^N be the set of points $\{\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_N\}$ in the plane. Let $\bar{p} = (p_1, p_2, \ldots, p_N)$ be the vector of constant weights, where

$$\begin{gathered} p_1 \geq p_2 \geq \cdots \geq p_N > 0, \\ \sum_{k \in K} p_k = 1, \quad K = \{1, 2, \ldots, N\}. \end{gathered} \tag{1}$$

Define the distance ρ from pursuer P located at $\bar{a}$ to the coalition of evaders $E^N = \{E_1, E_2, \ldots, E_N\}$ located at B^N as

$$\rho(P, E^N) = \sum_{k \in K} p_k |\bar{b}_k - \bar{a}|. \tag{2}$$

Components of $\bar{p}$ may reflect the additive "importance" of the corresponding evaders to P. Also (see, e.g., [12], [13]) if there is only one specific member ("true target") of the coalition, p_k may reflect *a priori* the probability for E_k to be precisely that true target. Then (2) represents the mean value of Euclidean distance from P to the true target.

Let $\bar{z}_e = (\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_N)$ be the coordinates of E^N and $\bar{z} = (\bar{z}_p, \bar{z}_1, \bar{z}_2, \ldots, \bar{z}_N)$ be the state. Let $\bar{z}_p^0$ and $\bar{z}_e^0 = (\bar{z}_1^0, \bar{z}_2^0, \ldots, \bar{z}_N^0)$ be initial positions of the players. Let the players move in the plane obeying the system of equations

$$\dot{\bar{z}}_i = \bar{u}_i, \quad \bar{z}_i|_{t=0} = \bar{z}_i^0, \quad i \in I = \{p, e\}, \tag{3}$$

where $\bar{u}_p$ and $\bar{u}_e = (\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_N)$ are the control variables such that

$$|\bar{u}_p| \leq 1, \; |\bar{u}_k| \leq \beta_k < 1, \; k \in K. \tag{4}$$

Let $\bar{z}(t, \bar{z}^0, \mathcal{S}_p, \mathcal{S}_e)$ be a solution of (3) for chosen strategies $\mathcal{S}_i, \; i \in I$. Let ρ be the average distance to E^N at $\bar{z}$ (see (1) and (2))

$$\rho(\bar{z}) = \sum_{k \in K} p_k |\bar{z}_k - \bar{z}_p|. \tag{5}$$

Let $\mathcal{S}_p^c(\bar{u}) = \bar{u}/|\bar{u}|$ and $\mathcal{S}_k^c(\bar{u}) = \beta_k \bar{u}/|\bar{u}|$, $k \in K$. Let M_{Φ_N} be the manifold where for every $\bar{z}^M$, a small enough $\varepsilon > 0$, and a constant vector $\bar{u}_p$, $|\bar{u}_p| \leq 1$, there exist a constant vector $\bar{u}_e = (\bar{u}_1, \bar{u}_2, \dots, \bar{u}_N)$, $|\bar{u}_k| \leq \beta_k$, $k \in K$, such that

$$\rho(\bar{z}(\varepsilon, \bar{z}^M, \mathcal{S}_p^c(\bar{u}_p), \mathcal{S}_e^c(\bar{u}_e))) > \rho(\bar{z}^M). \tag{6}$$

This condition means that if P keeps playing at any state $\bar{z}^M \in \mathsf{M}_{\Phi_N}$, E^N can increase the average distance by the instant $t = \varepsilon > 0$ without fail.

Let $T_{\Phi_N} = T_{\Phi_N}(\bar{z}^0, \mathcal{S}_p, \mathcal{S}_e)$ be the first instant when the corresponding solution of (3) meets M_{Φ_N},

$$T_{\Phi_N} = \min\{t :\ \bar{z}(t, \bar{z}^0, \mathcal{S}_p, \mathcal{S}_e) \in \mathsf{M}_{\Phi_N}\}. \tag{7}$$

Denote as Φ_N the game with the playing space $\mathsf{Z}_{\Phi_N} = \mathbb{R}^{2N+2}$, the terminal manifold M_{Φ_N} and the payoff

$$\mathcal{P}_{\Phi_N}(\bar{z}^0, \mathcal{S}_p, \mathcal{S}_e) = \rho(\bar{z}(T_{\Phi_N}, \bar{z}^0, \mathcal{S}_p, \mathcal{S}_e)). \tag{8}$$

Describe several conditions satisfied by the states on M_{Φ_N}. Let

$$\beta = \sum_{k \in K} \beta_k p_k \tag{9}$$

be the average speed of E^N, $\mathsf{M}_{\Phi_N}^{\emptyset}$ be the part of M_{Φ_N} where $\forall k \in K :\ \bar{z}_k \neq \bar{z}_p$, and $\bar{e}_k$ be the unit vector

$$\bar{e}_k = \frac{\bar{z}_k - \bar{z}_p}{|\bar{z}_k - \bar{z}_p|}, \quad k \in K. \tag{10}$$

Proposition 1. *If $\bar{z} \in \mathsf{M}_{\Phi_N}^{\emptyset}$ then (see (10))*

$$\left| \sum_{k \in K} p_k \bar{e}_k \right| \leq \beta. \tag{11}$$

Proof. Let $\bar{z} \in \mathsf{M}_{\Phi_N}^{\emptyset}$ at $t = 0$ and P move at the angle φ_1 for the period $\varepsilon > 0$. By the instant $t = \varepsilon$, E^N can ensure the average distance equal to

$$\beta\varepsilon + \sum_{k \in K} p_k |\bar{z}_k - \bar{z}_p - \varepsilon \bar{e}(\varphi_1)|. \tag{12}$$

Thus, (11) is equivalent to the condition

$$\sum_{k \in K} p_k \frac{\partial |\bar{z}_k - \bar{z}_p|}{\partial \bar{e}(\varphi_1)} \leq \beta, \tag{13}$$

or

$$\sum_{k \in K} p_k \bar{e}_k \cdot \bar{e}(\varphi_1) \leq \beta. \tag{14}$$

However, even with the best choice,

$$\bar{e}(\varphi_1) = \sum_{k\in K} p_k \bar{e}_k \Big/ \Big|\sum_{k\in K} p_k \bar{e}_k\Big|, \tag{15}$$

P cannot shorten or preserve the initial average distance. □

Similarly, let $J \subset K$ and $\mathsf{M}^J_{\Phi_N}$ be the part of M_{Φ_N} where $\forall k \in K\backslash J : \ \bar{z}_k \neq \bar{z}_p$ and $\forall j \in J : \ \bar{z}_j = \bar{z}_p$.

Proposition 2. *If $\bar{z} \in \mathsf{M}^J_{\Phi_N}$ then (see (10))*

$$\Big|\sum_{k\in K\backslash J} p_k \bar{e}_k\Big| \leq \beta + \sum_{j\in J} p_j. \tag{16}$$

Using Propositions 1 and 2 for particular values of N, we can get more detailed descriptions of M_{Φ_N} under additional conditions imposed on the parameters of Φ_N. Let us proceed with Φ_2 in this direction; see also [1], [12], [13].

Let along with (1)

$$p_1 > p_2. \tag{17}$$

According to (11), P has to terminate Φ_2 only on $\mathsf{M}^{\{1,2\}}_{\Phi_2}$, $\mathsf{M}^{\{1\}}_{\Phi_2}$, $\mathsf{M}^{\{2\}}_{\Phi_2}$ or $\mathsf{M}^{\emptyset}_{\Phi_2}$. Since $\mathsf{M}^{\{1,2\}}_{\Phi_2}$ cannot be reached from the outside, exclude this sub-manifold from further consideration. First, it easily follows that for any parameters of the game

$$\mathsf{M}^{\{1\}}_{\Phi_2} \neq \emptyset. \tag{18}$$

In fact, for $J = \{1\}$, (16) is equivalent to $p_2 \leq \beta + p_1$, which in turn is valid due to (1) and (17). Second, since for $J = \{2\}$, (16) is equivalent to

$$\beta \geq p_1 - p_2, \tag{19}$$

the condition

$$\mathsf{M}^{\{2\}}_{\Phi_2} \neq \emptyset \tag{20}$$

holds only if the game parameters meet (19).

Let $\bar{z} \in \mathsf{M}^{\emptyset}_{\Phi_N}$ and (see (10))

$$\bar{e}_k = \bar{e}(\gamma_k), \ \ k \in K. \tag{21}$$

For $N = 2$, (11) means that

$$\cos(\gamma_1 - \gamma_2) \leq \frac{\beta^2 - p_1^2 - p_2^2}{2p_1p_2}. \tag{22}$$

If

$$\beta < p_1 - p_2, \tag{23}$$

then the right-hand part of (22) is less than minus one. Hence, (22) is never satisfied. Therefore, in this case,

$$\mathsf{M}_{\Phi_2} = \mathsf{M}_{\Phi_2}^{\{1\}}. \tag{24}$$

Finally, P terminates the game definitely on $\mathsf{M}_{\Phi_2}^{\{1\}}$ if the parameters meet (23), and on one of three manifolds $\mathsf{M}_{\Phi_2}^{\{1\}}$, $\mathsf{M}_{\Phi_2}^{\{2\}}$ or $\mathsf{M}_{\Phi_2}^{\emptyset}$ if (19) holds. In the latter case, Φ_2 becomes a game of *alternative pursuit*; see, e.g., [16]. Let $\Phi_2^{\{1\}}$, $\Phi_2^{\{2\}}$ and $\Phi_2^{\emptyset}$ be three associated games with the payoffs equal to the average distance to E^2 at the first instant when the state reaches *fixed* manifolds $\mathsf{M}_{\Phi_2}^{\{1\}}$, $\mathsf{M}_{\Phi_2}^{\{2\}}$ and $\mathsf{M}_{\Phi_2}^{\emptyset}$ correspondingly.

Remark 1. For $N = 2$ and $p_1 = p_2 = 1/2$, rewrite (22) as

$$\cos(\gamma_1 - \gamma_2) \leq \frac{(\beta_1 + \beta_2)^2 - 2}{2}. \tag{25}$$

If in addition $\beta_1 = \beta_2 = \beta$, then on $\mathsf{M}_{\Phi_2}^{\emptyset}$ we have

$$\cos(\gamma_1 - \gamma_2) \leq 2\beta^2 - 1. \tag{26}$$

This condition is also met on the terminal manifold of the game Λ_2, where the distance to E^2 is defined as the Euclidean distance to the further evader [15].

3 Pursuit with Terminal Manifolds $\mathsf{M}_{\Phi_2}^{\{1\}}$ and $\mathsf{M}_{\Phi_2}^{\{2\}}$

Consider $\Phi_2^{\{j\}}$ in the playing space $\mathsf{Z}_{\Phi_2}^{\{j\}} = \mathsf{Z}_{\Phi_2} \backslash \mathsf{M}_{\Phi_2}^{\{j\}}$, $j \in J$. It is evident that the following relation of strategic equivalence holds:

$$\Phi_2^{\{2\}}(1, \beta_1, \beta_2, p_1, p_2) \underset{\{(P,P),(E_1,E_2),(E_2,E_1)\}}{\overset{(\mathsf{Z}_{\Phi_2}^{\{1\}}, \bar{g}^{1\rightleftarrows 2})}{\sim}} \Phi_2^{\{1\}}(1, \beta_2, \beta_1, p_2, p_1),$$

where $\bar{g}^{1\rightleftarrows 2} : (\bar{z}_p, \bar{z}_1, \bar{z}_2) \to (\bar{z}_p, \bar{z}_2, \bar{z}_1)$.

Let $\Delta_{j,3-j}(1, \beta_1, \beta_2, R_j)$ be the game with the payoff equal to the Euclidian distance to E_{3-j} at an instant of E_j's capture with radius R_j [17]. Let $\Delta_{j,3-j}^0(1, \beta_1, \beta_2)$ be the game $\Delta_{j,3-j}(1, \beta_1, \beta_2, R_j)$ for $R_j = 0$, $j \in J$. Obviously,

$$\Phi_2^{\{j\}}(1, \beta_1, \beta_2, p_1, p_2) \underset{\{(P,P),(E_1,E_1),(E_2,E_2)\}}{\overset{(\mathsf{Z}_{\Phi_2}^{\{j\}}, \bar{g}^{\equiv})}{\sim}} \Delta_{j,3-j}^0(1, \beta_1, \beta_2), \quad j \in J,$$

where $\bar{g}^{\equiv} : \bar{z} \to \bar{z}$. Hence, to get a solution of $\Phi_2^{\{j\}}$, $j \in J$, it is sufficient to study $\Delta_{1,2}^0$ for various parameters. In turn, a solution of this game may be obtained as the limiting case as $R_1 \to 0+$. However, since in $\Delta_{1,2}(1, \beta_1, \beta_2, R_1)$ for $R_1 = 0$, the independent variable α identically equals zero, it is necessary to reformulate all results from [17] using, e.g., ψ_1 instead of α.

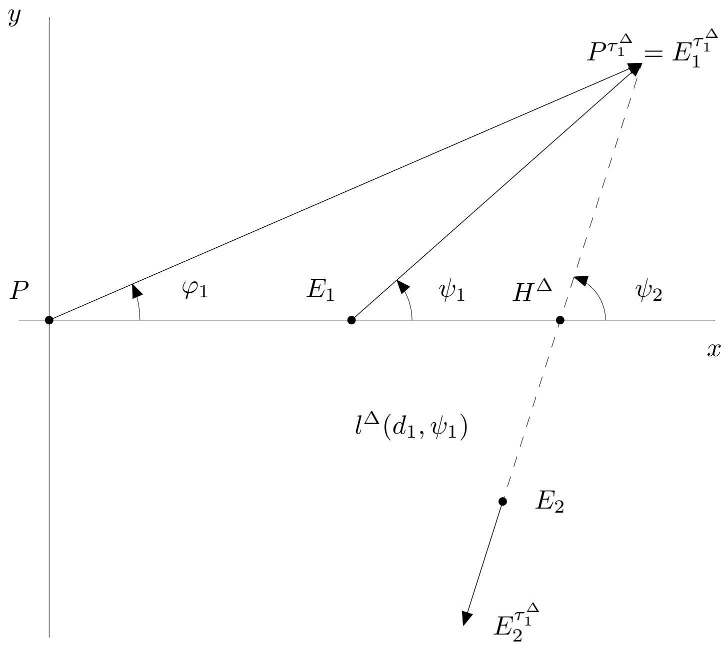


Figure 1: Geometry of optimal pursuit in $\Delta^0_{1,2}$.

Let P be at the origin O and let E_1 lie on the positive x-axis. Let φ_1, ψ_1, $\psi_2+\pi$ be the angles made by the optimal directions of players with the x-axis; see Figure 1. Let

$$\begin{aligned} \varphi_1^\Delta(\psi_1) &= \arcsin(\beta_1 \sin\psi_1), \\ \varphi_2^\Delta(\psi_1) &= \arcsin(\beta_2 \sin\psi_1), \\ \psi_2^\Delta(\psi_1) &= \psi_1 + \varphi_1^\Delta + \varphi_2^\Delta, \\ \tau_1^\Delta(d_1, \psi_1) &= d_1 \sin\psi_1 / \sin(\psi_1 - \varphi_1^\Delta), \\ h^\Delta(d_1, \psi_1) &= \tau_1^\Delta \sin(\psi_2^\Delta - \varphi_1^\Delta)/\sin\psi_2^\Delta, \end{aligned} \tag{27}$$

and

$$\Psi_1^\Delta = \{|\psi_1| < \psi_1^\Delta\}, \tag{28}$$

where ψ_1^Δ is the smallest positive root of the equation

$$\psi_2^\Delta(\psi_1) = \pi. \tag{29}$$

Let

$$h_0^\Delta(d_1) = \lim_{\psi_1 \to 0+} h^\Delta(d_1, \psi_1). \tag{30}$$

Obviously,

$$h_0^\Delta(d_1) = d_1(1+\beta_2)/(1-\beta_1)/(1+\beta_1+\beta_2). \tag{31}$$

By $\mathcal{L}^\Delta$ denote the mapping that takes each (d_1, ψ_1) to the ray l^Δ with vertex $H^\Delta = h^\Delta \bar{e}(0)$ and inclination $\psi_2^\Delta + \pi$.

Proposition 3. *For any given $d_1 > 0$, the mapping $\mathcal{L}^\Delta$ is such that*

- *every point of the lower half-plane lies on a single ray $l^\Delta(d_1, \psi_1)$ with $0 < \psi_1 < \psi_1^\Delta$,*
- *every point of the upper half-plane lies on a single ray $l^\Delta(d_1, \psi_1)$ with $-\psi_1^\Delta < \psi_1 < 0$,*
- *every point on the x-axis right of H_0^Δ lies simultaneously on two rays from the family*

$$\{l^\Delta(d_1, \psi_1)\}_{\psi_1 \in \Psi_1^\Delta} \tag{32}$$

  with the values of ψ_1 that differ only by their signs,
- *every point of the x-axis left of H_0^Δ lies on the ray $l^\Delta(d_1, 0)$.*

Corollary 1. *For any given $d_1 > 0$, the mapping $\mathcal{L}^\Delta$ is such that the rays (32) do not intersect each other and cover the plane of possible E_2 positions completely excluding the part of the x-axis right of H_0^Δ.*

Proposition 4. *Let (d_1, d_2, γ) describe $\bar{z}$ in the reduced space and $\psi_1 \in \Psi_1^\Delta$ be chosen such that*

$$E_2 \in l^\Delta(d_1, \psi_1). \tag{33}$$

Then, in $\Delta_{1,2}^0$ (see (27))

- *the optimal controls of P, E_1 and E_2 in the reduced space are defined by the angles φ_1^Δ, ψ_1 and $\psi_2^\Delta + \pi$,*
- *the optimal duration $\tau_{\Delta_{1,2}^0}(\bar{z})$ equals τ_1^Δ,*
- *the value $V_{\Delta_{1,2}^0}(\bar{z})$ is represented by*

$$U_{\Delta_{1,2}^0}(d_1, d_2, \gamma, \psi_1) = d_2 \eta_{\Delta_{1,2}^0}(d_1, d_2, \gamma, \psi_1), \tag{34}$$

  where

$$\eta_{\Delta_{1,2}^0}(d_1, d_2, \gamma, \psi_1) = \sin(\psi_2^\Delta - \psi_1 - \gamma)/\sin\psi_1, \tag{35}$$

- *the solution is regular with the exception of the dispersal surface that is projected into the part of x-axis right of H_0^Δ; see Figure 2.*

Proposition 5. *At any regular state of $\Delta_{1,2}^0$ with reduced coordinates (d_1, d_2, γ) there exist a unique ψ_1 such that (33) is satisfied,*

$$\operatorname{sign}\psi_1 = -\operatorname{sign}\gamma, \tag{36}$$

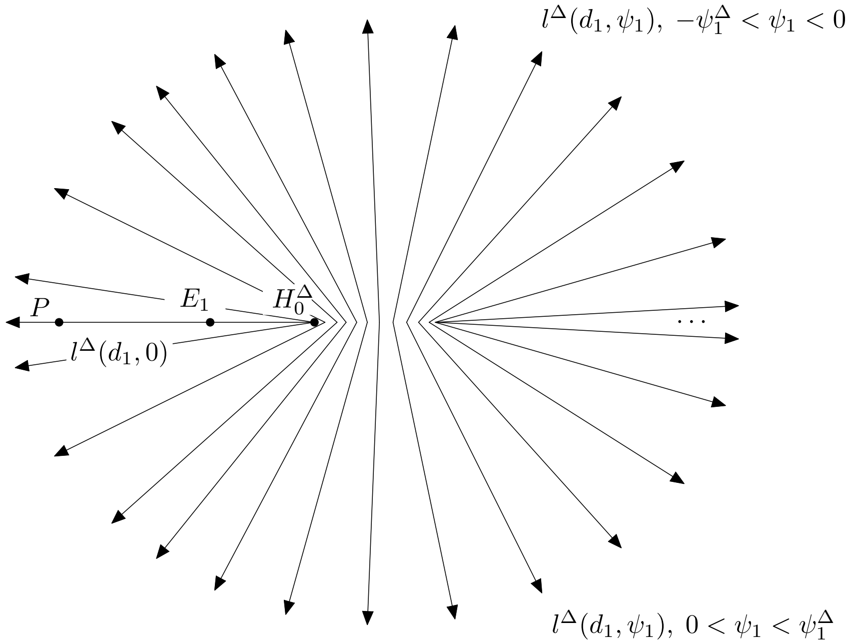


Figure 2: Field of E_2's optimal trajectories in $\Delta^0_{1,2}$.

and

$$\mathcal{G}_{\Delta^0_{1,2}}(d_1, d_2, \gamma, \psi_1) = 0, \tag{37}$$

where

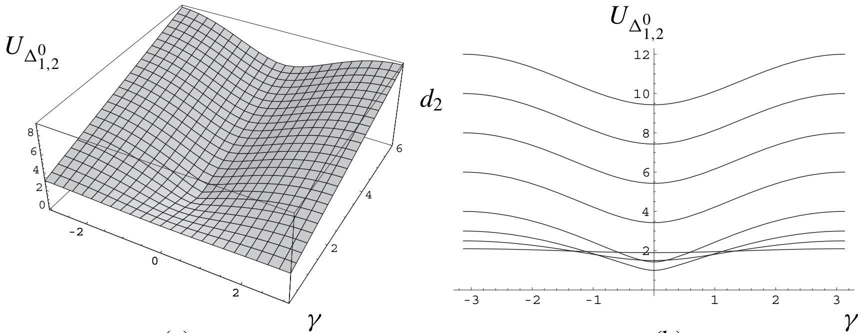
$$\begin{aligned}\mathcal{G}_{\Delta^0_{1,2}}(d_1, d_2, \gamma, \psi_1) &= d_1 \sin \psi_1 \sin(\psi_2^{\Delta} - \varphi_1^{\Delta}) \\ &\quad - d_2 \sin(\psi_1 - \varphi_1^{\Delta}) \sin(\psi_2^{\Delta} - \gamma).\end{aligned} \tag{38}$$

Note that if all players are in a straight line ($\gamma = 0$ or $\gamma = \pi$) and in the reduced space E_2 is located on the x-axis left of H_0^{Δ}, then $E_2 \in l^{\Delta}(d_1, 0)$ and the relations (34) and (35) are to be viewed as limits as $\psi_1 \to 0+$ with regard to (37). At these states

$$U_{\Delta^0_{1,2}}(d_1, d_2, \gamma, 0) = \begin{cases} d_2 + d_1 \dfrac{1+\beta_2}{1-\beta_1} & \textit{if}\ \gamma = \pi, \quad d_2 > 0, \\ -d_2 + d_1 \dfrac{1+\beta_2}{1-\beta_1} & \textit{if}\ \gamma = 0, \quad 0 \le d_2 < h_0^{\Delta}. \end{cases} \tag{39}$$

Figure 2 shows the field of E_2's optimal trajectories in $\Delta^0_{1,2}$.

Figure 3 shows the projection of $U_{\Delta^0_{1,2}}$ for $d_1 = 1$, $0 \le d_2 \le 6$, $|\gamma| < \pi$, $\beta_1 = 0.6$, $\beta_2 = 0.2$ and its cross sections for $d_2 = 0.1, 1, 2, 3, 4, 5, 6$.



(a) (b)

Figure 3: Surface of $U_{\Delta^0_{1,2}}$ and its cross sections.

4 Pursuit with Terminal Manifold $\mathsf{M}^{\emptyset}_{\Phi_2}$

Consider $\Phi_2^{\emptyset}$ in the playing space $\mathsf{Z}^{\emptyset}_{\Phi_2} = \mathsf{Z}_{\Phi_2} \backslash \mathsf{M}^{\emptyset}_{\Phi_2}$.

Let $\gamma^{\emptyset}$, $0 < \gamma^{\emptyset} < \pi$, satisfy the condition

$$\cos \gamma^{\emptyset} = \frac{\beta^2 - p_1^2 - p_2^2}{2 p_1 p_2}. \tag{40}$$

Obviously (see Proposition 1),

$$\partial \mathsf{M}^{\emptyset}_{\Phi_2} = \{\bar{z} : \arccos(\bar{z}_1 - \bar{z}_p) \cdot (\bar{z}_2 - \bar{z}_p)/|\bar{z}_1 - \bar{z}_p|/|\bar{z}_2 - \bar{z}_p| = \gamma^{\emptyset}\}.$$

Let us study $\Phi_2^{\emptyset}$ using the generalized Isaacs approach; see, e.g., [4]–[7], [9]. Let $V_{\Phi_2^{\emptyset}}(\bar{z})$ be the value of $\Phi_2^{\emptyset}$ in $\mathsf{Z}^{\emptyset}_{\Phi_2}$. By $V_{\bar{z}_i} = \partial V_{\Phi_2^{\emptyset}} / \partial \bar{z}_i$, $i \in I$, denote the adjoint variables. In the neighborhood of a regular point, smooth parts of $V_{\Phi_2^{\emptyset}}$ meet the Isaacs equation

$$\min_{\bar{u}_p} \max_{\{\bar{u}_1, \bar{u}_2\}} \{V_{\bar{z}_p} \cdot \bar{u}_p + V_{\bar{z}_1} \cdot \bar{u}_1 + V_{\bar{z}_2} \cdot \bar{u}_2\} = 0. \tag{41}$$

Suppose $V_{\bar{z}_i} \neq \bar{0}$, $i \in I$. Let $\bar{e}^t_p$, $\bar{e}^t_1$ and $\bar{e}^t_2$ be the unit vectors parallel to the optimal velocities of P, E_1 and E_2 at instant $t \geq 0$. It follows from (41) that $\bar{e}^t_p$, $\bar{e}^t_1$ and $\bar{e}^t_2$ are directed as $-V_{\bar{z}_p}$, $V_{\bar{z}_1}$ and $V_{\bar{z}_2}$. Besides,

$$-|V_{\bar{z}_p}| + \beta_1 |V_{\bar{z}_1}| + \beta_2 |V_{\bar{z}_2}| = 0. \tag{42}$$

The characteristic equations for (41) are represented as

$$\begin{aligned} \dot{\bar{z}}_p &= -V_{\bar{z}_p}/|V_{\bar{z}_p}|, \\ \dot{\bar{z}}_1 &= \beta_1 V_{\bar{z}_1}/|V_{\bar{z}_1}|, \\ \dot{\bar{z}}_2 &= \beta_2 V_{\bar{z}_2}/|V_{\bar{z}_2}|, \\ \dot{V}_{\bar{z}_i} &= \bar{0}, \quad i \in I. \end{aligned} \tag{43}$$

Using (43), we get that $\bar{e}_p^t$, $\bar{e}_1^t$ and $\bar{e}_2^t$ are constants along the optimal trajectories. Denote their values as $\bar{e}_p$, $\bar{e}_1$ and $\bar{e}_2$.

Parameterize $\partial\mathsf{M}_{\Phi_2}^{\emptyset}$ as

$$\begin{aligned}\bar{z}_p &= \bar{s}_p,\\ \bar{z}_1 &= \bar{s}_p + s_1\bar{e}(s_0),\\ \bar{z}_2 &= \bar{s}_p + s_2\bar{e}(s_0-\gamma^{\emptyset}),\end{aligned}\tag{44}$$

where $\bar{s}_p = (s_{px}, s_{py})$, $s_k \geq 0$, $k \in K$, and $\bar{e}(s)$ is the unit vector $(\cos s, \sin s)$.

The value of $\Phi_2^{\emptyset}$ on (44) is described by the expression

$$V_{\Phi_2^{\emptyset}}|_{\bar{z}\in\partial\mathsf{M}_{\Phi_2}^{\emptyset}} = p_1s_1 + p_2s_2.\tag{45}$$

Differentiating both parts of (45) with respect to $\bar{s}_p$, s_1, s_2, s_0, we get

$$\begin{aligned}V_{\bar{z}_p} + V_{\bar{z}_1} + V_{\bar{z}_2} &= \bar{0},\\ V_{\bar{z}_1}\cdot\bar{e}(s_0) &= p_1,\\ V_{\bar{z}_2}\cdot\bar{e}(s_0-\gamma^{\emptyset}) &= p_2,\\ V_{\bar{z}_1}\cdot s_1\bar{e}_\perp(s_0) + V_{\bar{z}_2}\cdot s_2\bar{e}_\perp(s_0-\gamma^{\emptyset}) &= 0,\end{aligned}\tag{46}$$

where $\bar{e}_\perp(s) = (-\sin s, \cos s)$.

It is clear that (42)–(46) are satisfied with

$$\begin{aligned}|V_{\bar{z}_p}| &= \beta,\\ |V_{\bar{z}_1}| &= p_1,\\ |V_{\bar{z}_2}| &= p_2,\\ \beta\bar{e}_p &= p_1\bar{e}_1 + p_2\bar{e}_2,\\ \bar{e}_1 &= \bar{e}(s_0),\\ \bar{e}_2 &= \bar{e}(s_0-\gamma^{\emptyset}).\end{aligned}\tag{47}$$

It follows from (47) that $\bar{e}_1$ and $\bar{e}_2$ make the angles $\gamma_1^{\emptyset}$ and $\gamma_2^{\emptyset}$ with $\bar{e}_p$, where

$$\begin{aligned}\cos\gamma_1^{\emptyset} &= \frac{\beta^2 + p_1^2 - p_2^2}{2\beta_1 p_1},\\ \cos\gamma_2^{\emptyset} &= \frac{\beta^2 - p_1^2 + p_2^2}{2\beta_2 p_2};\end{aligned}\tag{48}$$

see Figure 4. Evidently, $\gamma_1^{\emptyset} + \gamma_2^{\emptyset} = \gamma^{\emptyset}$.

In the reduced space, let P be located at the origin and let E_1 lie on the positive x-axis. Let φ_1, ψ_1 and $\psi_2 + \pi$ be the angles made by the optimal directions of P, E_1 and E_2; see Figures 5–11.

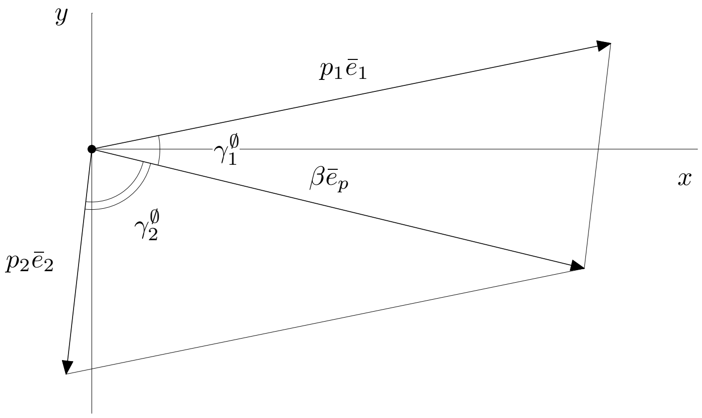


Figure 4: Collocation of $\beta\bar{e}_p$, $p_1\bar{e}_1$ and $p_2\bar{e}_2$.

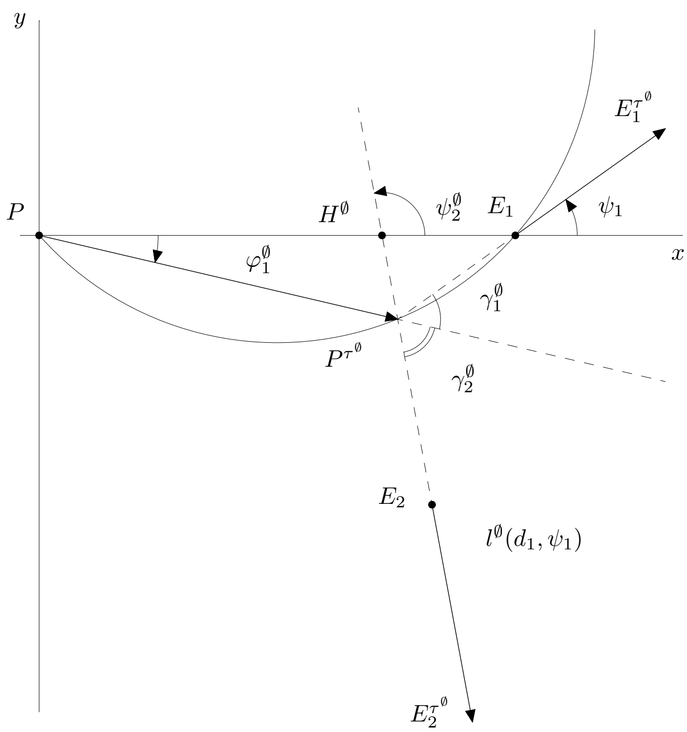


Figure 5: Geometry of optimal pursuit in $\Phi_2^\emptyset$ for $\mathsf{A}_{\Phi_2}^\emptyset$ and $0<\psi_1<\gamma_1^\emptyset$.

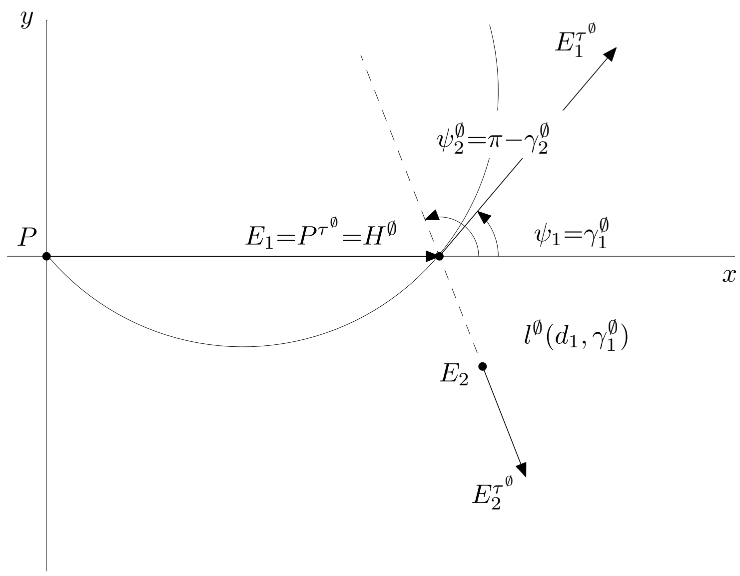


Figure 6: Geometry of optimal pursuit in $\Phi_2^\emptyset$ for $\mathsf{A}_{\Phi_2}^\emptyset$ and $\psi_1 = \gamma_1^\emptyset$.

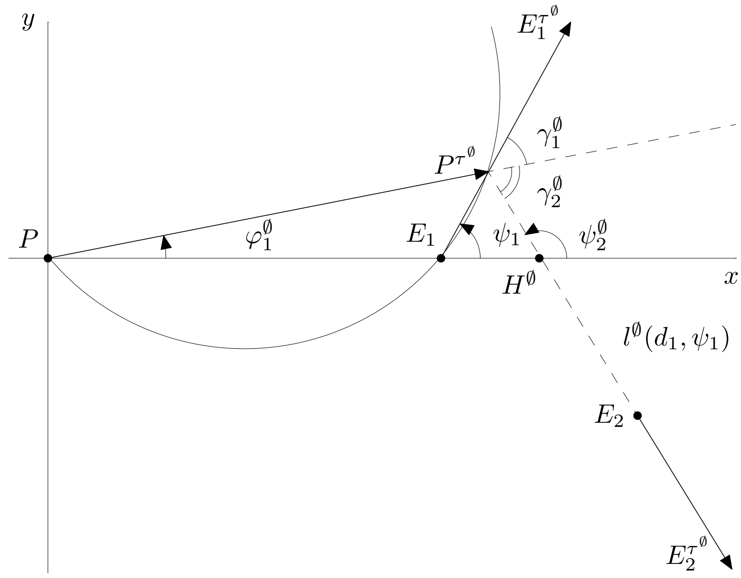


Figure 7: Geometry of optimal pursuit in $\Phi_2^\emptyset$ for $\mathsf{A}_{\Phi_2}^\emptyset$ and $\gamma_1^\emptyset < \psi_1 < \gamma^\emptyset$.

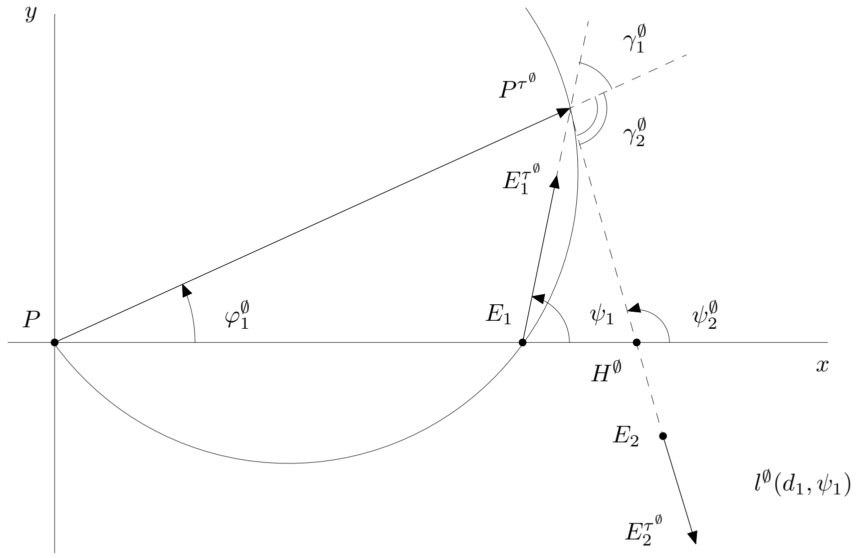


Figure 8: Geometry of optimal pursuit in $\Phi_2^{\emptyset}$ for $\mathsf{A}_{\Phi_2}^{\{1\}}$.

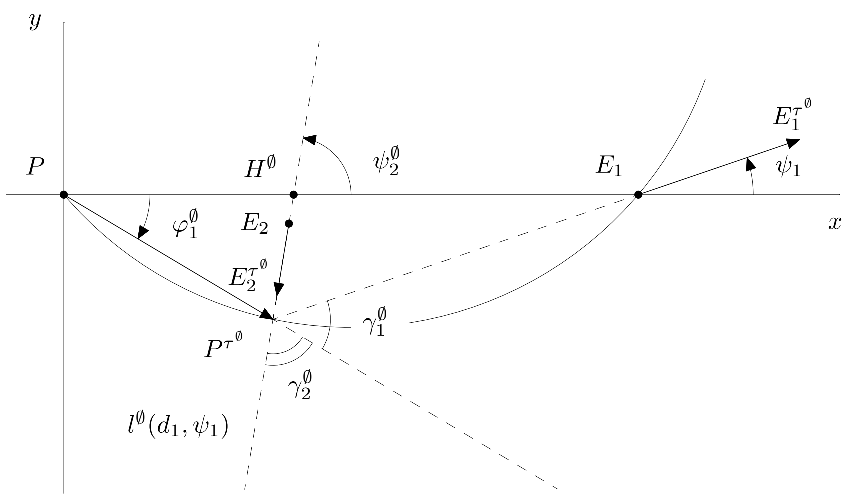


Figure 9: Geometry of optimal pursuit in $\Phi_2^{\emptyset}$ for $\mathsf{A}_{\Phi_2}^{\{2\}}$.

Let

$$\begin{aligned} \varphi_1^{\emptyset}(\psi_1) &= -\gamma_1^{\emptyset} \operatorname{sign} \psi_1 + \psi_1, \\ \psi_2^{\emptyset}(\psi_1) &= -(\pi - \gamma^{\emptyset}) \operatorname{sign} \psi_1 + \psi_1, \\ \tau^{\emptyset}(d_1, \psi_1) &= d_1 \sin |\psi_1| / \sin \gamma_1^{\emptyset}, \\ h^{\emptyset}(d_1, \psi_1) &= \tau_1^{\emptyset} \sin \gamma_2^{\emptyset} / \sin |\psi_2^{\emptyset}|. \end{aligned} \tag{49}$$

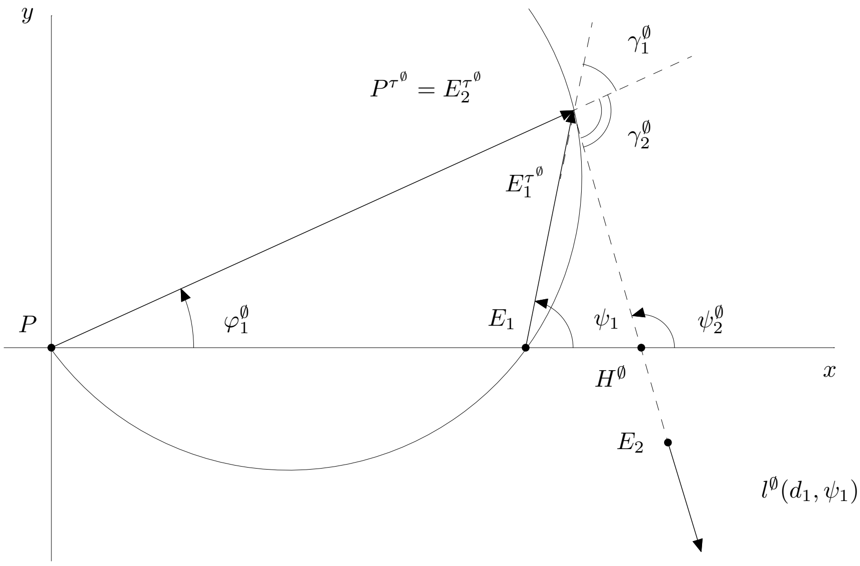


Figure 10: Geometry of optimal pursuit in $\Phi_2^{\emptyset}$ for $\partial \mathsf{A}_{\Phi_2}^{\{1\}}$.

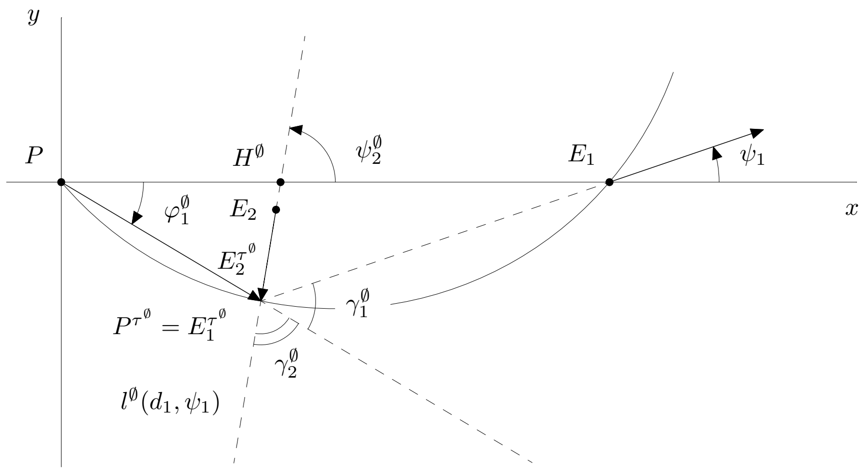


Figure 11: Geometry of optimal pursuit in $\Phi_2^{\emptyset}$ for $\partial \mathsf{A}_{\Phi_2}^{\{2\}}$.

Let

$$\Psi_1^{\emptyset} = \{|\psi_1| < \psi_1^{\emptyset},\ \psi_1 \neq 0\}, \tag{50}$$

where $\psi_1^{\emptyset}$ is the smallest positive root of the equation

$$\psi_2^{\emptyset}(\psi_1) = \pi. \tag{51}$$

Let $\bar{z}$ be an initial state with reduced coordinates (d_1, d_2, γ) and the players move at $\varphi_1^\emptyset$, ψ_1 and $\psi_2^\emptyset$ for the period $\tau^\emptyset$. The set of P's terminal positions $P^{\tau^\emptyset}$ in the reduced space may be parameterized as

$$\tau^\emptyset \bar{e}(\varphi_1^\emptyset), \quad \psi_1 \in \Psi_1^\emptyset. \tag{52}$$

It is easy to prove that (52) is the intersection of two circles of radius $d_1/2/\sin\gamma_1^\emptyset/2$ centered at $d_1/2\bar{e}(\pm(\pi - \gamma_1^\emptyset)/2)$.

Divide the playing space into several parts depending on the collocation of the players at the instant $t = \tau^\emptyset$. By $\mathsf{A}_{\Phi_2}^\emptyset$ denote the set with $\angle E_1^{\tau^\emptyset} P^{\tau^\emptyset} E_2^{\tau^\emptyset} = \gamma^\emptyset$ and $P^{\tau^\emptyset} \neq E_1^{\tau^\emptyset} \neq E_2^{\tau^\emptyset}$; see Figures 5–7. Similarly, let $\mathsf{A}_{\Phi_2}^{\{j\}}$ be the set with $\angle E_1^{\tau^\emptyset} P^{\tau^\emptyset} E_2^{\tau^\emptyset} = \pi - \gamma^\emptyset$ (it happens because E_j fails to pass through $P^{\tau^\emptyset}$ by $t = \tau^\emptyset$, $j \in J$); see Figures 8 and 9. And let $\partial\mathsf{A}_{\Phi_2}^{\{j\}}$ be the boundary of $\mathsf{A}_{\Phi_2}^{\{j\}}$ where $\angle E_1^{\tau^\emptyset} P^{\tau^\emptyset} E_2^{\tau^\emptyset} = \gamma^\emptyset$ and $P^{\tau^\emptyset} = E_j^{\tau^\emptyset}$, $j \in J$; see Figures 10 and 11.

In the reduced space, $\partial\mathsf{A}_{\Phi_2}^{\{1\}}$ is formed by two rays

$$l^\emptyset(d_1, \pm\psi_1^{\{1\}}),$$

where $\psi_1^{\{1\}}$ is the smallest root of the equation

$$\tau^\emptyset \bar{e}(\varphi_1^\emptyset) = d_1 \bar{e}(0) + \beta_1 \tau^\emptyset \bar{e}(\psi_1),$$

and $\partial\mathsf{A}_{\Phi_2}^{\{2\}}$ is parameterized as

$$\tau^\emptyset \bar{e}(\varphi_1^\emptyset) + \beta_2 \tau^\emptyset \bar{e}(\psi_2^\emptyset), \quad \psi_1 \in \Psi_1^{\{2\}},$$

where

$$\Psi_1^{\{2\}} = \{\psi_1 : |\psi_1| \le \gamma_1^\emptyset, \psi_1 \neq 0\}.$$

By $\mathcal{L}^\emptyset$ denote the mapping that takes (d_1, ψ_1) to the ray $l^\emptyset$ with vertex $H^\emptyset = h^\emptyset \bar{e}(0)$ and inclination $\psi_2^\emptyset$; see Figures 5–11.

The projection $M_{\Phi_2}^\emptyset$ of $\mathsf{M}_{\Phi_2}^\emptyset$ in the reduced space is the cone bounded by two rays passing through the origin and making the angles $\pi - \gamma^\emptyset$ and $-\pi + \gamma^\emptyset$.

Proposition 6. *For any given $d_1 > 0$, the mapping $\mathcal{L}^\emptyset$ is such that*

- *every point of the lower half-plane outside of $M_{\Phi_2}^\emptyset$ lies on a single ray $l^\emptyset(d_1, \psi_1)$ with $0 < \psi_1 < \psi_1^\emptyset$,*
- *every point of the lower half-plane outside of $M_{\Phi_2}^\emptyset$ lies on a single ray $l^\emptyset(d_1, \psi_1)$ with $0 < \psi_1 < \psi_1^\emptyset$,*

- *every point of the upper half-plane outside of $M^{\emptyset}_{\Phi_2}$ lies on a single ray $l^{\emptyset}(d_1, \psi_1)$ with $-\psi^{\emptyset}_1 < \psi_1 < 0$,*
- *every point of the positive x-axis lies simultaneously on two rays from the family*
$$\{l^{\emptyset}(d_1, \psi_1)\}_{\psi_1 \in \Psi^{\emptyset}_1} \tag{53}$$
with the values of ψ_1 that differ only by their signs.

Corollary 2. *For any given $d_1 > 0$, the mapping $\mathcal{L}^{\emptyset}$ is such that the rays (53) do not intersect each other and cover the plane of possible E_2 positions completely excluding the cone $M^{\emptyset}_{\Phi_2}$.*

Proposition 7. *Let (d_1, d_2, γ) describe $\bar{z}$ in the reduced space and let ψ_1 be chosen such that*
$$E_2 \in l^{\emptyset}(d_1, \psi_1). \tag{54}$$
Then, in $\Phi^{\emptyset}_2$ (see (49))

- *the optimal controls of P, E_1 and E_2 in the reduced space are defined by the angles $\varphi^{\emptyset}_1$, ψ_1 and $\psi^{\emptyset}_2$,*
- *the optimal duration $\tau_{\Phi^{\emptyset}_2}(\bar{z})$ equals $\tau^{\emptyset}$,*
- *the value $V_{\Phi^{\emptyset}_2}(\bar{z})$ is represented by*
$$U_{\Phi^{\emptyset}_2}(d_1, d_2, \gamma, \psi_1) = \rho(\bar{z}^{\tau^{\emptyset}}), \tag{55}$$
where $\bar{z}^{\tau^{\emptyset}} = (\tau^{\emptyset}\bar{e}(\varphi^{\emptyset}_1), d_1\bar{e}(0) + \beta_1\tau^{\emptyset}\bar{e}(\psi_1), d_2\bar{e}(\gamma) - \beta_2\tau^{\emptyset}\bar{e}(\psi^{\emptyset}_2))$,
- *the solution is regular with the exception of the dispersal surface that is projected into a part of the positive x-axis.*

Proposition 8. *At any regular state of $\Phi^{\emptyset}_2$ with reduced variables (d_1, d_2, γ) there exist a unique ψ_1 such that (54) is satisfied,*
$$\operatorname{sign}\psi_1 = -\operatorname{sign}\gamma, \tag{56}$$
and
$$\mathcal{G}_{\Phi^{\emptyset}_2}(d_1, d_2, \gamma, \psi_1) = 0, \tag{57}$$
where
$$\mathcal{G}_{\Phi^{\emptyset}_2}(d_1, d_2, \gamma, \psi_1) = d_1 \sin\psi_1 \sin(\psi^{\emptyset}_2 - \varphi^{\emptyset}_1) - d_2 \sin(\psi_1 - \varphi^{\emptyset}_1) \sin(\psi^{\emptyset}_2 - \gamma). \tag{58}$$

The optimal duration of $\Phi^{\emptyset}_2$ in $M^{\emptyset}_{\Phi_2}$ equals zero and the average distance to the coalition E^2 is assigned to the value $V_{\Phi^{\emptyset}_2}$.

Figure 12 shows the field of E_2's optimal trajectories in $\Phi^{\emptyset}_2$.

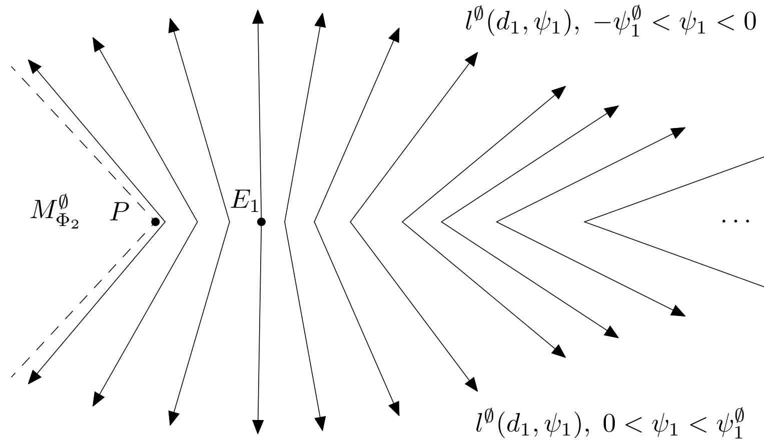


Figure 12: Field of E_2's optimal trajectories in $\Phi_2^{\emptyset}$.

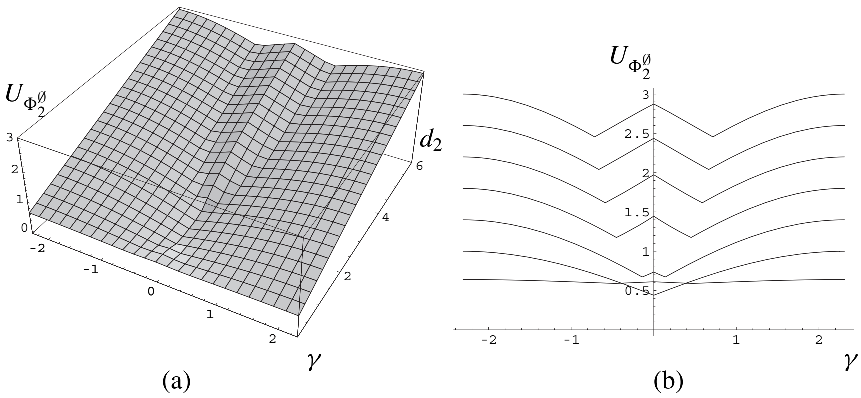


Figure 13: Surface of $U_{\Phi_2^{\emptyset}}$ and its cross sections.

Figure 13 shows the projection of $U_{\Phi_2^{\emptyset}}$ for $d_1 = 1$, $0 \le d_2 \le 6$, $|\gamma| < \gamma^{\emptyset}$, $\beta_1 = 0.6$, $\beta_2 = 0.2$, $p_1 = 0.6$, $p_2 = 0.4$ and its cross sections for $d_2 = 0.1, 1, 2, 3, 4, 5, 6$.

5 Alternative Pursuit

Consider Φ_2 in the playing space Z_{Φ_2} when P chooses the terminal manifold $\mathsf{M}_{\Phi_2}^{\{1\}}$, $\mathsf{M}_{\Phi_2}^{\{2\}}$ or $\mathsf{M}_{\Phi_2}^{\emptyset}$ that secures the least average distance. Suppose (17) and (19) hold. Then all domains of attraction $\mathsf{A}_{\Phi_2}^{\{1\}}$, $\mathsf{A}_{\Phi_2}^{\{2\}}$ and $\mathsf{A}_{\Phi_2}^{\emptyset}$ for $\mathsf{M}_{\Phi_2}^{\{1\}}$, $\mathsf{M}_{\Phi_2}^{\{2\}}$ and $\mathsf{M}_{\Phi_2}^{\emptyset}$ are

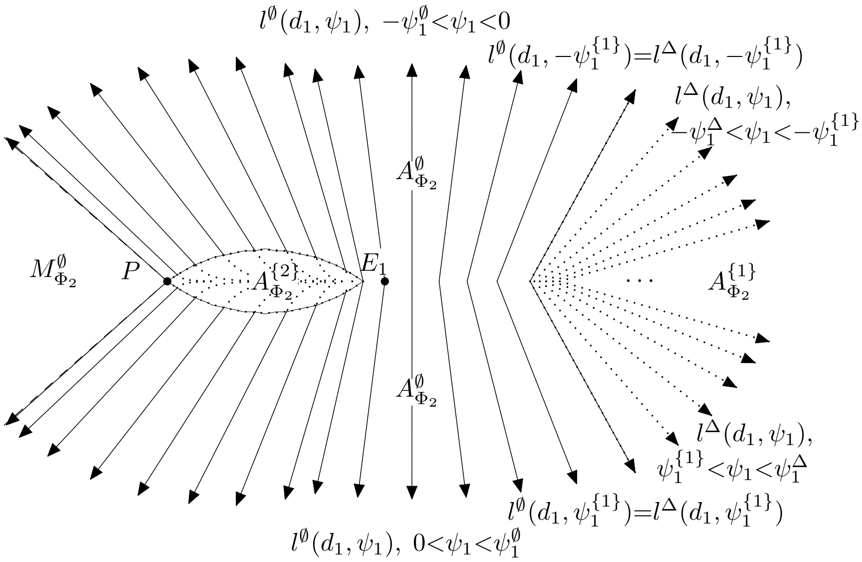


Figure 14: Field of E_2's optimal trajectories in Φ_2.

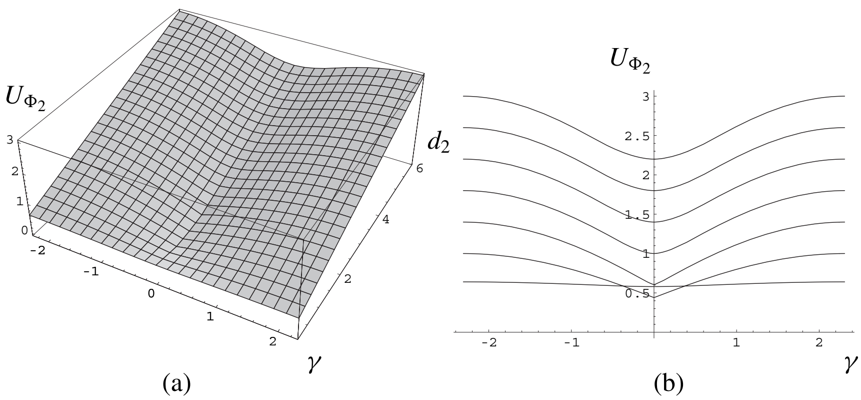


Figure 15: Surface of U_{Φ_2} and its cross sections.

nonempty. Let $V_{\Phi_2}(\bar{z})$ be the value of Φ_2. Obviously, the following relations of strategic equivalence are satisfied:

$$\begin{aligned}
&\Phi_2 \underset{\{(P,P),(E_1,E_1),(E_2,E_2)\}}{\overset{(\mathsf{A}_{\Phi_2}^{\emptyset},\bar{g}^{\equiv})}{\sim}} \Phi_2^{\emptyset},\\
&\Phi_2 \underset{\{(P,P),(E_1,E_1),(E_2,E_2)\}}{\overset{(\mathsf{A}_{\Phi_2}^{\{j\}},\bar{g}^{\equiv})}{\sim}} \Phi_2^{\{j\}}, \quad j \in J,
\end{aligned} \tag{59}$$

and

$$V_{\Phi_2}(\bar{z}) = \begin{cases} \min\{V^{\emptyset}_{\Phi_2}(\bar{z}),\, p_2 V^{\Delta}_{1,2}(\bar{z}),\, p_1 V^{\Delta}_{2,1}(\bar{z})\} & \text{if } \bar{z} \notin \mathsf{M}^{\emptyset}_{\Phi_2}, \\ \rho(\bar{z}) & \text{otherwise.} \end{cases} \tag{60}$$

Figure 14 shows the field of E_2's optimal trajectories in Φ_2 for nonempty $\mathsf{A}^{\{1\}}_{\Phi_2}$, $\mathsf{A}^{\{2\}}_{\Phi_2}$ and $\mathsf{A}^{\emptyset}_{\Phi_2}$.

Figure 15 shows the projection of U_{Φ_2} that represents V_{Φ_2} in the reduced space for $d_1 = 1$, $0.1 \le d_2 \le 6$, $|\gamma| < \gamma^{\emptyset}$, $\beta_1 = 0.6$, $\beta_2 = 0.2$, $p_1 = 0.6$, $p_2 = 0.4$ and its cross sections for $d_2 = 0.1, 1, 2, 3, 4, 5, 6$.

References

[1] Abramyantz T., Maslov E., Open-Loop coplanar pursuit of a multiple-target, *Podstawy Sterowania*, 11 (2), 149–160, 1981.

[2] Abramyants T.G., Volkovinskii M.I., Maslov E.P., Petrovskii A.M., A planar pursuit game with a limited number of corrections of the paths of the pursuer, *Avtomat. Telemekh.*, 7, 31–39, 1972.

[3] Berkowitz L.D., A variational approach to differential games, *Annals of Mathematics Study*, 52, 127–174, 1964.

[4] Bernhard P., Singular surfaces in differential games: An introduction. In P. Hargedorn, H.W. Knobloch, G.H. Olsder (Eds.) Differential Games and Applications, *Springer Lecture Notes in Information and Control Sciences*, Berlin: Springer, 3, 1–33, 1977.

[5] Bernhard P., Differential games: Introduction. In M.G. Singh (Ed.) *System and Control Encyclopedia*, Pergamon Press, Oxford, 1009–1010, 1987.

[6] Bernhard P., Differential games: Isaacs equation. In M.G. Singh (Ed.) *System and Control Encyclopedia*, Pergamon Press, Oxford, 1010–1017, 1987.

[7] Bernhard P., Closed-loop differential games. In M.G. Singh (Ed.) *System and Control Encyclopedia*, Pergamon Press, Oxford, 1004–1009, 1987.

[8] Gabasov R., Kirillova F.M., Some applications of functional analysis in the theory of optimal systems, *Izv. Akad. Nauk SSSR, Tekh. Kibern.*, 4, 3–13, 1966.

[9] Isaacs R., *Differential Games*, Wiley, New York, 1965.

[10] Kurzhanskii A.B., Differential games of approach in systems with delay, *Differents. Uravn.*, VII (8), 1389–1409, 1971.

[11] Kurzhanskii A.B., Osipov Yu. S., On some problems of open-loop pursuit in linear systems, *Izv. Akad. Nauk SSSR, Tekh. Kibern.*, 3, 18–29, 1970.

[12] Maslov E.P., Rubinovich E.Ya., Differential games with a multiple target, *Itogi Nauki Tekhn., Tekh. Kibern.*, 32, 32–59, 1991.

[13] Ol'shanskii V.K., Rubinovich E. Ya., Simple differential games of pursuit of a system of two evaders, *Avtomat. Telemekh.*, 1, 24–34, 1974.

[14] Shevchenko I.I., On open-loop pursuit-evasion, *Methods for Investigation of Nonlinear Control Systems*, Nauka, Moscow, 80–83, 1983.

[15] Shevchenko I.I., Approaching a coalition of evaders, *Avtomat. Telemekh.*, 1, 47–55, 1986.

[16] Shevchenko I.I., *Geometry of Alternative Pursuit*, Far Eastern State University Publishing, Vladivostok, 2003.

[17] Shevchenko I.I., Minimizing the distance to one evader while chasing another, *Computers and Mathematics with Applications*, 47 (12), 1827–1855, 2004.

PART IV
Evolutionary Game Theory and Applications

Adaptive Dynamics Based on Ecological Stability

József Garay
Research Group of Ecology and Theoretical Biology
Department of Plant Taxonomy and Ecology
Hungarian Academy of Science and Eötvös University
Pázmány P. sétány 1/C
H-1117 Budapest, Hungary
garayj@ludens.elte.hu

Abstract

An important step in coevolution occurs when a new mutant clone arises in a resident population of interacting individuals. Then, according to the ecological density dynamics resulting from the ecological interaction of individuals, mutants will go extinct or replace some resident clone or work their way into the resident system. One of the main points of this picture is that the outcome of the selection process is determined by ecological dynamics. For simplicity, we start out one from resident species described by a logistic model, in which the interaction parameters depend on the phenotypes of the interacting individuals. Using dynamic stability analysis we will answer the following purely ecological questions: After the appearance of a mutant clone,

(1) what kind of mutant cannot invade the resident population,
(2) and what kind of mutant can invade the resident population?
(3) what kind of mutant is able to substitute the resident clone,
(4) and when does a stable coexistence arise?

We assume that the system of mutants and residents can be modelled by a Lotka–Volterra system. We will suppose that the phenotype space is a subset of R^n and the interaction function describing the dependence of the parameters of the Lotka–Volterra dynamics on the phenotypes of the interacting individuals is smooth and mutation is small. We shall answer the preceding questions in terms of possible mutation directions in the phenotype space, based on the analysis of ecological stability. Our approach establishes a connection between adaptive dynamics and dynamical evolutionary stability.

1 Introduction

Let us consider a population of interacting asexual individuals that cannot change their phenotypes; for instance, their behaviour strategies are genetically fixed. Assume that mutation is a rare event and selection is fast, thus the selection process has enough time to eliminate the less fit phenotypes before any new mutation occurs.

A basic question of the theory of evolutionary stability is the following: What kind of mutant cannot invade the resident population? Maynard Smith and Price [13] have introduced the concept of ESS (evolutionarily stable strategy): "*An ESS is a strategy such that, if most of the members of the population adopt it, there is no mutant strategy that would have a higher reproduction fitness.*"[1]

An important question in coevolution is: In what direction does the phenotype change? This question is rather complex since it can be divided according to questions 2–4 of the Abstract. Most papers dealing with the theoretical aspects of this complex question apply what is called the adaptive dynamics approach, see, e.g., Marrow *et al.* [12] and Dieckmann and Law [7]. Because the model setup of this paper is almost the same as that of the adaptive dynamics approach, we shortly recall the latter and its most important statements.[2]

Let us consider the following coevolutionary dynamics:

$$\dot{n}_1 = n_1 W(x, y, n_1, n_2)$$
$$\dot{n}_2 = n_2 W(y, x, n_1, n_2),$$

where x, y are the phenotypes of the resident and the mutant, n_1, n_2 the density of the resident and the mutant and $W(x, y, n_1, n_2)$ and $W(y, x, n_1, n_2)$[3] denote the fitness function of the resident and the mutant, respectively. If there is no mutant in the system, we get the resident dynamics of the following form:

$$\dot{n}_1 = n_1 w(x, n_1). \tag{1}$$

Let n_1^* denote the equilibrium of the resident dynamics, $w(x, n_1^*) = 0$. If we suppose that the mutant is very rare, the sign of the invasive fitness of the mutant, $W(y, x, n_1^*, 0)$, determines what will happen to this mutant. For instance, if $W(y, x, n_1^*, 0) > 0$, or $\frac{W(y,x,n_1^*,0)}{y}\big|_{y=x} > 0$, then the mutant will spread. In the approach of adaptive dynamics x^* is called a "*singular strategy*" if $\frac{W(y,x^*,n_1^*,0)}{y}\big|_{y=x^*} = 0$. There are two types of singular strategies. If $\frac{\partial^2 W(y,x^*,n_1^*,0)}{\partial y^2}\big|_{y=x^*} < 0$, x^* is called an *ESS*. If $\frac{\partial^2 W(y,x^*,n_1^*,0)}{\partial y^2}\big|_{y=x^*} > 0$, x^* is said to be a "*branching point*." If the system starts at a branching point, a monomorphic system becomes dimorphic. Moreover, Geritz *et al.* [10] noted that: "*Close to a singular strategy it may happen that* $W(y, x^*, n_1^*, 0) > 0$ *and* $W(x^*, y, 0, n_2^*) > 0$, *so that both x and y are protected against extinction, and the population necessarily becomes dimorphic*."

[1] In mathematical terms, let us introduce a fitness function $W : S \times S \to R$, where S is the phenotype space. $x^* \in S$ is an ESS if $W(x^*, x^*) \geq W(y, x^*)$ for all mutant strategies $y \in S$, and if for a mutant strategy $W(x^*, x^*) = W(y, x^*)$ then $W(x^*, y) > W(y, y)$.

[2] For transparency, here we use the notation of our treatment rather than the usual notation of the adaptive approach.

[3] Here we follow the notation of game theory; namely, the fitness function denotes the fitness of the first phenotype variable.

The above-mentioned approaches are the same in that each considers only one species and the mathematical formalization is based on the notion of fitness, specifically, the fitness of residents with that of mutants. The main defect of these approaches is that none of them resulted in a general theory of density-dependent coevolutionary theory (cf. Hammerstein and Riechert [11] and Meszéna *et al.* [14]). In our view, in a density-dependent coevolutionary system containing more than one species only the dynamical approach can treat our initial question. The reason is that the future of the mutant does not depend only on the interaction with the resident but also on other species' resident and mutant phenotypes, i.e., essentially on the dynamic behaviour of the whole ecosystem. In earlier papers (Garay and Varga [9], Cressman *et al.* [5], and Cressman and Garay [3], [4]) we introduced an evolutionary stability notion based on the asymptotic behaviour of the selecting dynamic system. Although this approach can be worked out for multi-species resident systems, in this article only a one-species resident system is considered, since our main modelling-methodological question is whether there is any difference between the earlier fitness-centred method and the dynamics-centred method presented in this paper. As we will see, there are many differences.

The G-function approach (Vincent *et al.* [16], Cohen *et al.* [2], and Vincent and Brown [15]), on the one hand, is very similar to the approach of Cressman and Garay [3], [4], since both use population dynamics to describe the selection process, and both can deal with situations of more than one species. On the other hand, the G-function approach is more general because it simultaneously considers strategy dynamics and population dynamics. However, in this chapter we will not investigate a strategy dynamics. Instead, we will start out from a monomorphic model, in which the resident type and any mutant type (with genetically fixed phenotypes) interact. Furthermore, according to the ecological dynamics, either of them dies out or a stable coexistence develops. The novelty in this setup is the possibility to investigate the case of neutral mutants, as well.

Now we list the basic biological assumptions. Let us suppose that mutation is rare, which has two meanings: the mutant has a very low density, and following the appearance of the mutant, the ecological interactions have enough time to eliminate the less fit phenotypes, or reach a locally asymptotically stable dimorphic equilibrium. Since we consider a coevolving system, in our model the interaction parameters depend on the phenotypes of the interacting individuals. Moreover, we assume that the ecological interactions can be described by a Lotka–Volterra model.

In the first part, we will overview the stability properties of the two-dimensional Lotka–Volterra model. In the second part, we will assume that the phenotype space can be modelled by a body of the n-dimensional Eucledian space, mutation is arbitrarily small (infinitesimal) and the interaction parameters of the Lotka–Volterra model are smooth functions (actually quadratic polynomials) of the phenotypes. Based on these assumptions we will characterize both evolutionary stability and the possible mutation directions. The latter means that for a given resident state we describe the direction of phenotypic change implying that mutants will spread, die out or infiltrate the system, respectively. This characterization is strictly based on local and global asymptotic behaviour of the ecological dynamics.

2 Ecological Selection

Since the basic idea of our approach is that the evolutionary change is determined by the local qualitative behaviour of the ecological dynamics, first we have to collect well-known results concerning the asymptotic behaviour of the two-dimensional Lotka–Volterra model.

Let us suppose that the logistic model describes a monomorphic resident population, in which individuals interact with each other. Suppose that the interactions between individuals depend on their genetically fixed phenotypes. Thus we have an interaction function $f : S \times S \to R$, where S is the set of possible phenotypes. Since we suppose that the resident population is monomorphic, each individual has a phenotype $x \in S$. Let us denote by $n_1 \in R$ the density of the resident individuals and let $r \in R$ be the basic fitness. Then we have the following resident model:

$$\dot{n}_1 = n_1(r + f(x, x)n_1). \tag{2}$$

Obviously, this resident population has a stable equilibrium density $\hat{n}_1 = -\frac{r}{f(x,x)}$ if and only if $r > 0$, and $f(x, x) < 0$. In this chapter these inequalities are always supposed.

Now assume that a mutant subpopulation arises with phenotype $y \in S$. Let $n_2 \in R$ denote the density of the y-mutants. For simplicity, we assume that the basic fitness of different types is the same. Thus we have the following coevolutionary Lotka–Volterra model:

$$\begin{aligned} \dot{n}_1 &= n_1(r + f(x, x)n_1 + f(x, y)n_2) \\ \dot{n}_2 &= n_2(r + f(y, x)n_1 + f(y, y)n_2). \end{aligned} \tag{3}$$

From an evolutionary viewpoint, the first question is when the mutant can, and when it cannot invade the resident population.

2.1 What Kind of Mutant Cannot Invade the Resident Population?

Under the assumption that the density of the mutants is small, one of the possible mathematical descriptions of uninvadability is that the *resident equilibrium*, $(\hat{n}_1, 0) = \big(-\frac{r}{f(x,x)}, 0\big)$, is locally asymptotically stable. We have two cases.

2.1.1 The first case corresponds to the starting point of the adaptive dynamics school: Assume that the mutant's fitness, $\frac{\dot{n}_2}{n_2} = r + f(y, x)\hat{n}_1$, is strictly negative near the resident equilibrium. Then the mutant obviously cannot invade. Clearly, if

$$f(x, x) > f(y, x), \tag{4}$$

then it is locally asymptotically stable.

2.1.2 Now suppose that the mutant is neutral, i.e., $\frac{\dot{n}_2}{n_2} = r + f(y, x)\hat{n}_1 = 0$. Then at the resident equilibrium the mutant has the same fitness as the resident, if

$$f(x, x) = f(y, x). \tag{5}$$

An easy calculation shows that the Jacobian of (2) at the resident equilibrium has a zero eigenvalue, thus we have to use the method of the centre manifold, which gives that if $\frac{rf(x,x)-rf(y,x)}{f(x,x)} = 0$ and $\frac{f(y,y)f(x,x)-f(x,y)f(y,x)}{f(x,x)} < 0$ the $(\hat{n}_1, 0)$ is locally asymptotically stable. Summing up,

$$f(x, x) = f(y, x) \quad \text{and} \quad f(y, y) > f(x, y), \tag{6}$$

guarantee the uninvadability of the resident population.

We note that if $f(x, x) - f(y, x) = 0$ and $f(y, y) - f(x, y) = 0$, then the n_1-isocline is a zero ray, containing equilibrium points.

From an evolutionary viewpoint, a more important case is when the mutant can invade the population.

2.2 What Kind of Mutant Can Invade the Resident Population?

From an ecological perspective, a successful invasion requires the resident equilibrium to be unstable. We have two cases.

2.2.1 Obviously, if the mutant has positive fitness near the resident equilibrium, i.e., $r + f(y, x)\hat{n}_1 > 0$, then the mutant will invade. In other words, if $f(x, x) < f(y, x)$, then $(\hat{n}_1, 0)$ is locally unstable.

2.2.2 In the case when the mutant is neutral at the resident equilibrium, i.e., $f(x, x) = f(y, x)$, applying the centre manifold technique we obtain

$$f(x, x) = f(y, x) \quad \text{and} \quad f(y, y) \geq f(x, y),$$

which imply that the resident equilibrium is unstable.

Remark 1. It is known that the ESS definition of Maynard Smith has the following form: if the mutant is rare (i.e., its initial density is small) then $x^* \in S$ is an ESS if the following two conditions hold:

Equilibrium condition:

$$f(x^*, x^*) \geq f(y, x^*) \text{ for all } y \in S.$$

Stability condition:

$$\text{if for } y \in S \; f(x^*, x^*) = f(y, x^*), \text{ then } f(x^*, y) > f(y, y).$$

After the successful invasion of mutants there are two possibilities. The first one is that invader mutants and residents coexist. Coexistence in the Lotka–Volterra model means that either there is a locally asymptotically stable rest point of the dynamics (2), or there is a stable limit cycle. In both cases there must exist an interior rest point of the dynamics (2). We will consider only the case of a locally asymptotically stable rest point. The second possibility is that the mutant replaces the resident population. Then if the interior rest point exists, it must be unstable.

2.3 What Kind of Mutant Can Coexist with the Resident?

From an ecological viewpoint, coexistence will arise if the resident equilibrium is unstable, and if there is an appropriate neighbourhood of the resident equilibrium such that each trajectory starting from this neighbourhood will not go to a locally asymptotically stable resident equilibrium.[4]

The first question is: When does an interior equilibrium exist? It is easy to see that, if $f(y,y)f(x,x) - f(x,y)f(y,x) \neq 0$, then the unique interior rest point has the following coordinates:

$$n_1^* = \frac{rf(x,y) - rf(y,y)}{f(y,y)f(x,x) - f(x,y)f(y,x)}$$
$$n_2^* = \frac{rf(y,x) - rf(x,x)}{f(y,y)f(x,x) - f(x,y)f(y,x)}.$$

For coexistence we require that each coordinate of the interior equilibrium be positive, $(n_1^*, n_2^*) \in R_+^2$, i.e.,

$$\begin{aligned} sign(f(x,y) - f(y,y)) &= sign(f(y,x) - f(x,x)) \\ &= sign(f(y,y)f(x,x) - f(x,y)f(y,x)) \neq 0. \end{aligned}$$

If $f(y,y)f(x,x) - f(x,y)f(y,x) = 0$, then there is no unique interior equilibrium. In this case each point of the graph of function $n_1 = -\frac{r}{f(x,x)} - \frac{f(y,y)}{f(x,x)} n_2$ is an equilibrium. This degenerate case is also excluded.

2.3.1 Stable coexistence

The second question is: When is an interior rest point locally asymptotically stable? First, let us consider the case when local asymptotic stability of $(n_1^*, n_2^*) \in R_+^2$ is guaranteed by the linearization method. An easy calculation shows that both eigenvalues of the Jacobian of (2) have a negative real part if and only if $n_1^* f(x,x) + n_2^* f(y,y) < 0$ and $f(y,y)f(x,x) - f(x,y)f(y,x) > 0$. From the latter inequality, for the existence of a unique interior equilibrium we have

$$f(x,y) - f(y,y) > 0,$$
$$f(y,x) - f(y,x) > 0.$$

From an ecological point of view, these inequalities mean the following: If the competition within clones is stronger than the competition between clones then there always exists a unique locally asymptotically stable interior equilibrium.

Observe that the latter inequality is the same as that in 2.2.1, guaranteeing the invadability by the mutant. Underlying this result is the fact that the Lotka–Volterra

[4]For general ecological dynamics, the situation is more complex since it is possible that in each neighbourhood of the resident equilibrium there exists an initial state from which the mutant goes extinct, and another initial state from which the system goes to coexistence.

model has linear isoclines.[5] In other words, in our two-dimensional Lotka–Volterra model, since the interior equilibrium is also globally asymptotically stable, the preceding conditions also guarantee invadability by the mutant.

Summing up the above four inequalities imply the existence and local asymptotic stability of $(n_1^*, n_2^*) \in R_+^2$.

2.4 What Kind of Mutant Can Replace the Resident Population?

The replacement process in ecological terms means that the resident equilibrium is unstable, and there is an appropriate neighbourhood of the resident equilibrium such that each trajectory starting in this neighbourhood arrives in the resident equilibrium. Consequently, the mutant equilibrium, $(0, \hat{n}_2) = \left(0, -\frac{r}{f(y,y)}\right)$, is necessarily asymptotically stable. Of course, for the existence of an equilibrium of the mutant clone we need $r > 0$ and $f(y, y) < 0$. Similar to the above resident equilibrium, for the mutant equilibrium we also have two cases.

a) If $r + f(x, y)\hat{n}_2 < 0$, then $(0, \hat{n}_2)$ obviously is locally asymptotically stable. Since $f(y, y) < 0$, we have $f(y, y) > f(x, y)$.

b) Now suppose that $r + f(x, y)\hat{n}_2 < 0$, i.e., $f(y, y) = f(x, y)$. Similar to the case of the resident equilibrium, the Jacobian has a zero eigenvalue, and again from centre manifold theory we get that the mutant equilibrium is locally asymptotically stable if and only if either $f(y, y) > f(x, y)$ or both $f(y, y) = f(x, y)$ and $f(y, y)f(x, x) - f(x, y)f(y, x) > 0$ are fulfilled. From the above we get that if the interior equilibrium is unstable, then both the resident and the mutant equilibria are locally asymptotically stable. So, in the two-dimensional Lotka–Volterra model the mutant cannot invade. Thus, in the case of replacement there is no interior rest point. We have the following three main possibilities.

2.4.1 The mutant has positive fitness both near the resident equilibrium and near the mutant equilibrium, which is guaranteed by

$$\begin{aligned} f(x, x) &< f(y, x), \\ f(y, y) &> f(x, y). \end{aligned}$$

2.4.2 The mutant has positive fitness near the resident equilibrium and is neutral at the mutant equilibrium, but the mutant equilibrium is locally asymptotically stable, which is guaranteed by

$$\begin{aligned} &f(x, x) < f(y, x), \\ &f(y, y) = f(y, x) \quad \text{and} \quad f(x, x) < f(y, x). \end{aligned}$$

[5]Obviously, if the ecological dynamics are more complex, then the condition for local asymptotic stability of the interior rest point may be independent from the condition of invadability by the mutant, which is a local condition for the resident equilibrium laying on the border of the positive orthant.

2.4.3 The mutant is neutral at the resident equilibrium, and the resident has negative fitness near the mutant equilibrium:

$$f(x,x) = f(y,x) \quad \text{and} \quad f(y,y) = f(x,y)$$
$$f(y,y) > f(x,y).$$

2.5 Adaptive Change in the Competitive Lotka–Volterra Model

We emphasise at this point that we do not assume that the mutation produces only an arbitrarily small change in phenotype. Now let us consider only the case of a Lotka–Volterra type competitive interaction, i.e., $f : S \times S \to R_-$. Observe that there is no assumption on the phenotypic space. If mutation is rare but arbitrary we have two different evolutionary scenarios: The first, more usual case is when the y-mutant can replace the x-resident if

$$\text{either } f(y,x) > f(x,x) \text{ and } f(y,y) \geq f(x,y),$$
$$\text{or } f(y,x) = f(x,x) \text{ and } f(y,y) > f(x,y).$$

The second rare case is when stable dimorphism evolves with a locally asymptotic ecological rest point. Observe that there is no assumption on the connection between resident and mutant phenotypes, thus a " branching" occurs if and only if $f(x,y) > f(y,y)$ and $f(y,x) > f(x,x)$.

In this setup, there is a discrete series of the phenotypes, thus there is no continuous dynamics in the phenotype space to describe the evolutionary change in phenotypes.

3 Multidimensional Interaction Function

Following the theory of continuous ESS and the theory of adaptive dynamics, let us suppose that $S \subset R^n$ is a body with piecewise smooth surface, and the interaction function $f : R^n \times R^n \to R$ is a quadratic polynomial smooth with respect to both phenotypes. Moreover, we also suppose that the evolution can change the phenotype in very small steps, described by the vectors of the set $T_x := \{y - x \in R^n \mid y \in S\}$[6] is an arbitrary small vector. We can suppose that f is negative on $S \times S$. The arbitrary small mutation is a strict restriction since all requirements become local conditions.

In the following we want to characterize the evolutionary event in terms of the interaction function f. We introduce the notation $\frac{\partial f}{\partial_{1i}}(z^1,z^2) := \frac{\partial f}{\partial z_i^1}(z^1,z^2)$, $\frac{\partial^2 f}{\partial_{1i}^2}(z^1,z^2) := \frac{\partial^2 f}{\partial (z_i^1)^2}(z^1,z^2)$, $\frac{\partial^2 f}{\partial_{1i}\partial_{2i}}(z^1,z^2) := \frac{\partial f}{\partial z_i^1 \partial z_i^2}(z^1,z^2)$, and so on. The gradient

[6] T_x is usually called the tangent cone of S at x.

$grad f(z^1, z^2) := \left(\frac{\partial f}{\partial z_1^1}(z^1, z^2), \dots, \frac{\partial f}{\partial z_n^1}(z^1, z^2), \frac{\partial f}{\partial z_1^2}(z^1, z^2), \dots, \frac{\partial f}{\partial z_n^2}(z^1, z^2)\right) \in$ R^{2n} in partitioned form will be written as

$grad f(z^1, z^2) := \left(grad_1 f(z^1, z^2),\ grad_2 f(z^1, z^2)\right) \in R^{2n}$.

Furthermore, the Hessian matrix of function f is the following:

$$\mathbf{F}(z^1, z^2) := \begin{pmatrix} \frac{\partial^2 f}{\partial_{11}^2} & \cdots & \frac{\partial^2 f}{\partial_{1n}\partial_{11}} & \frac{\partial^2 f}{\partial_{21}\partial_{11}} & \cdots & \frac{\partial^2 f}{\partial_{2n}\partial_{11}} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial_{11}\partial_{1n}} & \cdots & \frac{\partial^2 f}{\partial_{1n}^2} & \frac{\partial^2 f}{\partial_{21}\partial_{1n}} & \cdots & \frac{\partial^2 f}{\partial_{2n}\partial_{1n}} \\ \frac{\partial^2 f}{\partial_{21}\partial_{11}} & \cdots & \frac{\partial^2 f}{\partial_{21}\partial_{1n}} & \frac{\partial^2 f}{\partial_{21}^2} & \cdots & \frac{\partial^2 f}{\partial_{2n}\partial_{21}} \\ \vdots & \vdots & & & \vdots & \vdots \\ \frac{\partial^2 f}{\partial_{21}\partial_{11}} & \cdots & \frac{\partial^2 f}{\partial_{2n}\partial_{1n}} & \frac{\partial^2 f}{\partial_{21}\partial_{2n}} & \cdots & \frac{\partial^2 f}{\partial_{2n}^2} \end{pmatrix} \in R^{2n \times 2n}.$$

Later we will use the following partitioned form of this Hessian:

$$\mathbf{F}(z^1, z^2) := \begin{pmatrix} \mathbf{F}_{11}(z^1, z^2) & \mathbf{F}_{12}(z^1, z^2) \\ \mathbf{F}_{21}(z^1, z^2) & \mathbf{F}_{22}(z^1, z^2) \end{pmatrix}.$$

Let $(0, \xi)$,$(\xi, 0)$, $(\xi, \xi) \in R^{2n}$ where $\xi \in R^n$ and 0 is the n-dimensional zero vector. Since for simplicity we assumed that the interaction function is a quadratic polynomial of its variables, using this notation, for the corresponding Taylor polynomial we get the following equalities:

$$f(x, y) = f(x, x) + \langle grad f(x, x), (0, \xi)\rangle + \frac{1}{2}(0, \xi)\mathbf{F}(x, x)\begin{pmatrix} 0 \\ \xi \end{pmatrix}$$
$$f(y, x) = f(x, x) + \langle grad f(x, x), (\xi, 0)\rangle + \frac{1}{2}(\xi, 0)\mathbf{F}(x, x)\begin{pmatrix} \xi \\ 0 \end{pmatrix}$$
$$f(y, y) = f(x, x) + \langle grad f(x, x), (\xi, \xi)\rangle + \frac{1}{2}(\xi, \xi)\mathbf{F}(x, x)\begin{pmatrix} \xi \\ \xi \end{pmatrix}.$$

4 Evolutionary Ecological Stability

Here we characterize the terminal state of evolution, in terms of a given interaction function f. We have already shown in 2.1.1 and 2.1.2 what kind of mutant cannot invade the resident population.

4.1. The first case again is, when the mutant's fitness is negative near the resident equilibrium. Then the condition of uninvadability, $f(x, x) > f(y, x)$, reads as follows:

either $\langle grad_1 f(x, x), \xi\rangle < 0$,
or $\langle grad_1 f(x, x), \xi\rangle = 0$ and $\xi \mathbf{F}_{11}(x, x)\xi < 0$.

Hence we obtain that if the interaction function f attains a strict maximum in its first variable at x, while the second variable is fixed at x, then there is no invader mutant. In game-theoretical terms this means that x is a strict best reply against itself.

4.2 As we have seen above, when the mutant is neutral, i.e., $f(x,x) = f(y,x)$, then $f(x,y) > f(y,y)$ guarantees the uninvadability of the resident population. At first glance, this possibility does not seem to be important since it is defined only by the equality of the mutant's and resident's fitness. Nevertheless, from evolutionary (cf. definitions of ESS) and mathematical point of view, this is a very important case, since if $f(x,x) = f(y,x)$ holds for all mutant phenotypes, we get that x is a Nash equilibrium. As mentioned above, the resident equilibrium is locally asymptotically stable, if $f(x,y) > f(y,y)$, which means

$$f(x,x) + \langle gradf(x,x), (0,\xi)\rangle + \frac{1}{2}(0,\xi)\mathbf{F}(x,x)\begin{pmatrix} 0 \\ \xi \end{pmatrix}$$
$$> f(x,x) + \langle gradf(x,x), (\xi,\xi)\rangle + \frac{1}{2}(\xi,\xi)\mathbf{F}(x,x)\begin{pmatrix} \xi \\ \xi \end{pmatrix}.$$

The dominant part of the latter inequality is

$$\langle grad_1 f(x,x), \xi\rangle \geq 0.$$

If $\langle grad_1 f(x,x), \xi\rangle = 0$ then $\xi\mathbf{F}_{22}(x,x)\xi > \xi\mathbf{F}_{11}(x,x)\xi + 2\xi\mathbf{F}_{12}(x,x)\xi + \xi\mathbf{F}_{22}(x,x)\xi$ thus

$$\xi\mathbf{F}_{11}(x,x)\xi + 2\xi\mathbf{F}_{12}(x,x)\xi < 0.$$

Following Maynard Smith's idea we shall say x^* is a Lotka–Volterra evolutionarily stable strategy (LV-ESS)[7] if the overwhelming majority of the population have phenotype x^* and no rare mutant can invade the population. Now we are in a position to describe Maynard Smith's ESS definition in terms of the interaction function.

For evolutionary stability, the conditions of 4.1 and 4.2 must be simultaneously satisfied for all possible phenotypic mutation directions. As we have seen $\langle grad_1 f(x,x), \xi\rangle \leq 0$ must hold. If ξ and $-\xi \in T_x$, then this inequality must hold for ξ and $-\xi$, as well, thus for ξ we have $\langle grad_1 f(x,x), \xi\rangle = 0$. Therefore, if $x^* \in intS$ then equality holds for all mutation directions ξ. Summarizing, if function f is a polynomial of the seconds degree, x^* is an LV-ESS if and only if the following conditions hold.

Equilibrium condition: for all $\xi \in T_x$ small enough

$$\langle grad_1 f(x^*,x^*), \xi\rangle \leq 0, \text{ and}$$
$$\text{if } \langle grad_1 f(x^*,x^*), \xi\rangle = 0 \text{ then } \xi\mathbf{F}_{11}(x,x)\xi \leq 0.$$

[7]Although LV-ESS is strictly based on the one-species ESS, we use a different name, since the concept of LV-ESS strongly depends on the ecological dynamics, cf. Cressman and Garay [3], [4]. The mutation is also supposed to be arbitrarily small, thus it is a local notion; moreover, the interaction coefficient is a smooth function of its variables.

Stability condition: If for some $\xi \in T_x$ $\langle grad_1 f(x^*, x^*), \xi\rangle = 0$ and $\xi \mathbf{F}_{11}(x^*, x^*)\xi = 0$, then

$$\xi \mathbf{F}_{12}(x^*, x^*)\xi \leq 0.$$

Clearly, the above Lotka–Volterra based, one-species LV-ESS is a continuous version of Maynard Smith's ESS. However, there is an important difference between the ESS of the adaptive dynamics approach and the LV-ESS, since the latter allows us to take into account the neutral mutants (when $\langle grad_1 f(x^*, x^*), \xi\rangle = 0$ and $\xi \mathbf{F}_{11}(x^*, x^*)\xi = 0$), when the stability depends on the resident-mutant interaction. This means that the ESS of the adaptive dynamical approach implies the LV-ESS, but not conversely.

5 Adaptive Changes

5.1 In Which Direction Can the Phenotypes Change?

As before, here we also have two possibilities for invadability.

5.1.1. Suppose that the mutant's fitness is positive. Similarly to 4.1, we get $\langle grad_1 f(x, x), \xi\rangle \geq 0$. If $\langle grad_1 f(x, x), \xi\rangle = 0$ then $\xi \mathbf{F}_{11}(x, x)\xi > 0$. We emphasise that this statement is known from the adaptive dynamical approach states. There is only one main difference between them; namely, in our setup there is no reason to suppose that the evolution follows the gradient of the fitness function at resident equilibrium, cf. Dickmann and Law [7]. As a matter of fact, our approach is strictly based on the assumption that there is only one monomorphic mutant population. That is the reason why the qualitative behaviour of the two-dimensional Lotka–Volterra model determines the evolutionary events. Since it can be supposed that mutation " chooses" mutant phenotypes at random, there is no reason to suppose that the direction of evolution is the same as the gradient of the fitness of the mutant.

5.1.2 Suppose that the mutant is neutral, as we have seen in 4.2. The neutrality condition reads $f(x, x) = f(y, x)$, i.e., $\langle grad_1 f(x, x), \xi\rangle = 0$ and $\xi \mathbf{F}_{11}(x, x)\xi = 0$. In the case of neutrality of the mutant, the resident equilibrium is locally unstable if $f(y, y) > f(x, y)$. In parallel with 4.2, we get $\xi \mathbf{F}_{11}(x, x)\xi + 2\xi \mathbf{F}_{12}(x, x)\xi > 0$. Observe that this possibility is neglected by the adaptive dynamics approach. Summarizing, a direction of successful phenotypic change of evolution is characterized by the following inequalities:

$\langle grad_1 f(x, x), \xi\rangle \geq 0$.

If $\langle grad_1 f(x, x), \xi\rangle = 0$ then $\xi \mathbf{F}_{11}(x, x)\xi \geq 0$.

If $\langle grad_1 f(x, x), \xi\rangle = 0$ and $\xi \mathbf{F}_{11}(x, x)\xi = 0$ then $\xi \mathbf{F}_{12}(x, x)\xi > 0$.

5.2 After the Successful Invasion

After the mutant successfully invades the resident population, there are two possibilities: either stable coexistence arises or the mutant replaces the resident population.

5.2.1 What Kind of Mutant can Coexist with the Resident?

As above, we consider only the case of the locally asymptotically stable rest point of the competition model (i.e., when for all $x, y \in S$ $f(x,x)$, $f(x,y)$, $f(y,x)$, $f(y,y) < 0$, and $r > 0$). As we have seen above, for the stable coexistence we need $f(x,y) - f(y,y) > 0$ and $f(y,x) - f(x,x) > 0$.

Inequality $f(x,y) - f(y,y) > 0$ holds if $\langle grad_1 f(x,x), \xi\rangle \leq 0$, and $\langle grad_1 f(x,x), \xi\rangle = 0$ implies $\xi \mathbf{F}_{11}(x,x)\xi + 2\xi\mathbf{F}_{12}(x,x)\xi < 0$.

Inequality $f(y,x) - f(x,x) > 0$ holds if $\langle grad_1 f(x,x), \xi\rangle \geq 0$, and $\langle grad_1 f(x,x), \xi\rangle = 0$ implies $\xi\mathbf{F}_{11}(x,x)\xi > 0$.

Summing up, in direction ξ speciation occurs if

$$\langle grad_1 f(x,x), \xi\rangle = 0, \xi\mathbf{F}_{11}(x,x)\xi > 0 \text{ and}$$
$$\xi\mathbf{F}_{11}(x,x)\xi + 2\xi\mathbf{F}_{12}(x,x)\xi < 0.$$

Observe that the latter two inequalities imply that $\xi\mathbf{F}_{12}(x,x)\xi < 0$. Moreover, the first and the second inequalities imply that the speciation state x must be a Nash equilibrium of the function $-f$, which means that the interaction function attains a minimum in its first variable at x, while its second variable is fixed at x.

The first two of the above three inequalities are the same as those of the adaptive dynamics approach, which guarantee "branching." Both of them start from the minimum fitness of the resident, where the fitness gradient is zero. The reason for this is that $f(x,y) - f(y,y) > 0$ implies that the resident can invade the mutant equilibrium and $f(y,x) - f(x,x) > 0$ implies that the mutant can invade the resident equilibrium (cf. Geritz *et al.* [10]). Although the conditions of branching are the same in the adaptive dynamics and our approach, this condition from our viewpoint guarantees the global asymptotic stability of the interior rest point; thus it also implies that the system is permanent. From a dynamical viewpoint, this is more important than the fact that the mutant and resident rest points are locally unstable.

Remark 2. Suppose that the interaction function f is linear in its first variable, i.e., $f(z^1, z^2) = z^1 g(z^2)$. It is easy to see that only the strict inequalities $(x-y)g(y) > 0$ and $(y-x)g(x) > 0$ together can guarantee the coexistence. Since the mutation is arbitrarily small, we have $g(y) = g(x) + o(x,\xi)$, where $o(x,\xi)\xrightarrow[\xi\to 0]{}0$. Thus $(x - y)(g(x) + o(x,\xi)) > 0$, implying $(x-y)g(x) \geq 0$, which is a contradiction with $(y-x)g(x) > 0$. This means that in the situation when the Lotka–Volterra model describes the ecological selection, mutation is arbitrarily small, the interaction function is smooth and linear in its first variable, then only the replacing process is possible, speciation is not.

5.2.2 What Kind of Mutant Can Replace the Resident Population?

As we have seen in Section 2.4., we have the following three possibilities.

5.2.2.1 The mutant has positive fitness both near the resident equilibrium and near the mutant equilibrium. In 2.4.1 we have seen that these are guaranteed by $f(y,y) > f(x,y)$ and $f(y,x) > f(x,x)$.

Inequality $f(y, y) > f(x, y)$ holds if $\langle grad_1 f(x, x), \xi\rangle \geq 0$, and $\langle grad_1 f(x, x), \xi\rangle = 0$ implies $\xi\mathbf{F}_{11}(x, x)\xi + 2\xi\mathbf{F}_{12}(x, x)\xi > 0$.

Inequality $f(y, x) > f(x, x)$ holds if $\langle grad_1 f(x, x), \xi\rangle \geq 0$, and $\langle grad_1 f(x, x), \xi\rangle = 0$ implies $\xi\mathbf{F}_{11}(x, x)\xi > 0$.

Summarizing, the resident dies out while the mutant persists, if the following inequalities hold:

either $\langle grad_1 f(x, x), \xi\rangle \geq 0$,
or $\langle grad_1 f(x, x), \xi\rangle = 0$ and
$\xi\mathbf{F}_{11}(x, x)\xi > 0$ with
$\xi\mathbf{F}_{11}(x, x)\xi + 2\xi\mathbf{F}_{12}(x, x)\xi > 0$.

5.2.2.2 The mutant has positive fitness near the resident equilibrium and the resident is neutral at the mutant equilibrium, but the mutant equilibrium is locally asymptotically stable. In 2.4.2 we have seen that these are guaranteed by

$$f(y, y) = f(x, y) \quad \text{and} \quad f(y, x) > f(x, x).$$

Inequality $f(y, x) > f(x, x)$ holds if either $\langle grad_1 f(x, x), \xi\rangle \geq 0$, or $\langle grad_1 f(x, x), \xi\rangle = 0$ and $\xi\mathbf{F}_{11}(x, x)\xi > 0$.

Moreover, $f(y, y) = f(x, y)$ holds if

$$\langle grad_1 f(x, x), \xi\rangle + \xi\mathbf{F}_{12}(x, x)\xi + \frac{1}{2}\xi\mathbf{F}_{22}(x, x)\xi = 0.$$

5.2.2.3 The mutant is neutral at the resident equilibrium and the resident has negative fitness near the mutant equilibrium, i.e., $f(x, x) = f(y, x)$ and $f(y, y) > f(x, y)$.

Equality $f(x, x) = f(y, x)$ holds if

$$\langle grad_1 f(x, x), \xi\rangle + \frac{1}{2}\xi\mathbf{F}_{11}(x, x)\xi = 0.$$

Inequality $f(y, y) > f(x, y)$ holds if either $\langle grad_1 f(x, x), \xi\rangle \geq 0$, or $\langle grad_1 f(x, x), \xi\rangle = 0$ and $\xi\mathbf{F}_{11}(x, x)\xi + 2\xi\mathbf{F}_{12}(x, x)\xi > 0$.

6 Discussion

This chapter is based on the following two conditions. First, mutation is a very rare event, which implies on the one hand, that the density of the mutant is very small, and on the other hand, that at most one mutant phenotype arises in each species at a time. Furthermore, before a new mutation occurs, the ecological selection has enough time to eliminate the less fit phenotypes. Moreover, the mutation is not generally arbitrarily small. Second, in coevolution the selection is determined by the ecological interaction of the different species. The main point is that the future of mutants is determined by all their ecological interactions, and not by the competition between them and their resident. The outcome of the ecological selection is given by the asymptotic behaviour of the ecological dynamics. The mutants' future is determined by the local asymptotic behaviour of the coevolutionary dynamics near the equilibrium $(n^*, 0) \in R^{2n}$. If the latter is unstable then the mutant can invade the resident system. After that, the mutant can replace the resident or infiltrate the system accordingly to the global behaviour of the ecological dynamics. Obviously, the coevolutionary outcome strictly depends on the ecological interactions. In this chapter we consider only the Lotka–Volterra type ecological model. If we consider another ecological model, the condition for qualitative behaviour of the rest point will change according to the new ecological dynamics.

In the first part of this chapter we considered arbitrary mutations, while in the second part, we followed the adaptive dynamics approach, supposing that the phenotypic space is a continuum and the interaction is a smooth function of the phenotypes. The presented dynamic approach can be considered as a generalization of the adaptive dynamics approach, since the cases investigated in the framework of adaptive dynamics (when the fitness of the mutant is strictly higher than that of the residents) are also considered in the dynamical approach, but the latter also deals with the case of neutral mutants (when the mutant has the same fitness as the resident). The natural generalization for a general multispecies coevolutionary system is also an important possibility.

Nevertheless, there is an essential difference between the approach presented here and the adaptive dynamics approach. Namely, since mutation is very rare, there is no reason to assume that evolution follows any gradient (cf. Vincent *et al.* [17]). In the adaptive dynamics approach it is stated that the evolutionary change in phenotype can be described by $\frac{dy}{dt} \sim \frac{\partial W(n,m,xy)}{\partial y}|_{x=y}$, see Dieckmann and Law [17]. In the gradient system it is assumed that mutation can find the best phenotypic direction, but it is supposed that mutants can "scan" all possible mutation directions, which means that all kinds of mutants occur. Thus at each non-ESS resident state we have a set of invadable phenotypes. As a result, a differential inclusion dynamics can be described by the qualitative behaviour of the coevolutionary process (cf. Cresmann and Hofbauer [6]).

Our approach is different, since we assume that mutation is rare, thus there is no reason to suppose that at all non-ESS states the "best" phenotypic change, i.e., the gradient, arises. In this view the final state of coevolution may depend strictly on random mutation events, if there are several different ESSs in the phenotype space.

Acknowledgments

This research supported by the Hungarian Scientific Research Fund OTKA (Project Number T037271, T049692). The final version of this manuscript was completed while the author was a Fellow at Konrad Lorenz Institute for Evolution and

Cognition Research. The author of this paper is a grantee of the János Bolyai Scholarship.

References

[1] Carr, J. (1981) *Application of Centre Manifold Theory.* Springer, Heidelberg.

[2] Cohen, Y., Vincent, T.L., Brown, J.S. (1993) A G-function approach to fitness minima, fitness maxima, evolutionary stable strategy and adaptive landscapes. *Evol. Ecol. Res.* **1**, 923–942.

[3] Cressman, R., Garay, J. (2003a) Evolutionary stability in Lotka–Volterra system. *J. Theor. Biol.* **222**, 233–245.

[4] Cressman, R., Garay, J. (2003b) Stability in N-species coevolutionary systems. *Theor. Pop. Biol.* **64**, 519–533.

[5] Cressman, R., Garay J., Hofbauer, J. (2001) Evolutionary stability concepts for N-species frequency dependent interactions. *J. Theor. Biol.* **211**, 1–10.

[6] Cressman, R., Hofbauer, J. (2005) Measure dynamics on a one-dimensional continuous train place: theoretical foundation for adaptive dynamics. *Theor. Pop. Biol.* **67**, 47–59.

[7] Dieckmann, U., Law R. (1996) The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34**, 579–612.

[8] Garay J. (2002) Many species partial adaptive dynamics. *BioSystems* **65**, 19–23.

[9] Garay J., Varga Z. (2000) Strict ESS for n-species system. *BioSystems* **56**, 131–137.

[10] Geritz, S.A.H., Metz, J.A.J., Kisdi, É., Meszéna, G. (1997) The dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett.* **78**, 2024–2027.

[11] Hammerstein, P., Riechert, S.E. (1988) Payoffs and strategies in territorial contests: ESS analyses of two ecotypes of the spider *Agelenopsis aperta. Evol. Ecol.* **2**, 115–138.

[12] Marrow, P., Dieckmann, U., Law, R. (1996) Evolutionary dynamics of predator-prey systems: an ecological perspective. *J. Math. Biol.* **34**, 556–578.

[13] Maynard Smith, J., Price, G.R. (1973) The logic of animal conflict. *Nature* **246**, 15–18.

[14] Meszéna, G., Kisdi É., Dieckmann, U., Geritz, S.A.H., Metz, J.A.J. (2001) Evolutionary optimisation models and matrix game in the unified perspective of adaptive dynamics. *Selection* **2**, 193–210.

[15] Vincent, T.L., Brown, J.S. (2005) *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics.* Cambridge Univ. Press, Cambridge, U.K.

[16] Vincent, T.L., Cohen, Y., Brown, J.S. (1993) Evolution via strategy dynamics. *Theor. Pop. Biol.* **44**, 149–176.

[17] Vincent, T.L., Van, M.V., Goh, B.S. (1996) Ecological stability, evolutionary stability and the ESS maximum principle. *Evol. Ecol.* **10**, 567–591.

Adaptive Dynamics, Resource Conversion Efficiency, and Species Diversity

William A. Mitchell
Department of Ecology and Organismal Biology
Indiana State University
Terre Haute, IN 47809, USA
wmitchell@indstate.edu

Abstract

Previous work on the theory of continuous evolutionary games and adaptive dynamics has shown that a species can evolve to an evolutionarily stable minimum on a frequency-dependent adaptive landscape (e.g., Brown and Pavlovic [7], and Abrams *et al.* [3],[4]). While such stable minima are convergent stable, they can be invaded by rare alternative strategies. The significance of such stable minima for biology is that they produce "disruptive selection" which can potentially lead to speciation (Metz *et al.* [20], Geritz *et al.* [16], Cohen *et al.* [11], and Mitchell [21]). Previous analyses of Lotka–Volterra competition communities indicate that stable minima and speciation events are more likely to occur when the underlying resource distribution is broader than the resource utilization functions of the competing species. Here, I present an analysis based on a resource-consumer model which allows individuals to adaptively vary resource use as a function of competitor density and strategy. I show that habitat specialization, stable minima, community invasibility, and sympatric speciation are more likely when individuals are more efficient at converting resources into viable offspring. Conversely, factors that inhibit conversion efficiency inhibit speciation and promote competitive exclusion. This model suggests possible links between species diversity and factors influencing the resource conversion efficiency, such as climate, habitat fragmentation, and environmental toxins.

Key words. G-function, disruptive selection, habitat specialization, species diversity, resource consumer model.

1 Introduction

Adaptive evolution can be represented on a surface called an adaptive landscape that graphs fitness as a function of evolutionary strategy. Adaptation then drives the evolution of the phenotype in the upslope direction of the adaptive landscape by selecting for strategies of higher fitness. When fitnesses of strategies are independent of the frequencies of strategies, the adaptive landscape is "rigid," and adaptation proceeds by "hill climbing" until a maximum is reached. This maximum is dynamically stable to perturbations of the equilibrium strategy. But when the fitness of a strategy is frequency dependent, the surface of the landscape becomes fluid, changing slope as the strategies of resident species evolve. In this case, there are at least two possible outcomes of adaptive dynamics. First, adaptive evolution may lead to an equilibrium which occurs at a maximum in the adaptive landscape, even though the position of the maximum may be frequency dependent. This maximum may be an *evolutionarily stable strategy* (Maynard Smith and Price [19]), in the sense that it is not invasible by alternative strategies. But the second possibility is that adaptive dynamics may produce an equilibrium which occurs at a minimum in the adaptive landscape (Eschel [14], Christiansen [10], Brown and Pavlovic [7], Abrams *et al.* [3], [4], Cohen *et al.* [11], Diekmann and Doebeli [12], and Kisdi [18]). While this type of equilibrium is "convergent stable" (Metz *et al.* [20]), it is not an ESS because it is invasible. Indeed, stable minima are invasible by rare strategies that are just slight deviations from the equilibrium strategy. Stable minima, as a result, are subject to disruptive selection that may lead to branching evolution by sympatric speciation, a possibility which has been theoretically examined for a variety of ecological circumstances (e.g., Diekmann and Doebeli [12], Kisdi [18], Doebeli and Diekmann [13], and Abrams [2]).

Recent investigations using Lotka–Volterra competition models demonstrate that resource competition can lead to stable minima and branching evolution under conditions of both symmetric and asymmetric competition (Diekmann and Doebeli [12], Kisdi [18], and Doebeli and Diekmann [13]). In these models, competition coefficients depend on the distance between phenotypes along a continuous axis of possible phenotypes. When resource distributions and resource utilization functions are Gaussian, explicit competition coefficients can be derived. One prediction of this model is that stable minima are more likely to result from adaptive dynamics when the variance of the resource utilization functions is narrow relative to the resource distribution. That is, branching evolution is more likely to occur when species' use of the resource spectrum is relatively specialized.

In previous theoretical work (Mitchell [21], Mitchell and Porter [23], and Mitchell *et al.* [22]), I used resource-consumer models of competition to analyze the adaptive evolution of habitat use. I considered competition communities in which strategies compete for resources on a continuum of habitat types, and individuals are allowed to allocate foraging effort adaptively. These models assumed that foragers pay a cost of exerting foraging effort (e.g., Abrams [1], and Mitchell *et al.* [22]). Different evolutionary strategies foraged most profitably in different habitat types, based on a difference in either foraging costs (Mitchell [21]) or resource encounter rates (Mitchell and Porter [23]). When a forager used a habitat for which it was not specialized, it paid a performance penalty, with the penalty increasing the further a habitat was from the preferred, or best, habitat. Based on habitat- and strategy-specific foraging costs and benefits, individual foragers flexibly allocated their foraging effort among habitat types. Competition in these models depended on both the density and behavior of competitors.

Analysis of these resource-consumer communities with behaviorally flexible foragers showed that stable minima were more likely to occur when: 1) the range of habitat types was broad, 2) travel cost among habitats was low, and 3) maintenance metabolic cost was low. The first of these results is similar to the prediction from Doebeli and Diekmann [13] discussed above; in both cases a broader range of resources (habitats) leads to stable minima. The second result has a similar basis; foragers are able to allocate more of their foraging effort to habitats in which they perform best when travel cost is low. This result reflects what behavioral ecologists have long known, that low travel cost allows active foragers to more selectively exploit resources (Stephens and Krebs [26]). Hence, at low travel cost, foragers in my model were effectively more specialized. The third result, that reduced maintenance cost makes stable minima more likely, is not as obvious. Reducing maintenance cost reduces the equilibrium resource levels across habitat types because foragers can cover their maintenance cost harvesting resources at a lower rate. The effect of lowering equilibrium resource densities when foragers have to pay a foraging cost is that it may no longer be profitable to exploit those habitats in which the forager pays a higher cost. Consequently, reduced maintenance cost results in behaviorally flexible foragers exploiting a narrower range of habitat types. Again, narrowing the range of habitat use by strategies makes stable minima more likely to occur.

While the resource-consumer models of competition on a habitat continuum discussed in the previous paragraph did permit individual foragers to flexibly allocate foraging effort, the function representing the trade-off in foraging ability in different habitats could not evolve—it was fixed throughout my previous analysis. In this analysis, I want to relax that restriction and allow evolution along a continuum of specialist-generalist, as well as evolution of the habitat to which the forager is specialized. Also, I will examine how the foragers' efficiency of converting harvested energy into offspring affects the adaptive landscape and resulting equilibrium. The reason I consider resource conversion efficiency is that it is a component of resource-consumer interactions that may vary predictably across broadly different environments (or climates). Thus, resource conversion efficiency may provide a handle on testing some predictions of adaptive dynamics.

2 The Model

2.1 Community Organization

Here, as in Mitchell and Porter [23], I will consider a community organized by the "discrete preferences" model of community structure (Rosenzweig [24]). In the discrete preference model, different strategies perform best in different habitats. In contrast, the "shared preference" model (Rosenzweig [24], and Mitchell [21]) is one in which all strategies perform the best in the same habitat, but differ in how well they perform there, and how rapidly their performance degenerates as habitat types become increasingly different. In the discrete preference model of

Mitchell and Porter, the penalty for foraging in habitats other than the preferred one increased nonlinearly (quadratically) with the difference between the two habitats as measured on a habitat axis. In that paper, the only evolutionary variable was the particular habitat type for which the strategy was specialized. But the penalty, or curvature of the quadratic, was not an evolutionary variable. It is reasonable to expect that evolution can also select along a specialist-generalist gradient, as well as along a habitat gradient. In the case of specialist-generalist evolution, two species may prefer the same habitat, but the specialist pays a lower foraging cost in that preferred habitat, with the trade-off that its foraging cost increases faster than the generalist's cost as habitats become more different than the preferred.

2.2 Constructing the Frequency-Dependent Fitness

Consider competing foragers that traverse an environment made up of patches of habitat drawn from a continuous distribution of habitat types. This seems reasonable for many real-life environments where habitat types are not finite but best defined by continuous variables (e.g., temperature, percent vegetation cover). As a forager moves through a patch it must spend some time in travel, even if it does not feed in the patch; this is the travel cost. If the individual feeds in the patch, then it must also spend some time harvesting the resources. While harvesting the resources, the forager experiences diminishing returns due to resource depletion. At some point the forager quits foraging and leaves the patch at some resource level (probably not zero). After the forager leaves the patch, it continues to travel through and feed in other patches. Patches are able to renew when they are not being foraged, which means that the initial resource level encountered by a forager depends on the density and behavior of resident foragers. The fitness of a forager depends on the rate at which it acquires resources as well as its ability to covert resources into offspring.

I use the above scenario to construct frequency-dependent fitnesses, and a G-function (Vincent *et al.* [28], and Cohen *et al.* [11]). Then I use the G-function to generate adaptive landscapes in order to investigate the role of resource conversion efficiency in promoting stable minima.

2.3 Foraging Cost

Denote the habitat type by the variable, z, which ranges in value from z_{min} to z_{max}. Let individuals be characterized by the evolutionary strategy pair (u, s), where u indicates the habitat for which the strategy is specialized, and s indicates the degree of specialization. When an individual forages in a patch of habitat, z, it pays a per unit time foraging cost, $c(u, s, z)$. So if $u = z'$, then (u, s) pays its lowest foraging cost in z'. The evolutionary variable, s, indicates the degree of specialization. Degree of specialization here refers to the efficiency of using the preferred habitat; I assume a "Jack of all trades, Master of none" principle, so that increased efficiency of using a preferred habitat (increased specialization) comes

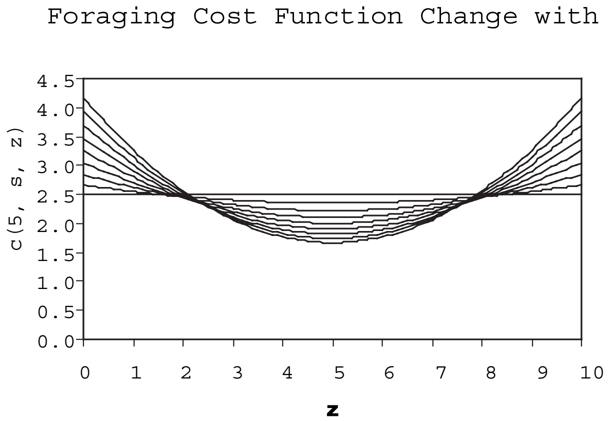


Figure 1: A series of foraging cost curves. Each curve denotes the foraging cost as a function of habitat, z, for a particular value of the evolutionary variable s. The value of s determines the degree of specialization on the preferred habitat. When $s = 0$, the cost curve is a horizontal line, indicating an absolute generalist strategy. For larger values of s, the cost curve reaches a minimum at the preferred habitat, and increases quadratically in increasingly different habitats. A specialization-generalization trade-off is evident, in that neither the specialist nor the generalist has the lowest cost across all habitats.

at the expense of reduced efficiency of using a nonpreferred habitat. If $s = 0$, the strategy is a complete generalist, and pays the same foraging cost in all habitats. If $s > 0$, then there is a penalty for not foraging in the habitat $z' = u$, and the penalty is greater for larger s.

I represent the foraging cost as

$$c(u, s, z) = \frac{k_0}{k_2 + s} + sk_1(u - z)^2, \tag{1}$$

where k_0, k_1, and k_2 are constants. The first term on the right-hand side represents the base foraging cost for a given degree of specialization, s. The second term represents the penalty paid for foraging in a patch of habitat different from the preferred habitat, as indicated by the variable u. Increasing the degree of specialization, s, reduces the base cost, but with the trade-off of a higher penalty for foraging in patches of unpreferred habitat. The combination of variables, (u, s), evolves in this model. A series of foraging cost curves is shown in Figure 1.

For the present analysis, I gave the constants in equation (1) the values $\mathrm{k}_0 = 1$, $\mathrm{k}_0 = 0.5$, and $\mathrm{k}_0 = 0.4$. This choice for the constants resulted in the specialization-generalization trade-off as discussed in the previous paragraph.

2.4 Resource Dynamics

I consider the situation in which patches of habitat are small enough so that they contain no more than one forager at a time. Furthermore, I assume that depletion occurs at a much faster rate than renewal, so that while the forager is in a patch of habitat z, the resource density declines at the rate

$$\frac{dR(z)}{dt} = -aR(z), \tag{2}$$

where a is the "area of search" that converts resource density into encounter rate of the forager with resources. Resource renewal occurs between patch visits by foragers. I assume for simplicity that resources renew according to the "re-growth" equation (Turchin [27]), but are kept harvested well below their carrying capacity. As a result, the resource renewal in a habitat patch is approximately linear in time,

$$\frac{dR(z)}{dt} = b. \tag{3}$$

Average resource densities across habitat types depend on the foraging behavior and density of the resident strategy or strategies, where a resident is defined as a strategy at an ecological (but not necessarily evolutionary) equilibrium. The foraging behavior of a resident individual determines the density of resources it leaves in a patch of habitat z, when it departs in search of another patch. In the following section I describe how a forager chooses this quitting resource density (or giving-up density; Brown [6]). The initial density of resources the next forager encounters is a function of the quitting resource density of the previous forager and the time which the patch has been allowed to renew. The renewal time is inversely proportional to the frequency of patch visits by resident foragers, which depends on the density of foragers actively searching, as opposed to foraging in patches. Since each forager spends a proportion of its time searching, and the remainder foraging, the density of searching foragers is the total number of foragers multiplied by the proportion of time spent searching. Let *prop* represent the proportion of resident individual engaged in searching, and let $R_{quit}(z)$ be the quitting resource density for a resident foraging in habitat z. Then the expected initial number of resources in a patch just encountered is

$$R_{init}(\hat{u}, \hat{s}, N, z) = R_{quit}(\hat{u}, \hat{s}, z) + \frac{b}{N * prop}. \tag{4}$$

2.5 The Average Rate of Energy Gain, Resource Conversion Efficiency, and Fitness

Traversing a habitat patch takes time, even if resources are not harvested. The time required to simply traverse a patch of habitat is the search time, T_s. If the forager harvests resources from the habitat patch, depleting those resources to a density $R_{quit}(z)$, then the patch foraging time can be calculated by solving equation (2) and rearranging terms. The solution to equation (2) describes an exponential decline of resources as a function of foraging time, t_f, so $R_{quit}(z)$ is related to $R_{init}(z)$ by

$$R_{quit} = R_{init} e^{-at_f}. \tag{5}$$

Rearranging terms to solve for foraging time yields

$$t_f(z) = (1/a)\ln(R_{init}/R_{quit}). \tag{6}$$

The net energy gain from the patch is energy acquired in the patch minus the energy spent foraging. The energy gained is the product of the total resources consumed from the patch and the per capita resource value. The energy expended in foraging is the product of the foraging time and the activity cost of foraging. The average rate of energy gain is the average net energy gain from a patch, divided by the sum of the search time and the average time spent in a habitat patch. Let $p(z)$ be the probability density function of habitat types, and let v denote the per capita resource value. Then the average rate of net energy gain of an individual with strategy (u, s) foraging in an environment where resource levels are determined by a resident strategy $(\hat{u}, \hat{s})$ is

$$\begin{aligned} &F(u, \hat{u}, s, \hat{s}, N) \\ &= \frac{\int_{z_{\min}}^{z_{\max}} p(z)\left[v(R_{init}(\hat{u}, \hat{s}, N, z) - R_{quit}(u, s, z)) - \frac{c(u,s,z)}{a}\ln\left(\frac{R_{init}(\hat{u},\hat{s},N,z)}{R_{quit}(u,s,z)}\right)\right]dz}{T_s + \frac{1}{a}\int_{z_{\min}}^{z_{\max}} p(z)\ln\left(\frac{R_{init}(\hat{u},\hat{s},N,z)}{R_{quit}(u,s,z)}\right)dz}. \end{aligned} \tag{7}$$

The fitness of the strategy (u, s) is equal to the product of the average rate of net energy gain and the conversion efficiency from energy to offspring minus the density-independent mortality rate. Let Q denote the resource conversion efficiency, and let M be the density-independent mortality rate, in which case the fitness of (u, s) is

$$\begin{aligned} &W(u, \hat{u}, s, \hat{s}, N, Q) \\ &= Q\frac{\int_{z_{\min}}^{z_{\max}} p(z)\left[v(R_{init}(z) - R_{quit}(z)) - \frac{c(u,s,z)}{a}\ln\left(\frac{R_{init}(z)}{R_{quit}(z)}\right)\right]dz}{T_s + \frac{1}{a}\int_{z_{\min}}^{z_{\max}} p(z)\ln\left(\frac{R_{init}(z)}{R_{quit}(z)}\right)dz} - M. \end{aligned} \tag{8}$$

Or, using $F(\cdot)$ for the average rate of energy gain,

$$W(u, \hat{u}, s, \hat{s}, N, Q) = QF(u, \hat{u}, s, \hat{s}, N, MC) - M. \tag{9}$$

One could also include an explicit maintenance metabolic cost, *mc*, representing the energy devoted to maintenance and replacement of tissues. In this case $W(\cdot)$ would be

$$W(u, \hat{u}, s, \hat{s}, N, Q) = Q(F(u, \hat{u}, s, \hat{s}, N, MC) - mc) - M. \tag{10}$$

In this case, the adaptive landscape can vary with *mc*. And in fact, *mc* is what Mitchell [21] and Mitchell and Porter [23] considered. But the effect of increasing Q is qualitatively similar to decreasing *mc* (or M); i.e., they all result in a decrease in $F(\cdot)$ at an ecological equilibrium. So for this study, I will use equation (8), and consider the role of Q influencing the adaptive landscape. For the results reported below, $M = 0.1$.

2.6 Foraging Behavior

Foraging behavior does not evolve in my model. Rather, I assume here that foraging behavior has already evolved to appropriately allocate effort based on foraging benefits and costs that may occur under a variety of conditions, including the possession of different morphological and/or physiological features that may influence those costs and benefits. The evolutionary variables, u and s, are such features affecting foraging costs and benefits.

The foraging behavior of an individual in a habitat, z, can be represented by the quitting resource density, $R_{quit}(u, s, z)$. In each habitat, z, the quitting resource density that maximizes the individual's net rate of energy gain, subject to the forager's evolutionary variables (u, s), and given the distribution of resource densities tailored by the N residents characterized by the evolutionary variables $(\hat{u}, \hat{s})$, is (Charnov 1976 and Brown [5])

$$R^*_{quit}(u, s, z) = \left(\frac{1}{av}\right)(c(u, s, z) + F(u, \hat{u}, s, \hat{s}, N, Q)). \tag{11}$$

2.7 *G*-function

Substituting the optimal quitting resource density for an individual with evolutionary variables (u, s) foraging in an environment with N residents with evolutionary variables $(\hat{u}, \hat{s})$ gives the G-function (Vincent *et al.* [28] and Cohen *et al.* [11]),

$$G(u, \hat{u}, s, \hat{s}, N, Q)$$
$$= Q\frac{\int_{z_{\min}}^{z_{\max}} p(z)\left[v(R_{init}(\hat{u}, \hat{s}, N, z) - R^*_{quit}(u, s, z)) - \frac{c(u,s,z)}{a}\ln\left(\frac{R_{init}(\hat{u},\hat{s},N,z)}{R^*_{quit}(u,s,z)}\right)\right]dz}{T_s + \frac{1}{a}\int_{z_{\min}}^{z_{\max}} p(z)\ln\left(\frac{R_{init}(\hat{u},\hat{s},N,z)}{R^*_{quit}(u,s,z)}\right)dz} - M. \tag{12}$$

Let $\hat{N}$ be the density of foragers with the evolutionary strategy $(\hat{u}, \hat{s})$, such that

$$G(\hat{u}, \hat{u}, \hat{s}, \hat{s}, \hat{N}, Q) = 0. \tag{13}$$

In other words, $\hat{N}$ is the equilibrium population density for strategy $(\hat{u}, \hat{s})$. Then, we can determine whether a rare strategy (u, s), can invade a resident strategy, $(\hat{u}, \hat{s})$, at an ecological equilibrium, by determining whether $G(u, \hat{u}, s, \hat{s}, \hat{N}, Q)$ is positive (can invade) or negative (cannot invade). For any particular resident strategy there is a surface that maps (u, s) to $G(u, \hat{u}, s, \hat{s}, \hat{N}, Q)$. This surface is the adaptive landscape. In the following section, I examine the effects of changing resource conversion efficiency. In particular, I show how resource conversion efficiency (Q) can influence the evolution of specialization and the occurrence of a stable minimum in the adaptive landscape, which may promote sympatric speciation.

3 Results

To study the effect of resource conversion efficiency (Q) when both the degree and location of specialization can evolve, I examine how a particular resident strategy, $(\hat{u}, \hat{s})$, at ecological equilibrium, $\hat{N}$, affects the adaptive landscape. I first consider the case in which the resident is an absolute generalist ($\hat{s} = 0$). Because the absolute generalist pays the same foraging cost in all habitats, its fitness does not change with u, which means that $G(\hat{u}, u, 0, 0, \hat{N}, Q) = 0$, for all values of u in the interval, $(u_{\min}, u_{\max})$. However, if $\hat{s} = 0$, we can ask whether there is selection to increase the value of s; in other words, is there selection pressure on an absolute generalist to become at least somewhat specialized with respect to some habitat. This question can be answered by looking at the slope of the adaptive landscape at $(\hat{u}, \hat{s})$, where $\hat{s} = 0$. If the slope is positive, then there is selection for some degree of specialization, and if small variations in s are produced as a result of mutation, then adaptive dynamics should result in the evolution of specialization. As Figure 2 indicates, the slope of the adaptive landscape in the s direction at $\hat{s} = 0$ depends on $\hat{u}$. Even though all values of $\hat{u}$ yield identical fitness to a resident generalist, small deviations from $\hat{s} = 0$ are more likely to yield positive fitness, and hence be able to invade, if $\hat{u}$ is closer to the intermediate value (e.g., $u = (z_{min} + z_{max})/2$) than if $\hat{u}$ is near either extreme ($\hat{u} = \mathrm{z}_{min}$ or $u = z_{max}$). Figure 2a shows a positive slope, and hence selection for increased specialization when $\hat{u} = 5$, which is the intermediate value ($z_{min} = 0$; $z_{max} = 10$). However, when $\hat{u} = 3$, the slope is negative (Figure 2b), and specialization will not result from adaptive dynamics.

Previous work on a similar model (Mitchell and Porter [23]), but in which the degree of specialization was fixed and habitat preference, u, was allowed to evolve, indicated that the value of u at a convergent stable equilibrium was in fact the intermediate value, $\hat{u} = (z_{\min} + z_{\max})/2$. The intuition behind this result is that at the intermediate value of u an individual pays the lowest foraging cost averaged over all habitats. So, when there is a single resident species, $\hat{u}$ converges to this intermediate value. If, on the other hand, a resident is specialized for one or the other end of the habitat spectrum, it does not benefit as much from its specialization. Consequently, it would be harder to initiate the evolution of specialization if the resident were to specialize near an extreme of the habitat spectrum. Figure 2 shows how, for the case of $\hat{s} = 0$, two different values of $\hat{u}$ (3 and 5), generate different adaptive landscapes in the vicinity of the resident strategy. When $\hat{u} = 5$, adaptive dynamics would drive the evolution of an increase in $\hat{s}$, whereas when $\hat{u} = 3$, adaptive dynamics would prevent the evolution of specialization.

What happens to the evolution of specialization as we increase the resource conversion efficiency, Q? An increase in Q from 0.033 to 0.10 can produce a qualitative change in the adaptive dynamics. For example, when $\hat{s} = 0$ and $\hat{u} = 3$ (Figure 2c), increasing Q changes the slope of the adaptive landscape at $\hat{s} = 0$ and $\hat{u} = 3$ from negative to positive with respect to the s axis (the slope with respect to the u axis remains zero). Thus, increasing the resource conversion efficiency increases the likelihood that adaptive dynamics can drive the evolution of specialization.

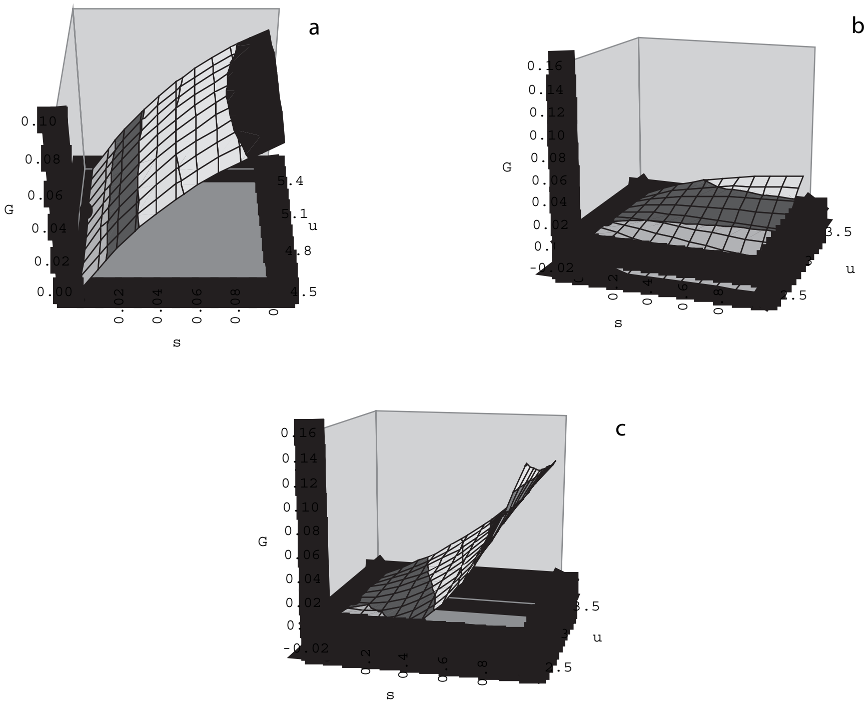


Figure 2: a) The adaptive landscape generated by a resident when $(\hat{u}, \hat{s}) = (5, 0)$. The resident is an absolute generalist because the specialization variable equals zero, but the fact that the habitat preference is 5 (i.e., for the "intermediate" habitat) means that small deviations in the specialization variable above zero will be selected for, as indicated by the positive slope of the adaptive landscape at the resident's position (indicated by the black dot). b) The adaptive landscape by a resident $(\hat{u}, \hat{s}) = (3, 0)$. In this case, the habitat preference is biased away from the intermediate value of 5, which means that deviations in the specialization variable are not as fit as the resident. Because the slope at the position of the resident (indicated by the black dot) is slightly negative, specialization will not evolve. c) The same resident as in Figure 2b, but with a greater resource conversion efficiency (0.1 as opposed to 0.033, in Figures 2a and 2b). The greater resource conversion efficiency increases the slope in the s direction so that it is positive. In this case, adaptive dynamics would drive increased specialization.

Whenever the resident possesses any degree of specialization; that is, whenever $\hat{s} > 0$, an inspection of the adaptive landscape reveals selection for higher u if $\hat{u} < (z_{min} + z_{max})/2$, and lower u if $\hat{u} > (z_{min} + z_{max})/2$. Thus, $\hat{u}$ converges on $(z_{min} + z_{max})/2$, so long as $\hat{s} > 0$. While $\hat{s}$ also converges to an equilibrium value, its value is not as simple to characterize as $\hat{u}$. Figure 3a shows the adaptive landscape generated by a resident strategy $(\hat{u}, \hat{s})$ at a convergent equilibrium, when the resource conversion efficiency is 0.033. This convergent equilibrium also happens to be a maximum in the adaptive landscape, and hence an evolutionarily

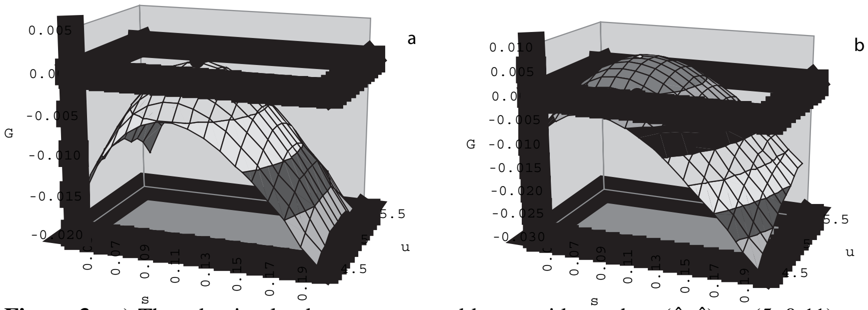


Figure 3: a) The adaptive landscape generated by a resident when $(\hat{u}, \hat{s}) = (5, 0.11)$, at a "low" resource conversion efficiency ($Q = 0.033$). The resident (black dot) resides on a stable maximum in the adaptive landscape. This maximum is an ESS. b) The adaptive landscape generated by a resident when $(\hat{u}, \hat{s}) = (4.8, 0.12)$, a small perturbation from the ESS displayed in Figure 3a. The perturbation resides on a hillside in the adaptive landscape, with the upslope in the direction of the ESS.

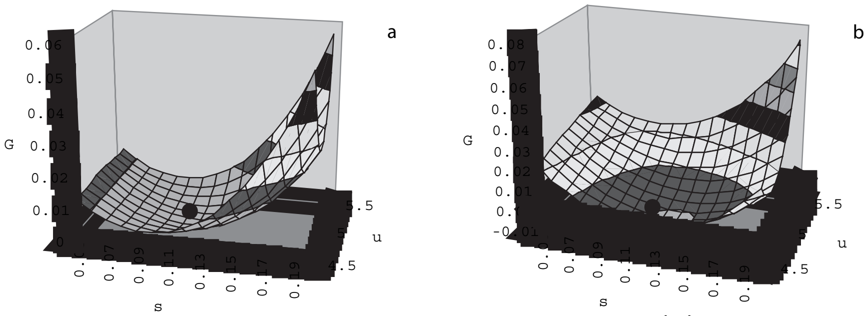


Figure 4: a) The adaptive landscape generated by a resident when $(\hat{u}, \hat{s}) = (5, 0.12)$, at a "high" resource conversion efficiency ($Q = 0.010$). The resident (black dot) resides on a stable minimum in the adaptive landscape. Also, the value of the specialization variable has increased slightly from 0.11 to 0.12 at the convergent stable equilibrium. b) The convergent stability of this equilibrium is seen in an example of a perturbation of the resident to $(\hat{u}, \hat{s}) = (4.8, 0.13)$. The perturbed resident generates an adaptive landscape that slopes upward in the direction of the equilibrium $(\hat{u}, \hat{s}) = (5, 0.12)$.

stable strategy (ESS). The convergent stability is evident by the fact that deviations from $(\hat{u}, \hat{s})$ produce a change in the adaptive landscape that select for (u, s) closer to the equilibrium, $(\hat{u}, \hat{s})$ (Figure 3b).

Increasing Q from 0.033 to 0.10 qualitatively changes the adaptive landscape and the equilibrium. One result of the increase in Q is to increase the degree of specialization, $\hat{s}$, from 0.11 to 0.12 at the equilibrium. The other result is that the equilibrium is now a minimum (Figure 4a). The adaptive surface is bowl-shaped, with the minimum occurring at $(\hat{u}, \hat{s}) = (5, 0.12)$. Figure 4b shows the adaptive landscape when the resident values for the evolutionary variables are perturbed

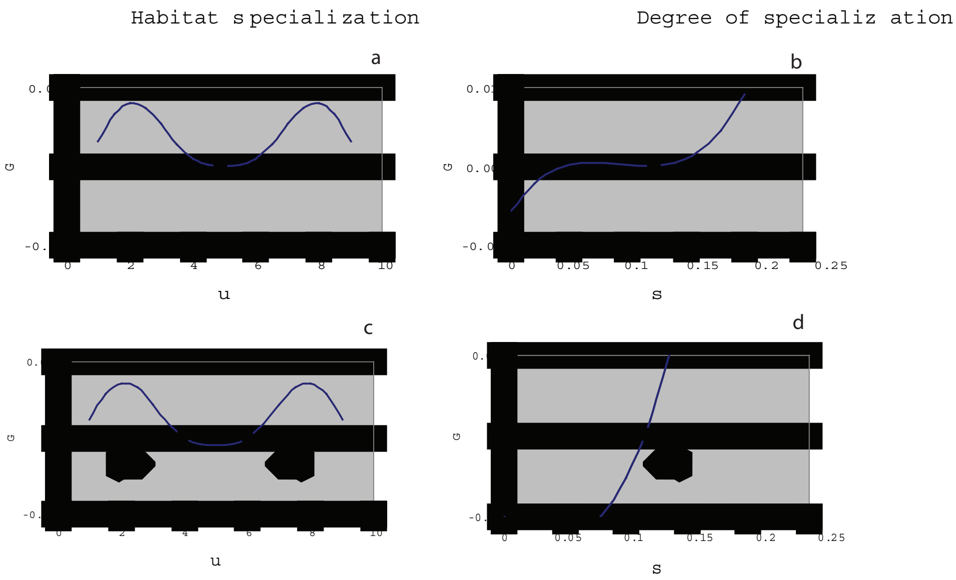


Figure 5: Disruptive and divergent adaptation at the stable minimum $(\hat{u}, \hat{s}) = (5, 0.12)$. Figures 5a and b show the minimum on the u and s axes, respectively. Figure 5c illustrates divergent selection in the case of invasion by a mutant u strategy. Figure 5d shows that once the two strategies are evolving away from each other on the u axis, there is selection for higher values of habitat specialization.

to $(\hat{u}, \hat{s}) = (4.8, 0.13)$. In this example, the upslope direction of the adaptive landscape is in the direction of (5, 0.12), or back toward the original equilibrium, suggesting that the equilibrium is a stable minimum.

The stable minimum at $(\hat{u}, \hat{s}) = (5, 0.12)$ is invasible (Figure 5). A new strategy with a slightly different value of either u or s may become established by either sympatric speciation or immigration. In either case, divergent selection drives the two strategies apart from one another on the adaptive landscape. Figure 5c illustrates this process when the invading strategy possesses a value of u different from the original resident. Once the second strategy invades, the minimum on the s axis disappears and adaptive evolution selects for higher values of habitat specialization (Figure 5d).

If for some reason habitat specialization is unable to evolve to higher values after the invasion of a second strategy, then the outcome of divergent selection is a 2-species community in which each species resides at a maximum on the u axis (Figure 6).

If s can adaptively increase, then the outcome of divergent selection is a 2-species community in which the species reside at minima on the u axis, and a maximum on the s axis (Figures 7a and b). This community is invasible by variant strategies of u. If a third species invades, the value of s is further increased by adaptive

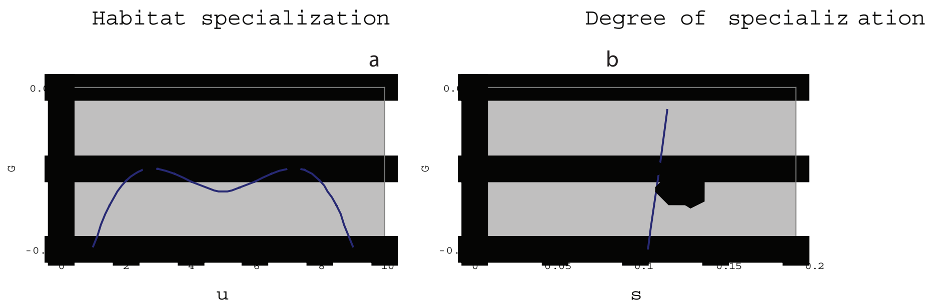


Figure 6: These figures illustrate the outcome of divergent evolution on the u axis from the stable minimum $(\hat{u}, \hat{s}) = (5, 0.12)$ if habitat specialization cannot evolve on the s axis. In this case, the evolutionarily stable community comprises two species.

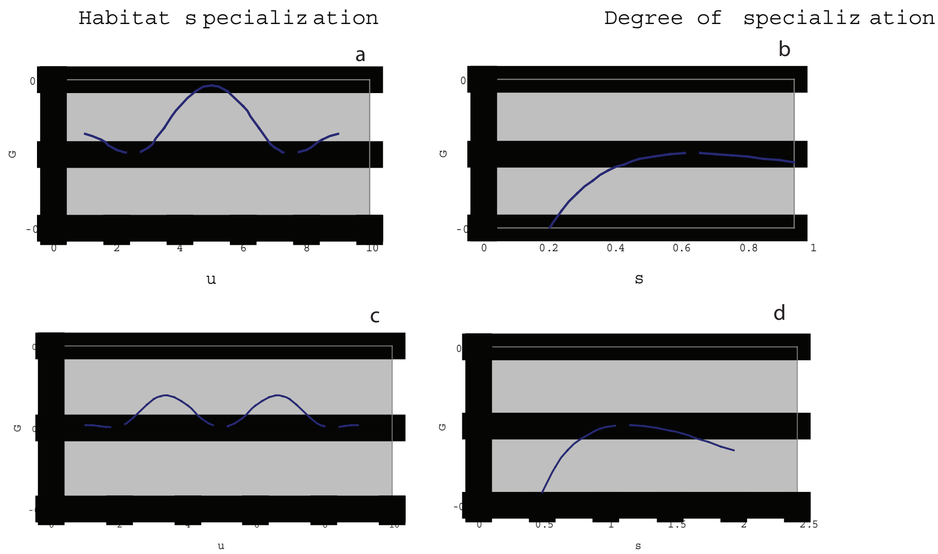


Figure 7: Figures 7a and b show the outcome of divergent evolution on both the u and s axes from the stable minimum $(\hat{u}, \hat{s}) = (5, 0.12)$ when habitat selection can evolve. The two resulting equilibria $((\hat{u}_1, \hat{s}_1) = (2.5, 0.65)$ and $(\hat{u}_2, \hat{s}_2) = (7.5, 0.65))$ reside at minima on the u axis (7a), and a maximum on the s axis. As a result of letting habitat specialization evolve to a higher value, the 2-species community is invasible. Figures 7c and d illustrate the outcome of evolution after the invasion of the two species community, when habitat specialization is allowed to evolve. The new 3-species community is still invasible, due to the evolution of increased habitat specialization.

evolution, and the outcome of adaptive evolution is a 3-species community. Each of the three species reside at minima on the u axis and a maximum on the s axis. As a result of increased specialization, this 3-species community is still invasible by speciation or immigration.

4 Discussion

For the resource-consumer model that I have analyzed, a stable minimum is more likely to occur on frequency-dependent adaptive landscapes when individuals are more efficient at converting resources to offspring. The explanation for this result depends on a series of cause and effect. First, the direct effect of increasing resource conversion efficiency is to reduce the average rate of energy intake required to sustain residents at ecological equilibrium. This means that the equilibrium level of resources across habitats decreases as the resource conversion efficiency increases. As resource levels across habitats decrease, it becomes less profitable for individuals to exploit habitats that are very different from their preferred habitat. This latter result is due to the existence of the foraging cost individuals must pay in order to exploit a habitat and the reduced resource level in the habitat. Because individuals in this resource-consumer model have flexible foraging behavior and can allocate their foraging effort among habitats, a reduction in habitat resource level will therefore tend to make foragers exploit a narrower range of habitats. Under the usual assumption of adaptive dynamics, that mutation is arbitrarily small, this behavioral specialization has an effect similar to the phenotypic specialization modeled by Doebeli and Dieckmann [13]—it can produce a stable minimum.

In order for the mechanism described in the previous paragraph to actually work, the behavioral specialization must be predicated on some degree of phenotypic specialization that enforces a trade-off in the foraging performance in different habitats. In the absence of any such trade-off (or any degree of specialization), there would be no adaptive reason for foragers to alter their relative use of habitats with a change in resource conversion efficiency. If there were only a slight degree of phenotypic specialization, then there would be a small behavioral response to increasing resource conversion efficiency. In order for a stable minimum to result, there needs to be enough of a phenotypic specialization to drive the behavioral specialization. In principle, it is possible that increasing the resource conversion efficiency could have resulted in an adaptive decrease in phenotypic specialization. If this had been the case, increasing resource conversion efficiency would have had little or no effect of producing a stable minimum. But this was not the case, as increasing resource conversion efficiency enhanced the evolution of phenotypic specialization, which in turn drove the behavioral specialization with the result of a stable minimum.

In the preceding analysis I investigated the effects of resource conversion efficiency on adaptive landscapes, but similar results would hold if I had considered the effects of maintenance cost or density-independent mortality. All of these parameters influence the equilibrium level of resources across habitats, so they all influence behavioral specialization, and the evolution of phenotypic specialization, with similar effects on the adaptive landscape. Of course, the direction of these effects may differ. Increasing resource conversion efficiency has qualitatively the same effects as decreasing maintenance cost or decreasing the density-independent mortality rate. Also, while I considered habitat specialization with respect to the foraging cost, similar results are obtained in the model when habitats differ with respect to the energy value of resource items (v in equations 7 and 8), or the per item encounter rate with resources (a in equations 2, 5, 6, 7, and 8).

One of the reasons why stable minima are interesting to biologists is the possibility that they may lead to branching evolution (e.g., Diekmann and Doebeli [12]). This possibility makes the study of stable minima a promising tool for biologists who want to understand patterns of species diversity (Rosenzweig [24]) and adaptive radiations (Schluter [25]). This is especially true if stable minima are more likely to be produced in some environments than others. The possibility that adaptive dynamics are influenced by the environment seems reasonable, given that the outcomes of some species interactions depend on environmental parameters such as temperature (Jiang and Morin [17]). Here is where resource-consumer models can be valuable. Because resource-consumer models typically include parameters, such as resource conversion efficiency, which vary with environmental conditions, they may provide a way of linking environmental conditions to the adaptive dynamics of branching speciation. If resource-consumer adaptive dynamics suggest that evolutionary branching is more likely when resources are more readily converted into offspring, then ecologists can use that result to make testable predictions about the relationship between species diversity and environmental parameters that influence resource conversion efficiency. I believe the model I have presented is one simple example of how to approach the goal of using resource-consumer models to understand patterns of biodiversity.

Acknowledgments

I wish to thank the organizers of the 11th International Symposium on Dynamic Games and Applications, held in Tucson, AZ, and in particular, Tom Vincent and Joel Brown for organizing the "Speciation and Extinction" session. I gratefully acknowledge the constructive suggestions of two anonymous reviewers.

References

[1] Abrams, P.A. 1991. Life history and the relationship between food availability and foraging effort. *Ecology*. 72:1242–1252.

[2] Abrams, P.A. 2003. Can adaptive evolution or behavior lead to diversification of traits determingin a trade-off between foraging gain and predation risk? *Evol. Ecol. Res*. 5:653–670.

[3] Abrams, P.A., Y. Harada, and H. Matsuda. 1993. On the relationships between quantitative genetics and ESS models. *Evolution*. 47:982–985.

[4] Abrams, P.A., H. Matsuda, and Y. Harada. 1993. Evolutionary unstable fitness maxima and stable fitness minima of continuous traits. *Evol. Ecol*. 7:465–487.

[5] Brown, J.S. 1988. Patch use as an indicator of habitat preference, predation risk, and competition. *Behavioral Ecology and Sociobiology*. 22:37–47.

[6] Brown, J.S. 1989. Desert rodent community structure: a test of four mechanisms of coexistence. *Ecological Monographs* 59:1–20.

[7] Brown, J.S. and N.B. Pavlovic. 1992. Evolution in heterogeneous envioronments: Effects of migration on habitat specialization. *Evol. Ecol*. 6:360–382.

[8] Brown, J.S. and T.L. Vincent. 1992. Organization of predator-prey communities as an evolutionary game. *Evolution.* 46:1269–1283.

[9] Charhov, E.L. 1976. Optimal Foraging, the marginal value theorem. *Theoretical Population Biol.* 9: 129–136.

[10] Christiansen, F.B. 1991. On conditions for evolutionary stability for a continuously varying character. *Amer. Nat.* 138:37–50.

[11] Cohen, Y., T.L. Vincent, and J.S. Brown. 1999. A G-function approach to fitness minima, fitness maxima, evolutionarily stable strategies and adaptive landscapes. *Evol. Ecol. Res*. 1:923–942.

[12] Diekmann, U. and M. Doebeli. 1999. On the origin of species by sympatric speciation. *Nature* 400:354–357.

[13] Doebeli, M. and U. Dieckmann. 2000. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *Amer. Nat.* 156:S77–S101.

[14] Eschel, I. 1983. Evolutionary and continuous stability. *J. Theor. Biol*. 103:99–111.

[15] Geritz, S.A.H., E. Kisdi, G. Meszena, and J.A.J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol*. 12:35–57.

[16] Geritz, S.A.H., J.A.J. Metz, E. Kisdi, and G. Meszena. 1997. The dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett*. 78:2024–2027.

[17] Jiang, Lin, and P.J. Morin. 2004. Temperature-dependent interactions explain unexpected responses to environmental warming in communities of competitors. *J. Anim. Ecol*. 73:560–576.

[18] Kisdi, E. 1999. Evolutionary branching under asymmetric competition. *J. Theor. Biol*. 197:149–162.

[19] Maynard Smith, J., and G.R. Price. 1973. The logic of animal conflict. *Nature* 246:15–18.

[20] Metz, J.A.J., S.A.H. Geritz, G. Meszena, F.J.A. Jacobs, and J.S. van Heerwaarden. 1996. pp. 183–231 *in* S.J. van Strien and S.M. Verduyn Lunel,

eds. *Stochastic and Spatial Structures of Dynamical Systems*. North-Holland, Amsterdam.

[21] Mitchell, W.A. 2000. Limits to species richness in a continuum of habitat heterogeneity: an ESS approach. *Evol. Ecol. Res.* 2:293–316.

[22] Mitchell, W.A., Z. Abramsky, B.P. Kotler, B. Pinshow, and J.S. Brown. 1990. The effect of competition on foraging activity in desert rodents: theory and experiments. *Ecology* 71:844–854.

[23] Mitchell, W.A. and W.P. Porter. 2001. Foraging games and species diversity. *Annales Zoologici Fennici* 38:89–98.

[24] Rosenzweig, M.L. 1995. *Species Diversity in Space and Time*. Cambridge University Press, Cambridge.

[25] Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford University Press. Oxford.

[26] Stephens, D.W. and J.R. Krebs. 1986. *Foraging Theory*. Princeton University Press. Princeton, NJ.

[27] Turchin, P. 2003. *Complex Population Dynamics*. Princeton University Press, Princeton, NJ.

[28] Vincent, T.L., Y. Cohen, and J.S. Brown. 1993. Evolution via strategy dynamics. *Theor. Pop. Biol.* 44:149–176.

Evolutionarily Stable Relative Abundance Distributions

Tania L. S. Vincent
Alaska Pacific University, 4101 University Drive
Anchorage, AK 99508, USA
octo8@ptialaska.net

Thomas L. Vincent
Department of Aerospace and Mechanical Engineering
University of Arizona
Tucson, AZ 85721-0119, USA
vincent@u.arizona.edu

Abstract

Modified versions of a well-known model for coexistence are used to examine the conditions that determine the relative abundance of species that are in an evolutionarily stable state. Relative abundance is a term used to refer to the ranking of the number of individuals present within trophically similar species in an ecosystem. We use the G-function approach to understand why relative abundance relationships take the form so often found in field data. We assume that the ecosystem is at or near an evolutionary equilibrium and seek evolutionarily stable strategies to identify a coalition of individual species. In order to have a coalition greater than one, the G-function must produce frequency dependence, implying that the fitness of any given individual depends on the strategies used by all individuals in the population. This is an essential element of the evolutionary game. Otherwise, evolution would drive the population to a single strategy (i.e., a coalition of one) that is an optimal or group fitness strategy. We start with a classical version of the Lotka–Volterra competition equation that is not frequency dependent and make it frequency dependent in three different ways, thus allowing for the modeling of relative abundance. The first two methods involve a single resource niche and rely on modifications of the competitive effects to provide for a coalition of two or more. These models yield relative abundance distribution curves that are generally convex and are not typical of most field data. The third method creates several resource niches, and the simulated results generally create concave curves that are much closer to the field data obtained for natural systems.

1 Introduction

It has been well over half a century since ecologists first noticed that the relative abundance of species within communities shows recurring patterns [i.e., few species are very abundant, some are common, but most are rare (Fisher, Corbet and Williams [5])]. This tantalizingly predictable observation begs an explanation. In the beginning, this pattern was explained by a set of species-abundance models [i.e., broken stick, log normal, log series, and geometric series (Whittaker [22])]. Although statistical in nature, these models could be considered in terms of resource partitioning where the abundance of a species is equivalent to the portion of the niche space it has preempted (Magurran [10]). According to Jones, Munday and Caley [7], "Various models predict different species abundance relationships (e.g., geometric series, log series, log normal, broken stick), depending on the underlying community-level process." They further note that, "In communities where one or few limiting resources control community structure, typically a species-poor community, species rank-abundance relationships should follow either a geometric or log series (May [13]). In species-rich communities, where large numbers of interacting factors control the ecology of participating species, central tendency should generate a log-normal distribution of species abundance (May [13], Sugihara [18]). Finally, if a single resource that controls community structure is shared equitably among species, a 'broken stick' distribution is expected (May [12]). The majority of species abundance relationships reported have been log-normal (Sugihara [18])."

At the turn of this century, a radically different way of explaining this pattern was presented by Hubbell [6]. Hubbell's "neutral model" explained this pattern as a result of a zero-sum game where individuals are drawn at random to fill a set number of community slots that are vacated upon an individual's death. This model assumes that individuals of all ecologically similar species are competitively equal, and that the chances of an individual filling a slot depend solely on the relative abundance of the species it belongs to, not on any special dispersal or competitive characteristics. By means of this "ecological drift" and meta-population dynamics, Hubbell's model fits actual relative abundance data to an amazing degree. In essence, Hubbell states that the relative abundance of communities do not need to be determined by a niche-based theory of competitive interactions nor of trade-offs in dispersal strategies.

Hubbell's revelation left the ecological world reeling. One of the main reasons for this is that his model dismisses several ecological and evolutionary tenants. First, Darwin tells us that evolution (and speciation) is nonneutral because heritable traits matter in the struggle for existence (an excellent example of this can be found in Benkman [1]). Second, immigration has been demonstrated as nonneutral [e.g., Fargione, Brown and Tilman [4], because established species more strongly inhibit introduced species more similar to themselves in resource-use traits. Finally, we know that commonness and rarity are nonneutral because ecological traits can affect distribution and abundance [e.g., (Murry, Tharall, Gill and Nicotra [15]); (Kelly and Woodward][8])]. Thus, it is not surprising that there have been several rebuttals to Hubbell's neutral model. For example, Magurran and Henderson [11] suggest that the left-skew pattern is the result of two different ecological processes acting at the same time, and it has been suggested by Fargione *et al.* [4] that the left-skew pattern is the result of ecological trade-offs; McGill [14] has even proposed that the original lognormal distribution is a better fit to the ecological data than the neutral model.

However, if we are ever to find a truly unified theory of species abundance, we need to find a model that (1) uses the rules of natural selection, (2) incorporates the ecological traits of individuals, (3) can vary parameters to predict any kind of relative abundance distribution, and (4) can not only tell us why that distribution occurs, but also which species will be dominant and which species will be rare. In this chapter we use a modeling tool that does all four. It is based on the G-function method of Vincent and Brown [20]. This method is applied to the relative abundance problem, and the results obtained are compared with some typical relative abundance data.

The G-function method has some assumptions in common with those of Hubbell [6]. First, the ecological community is identical to Hubbell's: "a group of trophically similar, sympatric species that actually or potentially compete in a local area for the same or similar resources." These are called evolutionarily identical individuals, and we can describe all such individuals using a single G-function. Second, as in [6], the probability of giving birth among these evolutionarily identical types is identical and they must likewise "obey exactly the same rules of ecological engagement." However, unlike [6], in our model individual traits (strategies) can influence the outcome of the "ecological engagement" in the manner of an evolutionary game. Furthermore, our model does not have any meta-communities and their attendant dynamics.

As a starting point, we make use of the well-known Lotka–Volterra competition model. It turns out that the standard version of this model is not helpful for producing evolutionarily stable relative abundance distributions. However, we show how some simple modifications can turn it into a useful tool. Not only do we obtain evolutionarily stable strategy (ESS) solutions composed of many species, but the modified models provide insight into why plots of the distribution of relative abundance among coevolved species take on the shapes that they do. In particular, we will show that modifications of the basic model will allow us to mimic data[1] such as that shown in Figures 1 and 2 where frequency (i.e., fraction of individuals belonging to each species) is plotted versus rank ($1 =$ highest rank, etc.).

The general shape of the relative abundance curves depicted in these figures is apparently quite common. For example, Jones *et al.* [7] show that abundance-rank curves for butterfly fishes in 10 different communities world wide all tend to have this same lognormal-like shape.

[1] Data come from USGS Patuxent Wildlife Research Center, 2003. North American Breeding Bird Survey internet data set, 9 Dec. 2003 (http://www.mbr-pwrc.usgs.gov/bbs/). It has been shown Link and Sauer [9] that these data can be used to calculate the "within route" relative abundances without bias.

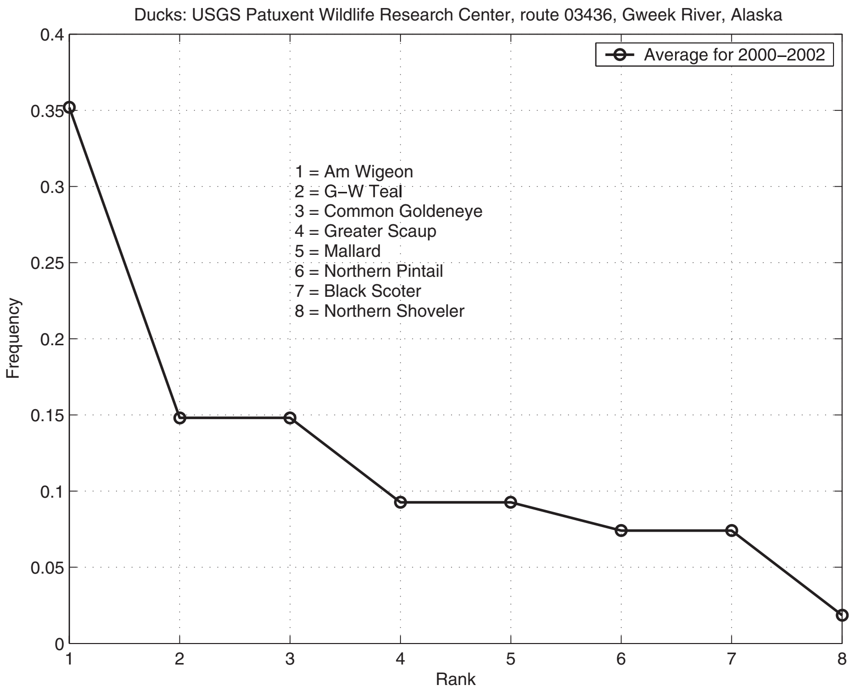


Figure 1: Relative abundance of duck.

2 Lotka–Volterra Competition Model

The Lotka–Volterra competition model has been extensively studied in an evolutionary setting Case [3], Rummel and Roughgarden [17], Vincent *et al.* [21]. Since this model is a useful paradigm for evolutionary processes, we will re-examine it here as a possible model for the study of relative abundance. In this basic model the strategies are scalars and the fitness function for a given evolutionarily identical type i, is given by

$$H_i(\mathbf{u}, \mathbf{x}) = r - \frac{r}{K(u_i)} \sum_{j=1}^{n_x} \alpha(u_i, u_j) x_j, \tag{1}$$

where n_x is the total number of different types currently in the community, r is the intrinsic growth rate common to all types, $K(u_i)$ is the carrying capacity of type i, and $\alpha(u_i, u_j)$ is the competitive effect of type j using strategy u_j on the fitness of individuals of type i using strategy u_i. The fitness function for any type i may be

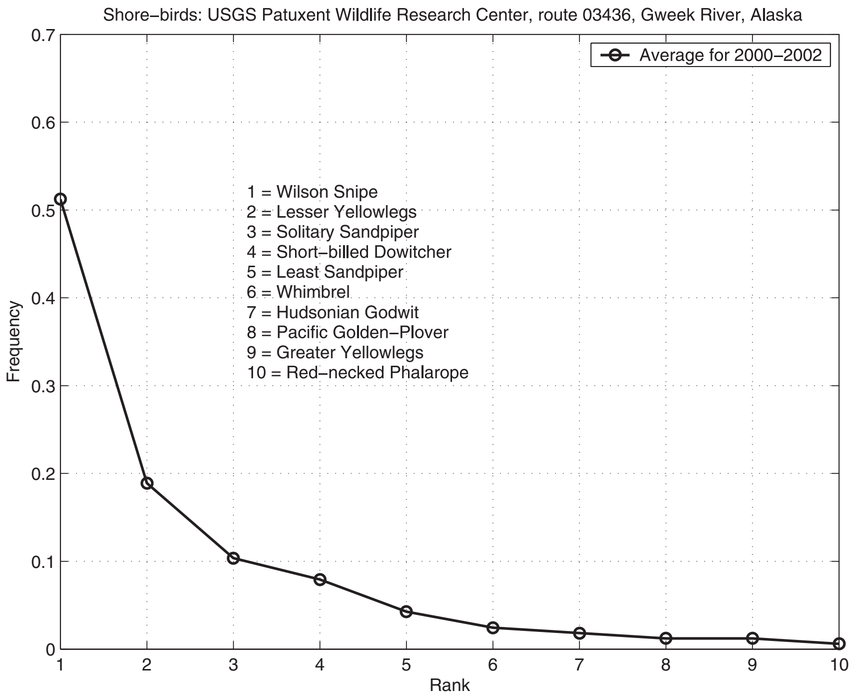


Figure 2: Relative abundance of shorebirds.

obtained from the following G-function by replacing v with u_i:

$$G(v, \mathbf{u}, \mathbf{x}) = r - \frac{r}{K(v)} \sum_{j=1}^{n_x} \alpha(v, u_j) x_j. \tag{2}$$

We will make use of the K and α functions used in Roughgarden [16]. That is,

$$K(v) = K_m \exp\left[-\frac{v^2}{2\sigma_k^2}\right] \tag{3}$$

and

$$\alpha(v, u_j) = \exp\left[-\frac{\left(v - u_j\right)^2}{2\sigma_a^2}\right]. \tag{4}$$

This model will be of little consequence for the study of relative abundance unless it is able to produce stable equilibrium populations composed of more than one species. In addition, the equilibrium populations of interest are those that are both ecologically stable and evolutionarily stable. As it stands, the above model is able to produce an ecological equilibrium of many species, but only a single species ESS (Vincent et al. [21]).

3 Limitations with the Basic Model

As a starting point, we assign the following parameters:

$$\begin{aligned} K_m &= 100 \\ r &= .25 \\ \sigma_\alpha &= 2 \\ \sigma_k &= \sqrt{2} \end{aligned}$$

with initial conditions $\mathbf{x} = [\ 20\ \ 20\ ]$ and strategies $\mathbf{u} = [\ -1\ \ 1.5\ ]$. Integrating the population dynamic equations[2] results in the ecologically stable equilibrium solution of $\mathbf{x}^* = [\ 65.53\ \ 26.98\ ]$. In fact, any number of different one- or two-species equilibrium solutions are possible. The two-species solutions require that the strategies be on either side of $u = 0$. While these solutions imply that ecological coexistence is possible, such solutions are not evolutionarily stable, except for the one-species solution of $u = 0$. Insight into the process is obtained by plots such as those of Figure 3. The left-hand frame of Figure 3 is a plot of the adaptive landscape (G vs. v for fixed $\mathbf{u}$ and $\mathbf{x}$). The right-hand frame is a plot of $K(v)$ (denoted K-curve) and $\sum \alpha(v, u_j)x_j^*$ (denoted S-curve). This latter plot helps in understanding how the adaptive landscape is formed. It follows from (2) that equilibrium ($G = 0$) is obtained only when the two curves cross or are tangent. The two curves intersect at values of v corresponding to the chosen strategies given above. Since the K-curve is higher than the S-curve between the two intersections, we see from the adaptive landscape that any strategy between the two values used in the simulation will result in a higher fitness. It is because of this that the chosen strategies are not evolutionarily stable. If we allow each of the strategies to evolve, they will both "climb" the adaptive landscape until they arrive at the peak ESS solution $u = 0$.

The adaptive landscape along with its associated component curves give some insight into the nature of the solution obtained. The K-curve is related to the amount of resource available, at a given strategy, for producing biomass. The S-curve is related to the utilization of the resource by the species strategies actually used. In those regions where the K-curve is above the S-curve, the resource is underutilized, allowing another "more efficient" species with a higher fitness to invade. In this example there exists an ESS coalition of one with

$$u_1 = 0, \quad x_1^* = 100.$$

The left-hand frame of Figure 4 illustrates that this solution satisfies the ESS maximum principle (Vincent and Brown [20]). Since the S-curve is always higher than the K-curve except at the point of tangency, the resource is totally utilized, resulting in a biomass greater than the sum of the biomasses obtained in the two-species

[2] When the strategies are known, the equilibrium solution for $\mathbf{x}$ can also be solved for directly from the population dynamic equations.

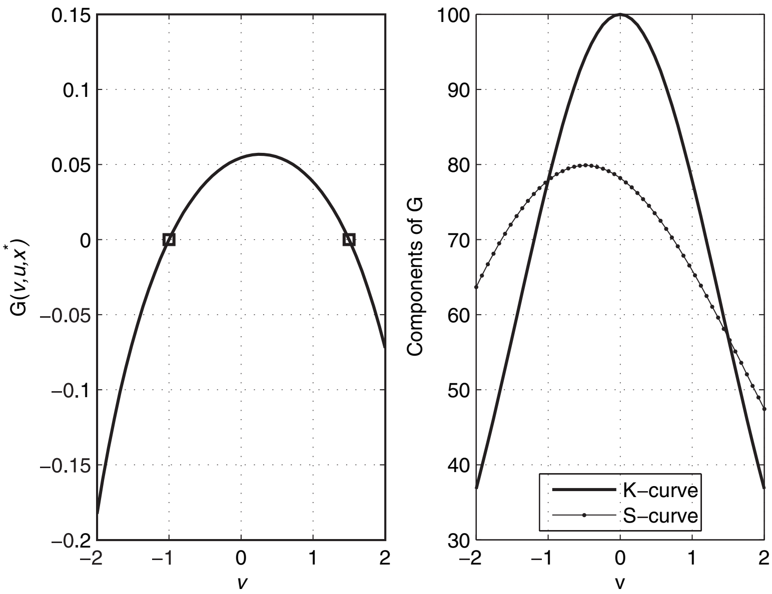


Figure 3: Coexistance of two species that are at an ecologically stable equilibrium, but are not evolutionarily stable.

ecological equilibrium case. Because the resource is totally utilized by the ESS solution, there is no chance of having evolutionarily stable solutions greater than one with this model except for the singular case that occurs when the K-curve and S-curve coincide.

Increasing σ_k flattens the K-curve and when $\sigma_k = 2$ it coincides with the S-curve, resulting in a flat adaptive landscape that is everywhere zero. In other words, in this unique and no doubt rare situation, the ESS maximum principle is satisfied by all values of u_i, allowing for any number of strategies. If we further increase σ_k to say $\sigma_k = \sqrt{5}$ we obtain the solution $u_1 = 0$, $x_1^* = 100$, which is no longer an ESS as illustrated in Figure 5. In fact, as illustrated in the left panel, equilibrium is at a local minimum on the adaptive landscape. The reason for this is that the S-curve is always lower than the K-curve except at the point of tangency. Prior to reaching equilibrium, the adaptive landscape is shaped more like panel one of Figure 4 with Darwinian dynamics driving evolution toward a peak on the landscape. However, as the strategy and population size approach equilibrium, the peak turns into a valley. As a consequence, if a second species is introduced at any other strategy it is able to invade and coexist with the first strategy on an ecological time scale. However, on an evolutionary time scale, the two species will eventually evolve to $\mathbf{u} =[\ 0.9005\ \ -0.9005\ ]$ at the population densities $\mathbf{x} =[\ 55.3265\ \ 55.3265\ ]$, each at a local minimum on the adaptive landscape.

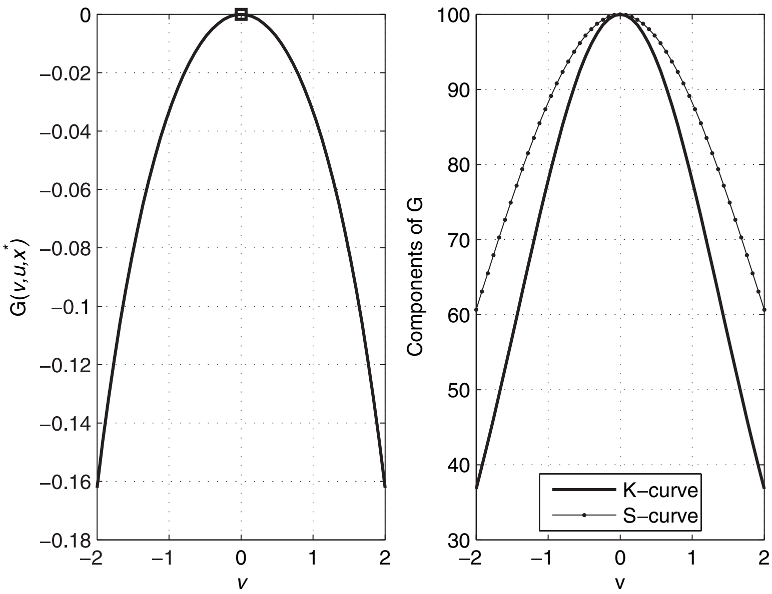


Figure 4: The standard model has only an ESS coalition of one.

This allows for the invasion of a third species, etc. Further increasing σ_k does not change the general nature of the fact that none of the solutions obtained with $\sigma_k > 2$ will be evolutionarily stable.

In conclusion, we find that with this model, after ruling out the singular case, no evolutionarily stable coexistence with other strategies is possible. However, it has been shown that by constraining the range of the strategy variable, it is possible to obtain an ESS coalition of two (Vincent et al. [21]). A more useful approach for studying relative abundance is to simply change the basic model.

4 Introduction of a Fixed Competitive Cost

A slight modification of the previous model allows for coexistence that makes for a convenient, yet simple model to study relative abundance. The modification involves adding a positive constant to the competition term. Without this term, choosing a v a long way from u_j results in a vanishingly small competition effect [see (4)]. Inclusion of this factor guarantees that there will be some minimum competitive effect with the introduction of any species, no matter how different that species may be from the focal strategy u_j. While this term is unaffected by strategy, this additional competitive effect is affected by both the number of species

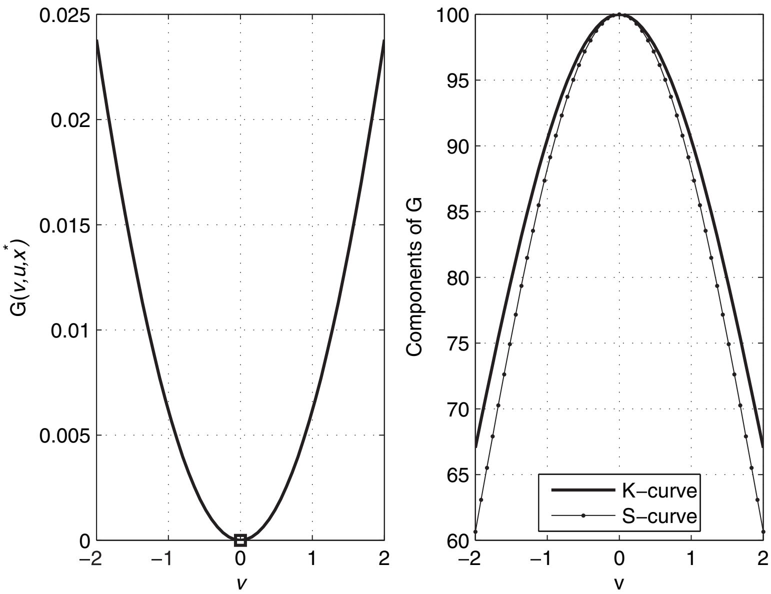


Figure 5: Increasing $\sigma_k = \sqrt{5}$ results in evolution to a local minimum.

in the community and the number of individuals in each species. For example, even if plants are competing for a limiting resource such as nitrogen, the addition of another species, no matter how different the nitrogen-garnering strategy, will still take up space, unusable to the original inhabitants.

This first modification of our model involves adding a competitive constant to α we will call α_{comp},

$$\alpha(v, u_j) = \alpha_{\text{comp}} + \exp\left[-\frac{\left(v - u_j\right)^2}{2\sigma_a^2}\right]. \tag{5}$$

By setting $\alpha_{\text{comp}} = 0$, we return to the classical model where there is either an ESS coalition of one with $u_1 = 0$ or no ESS. Figures 4 and 5 illustrate this situation. When $\alpha_{\text{comp}} > 0$, interference competition is added to the classic Lotka–Volterra model, resulting in an effect for a community of species comparable to logistic growth in a single species. By allowing a "carrying capacity" logistic effect from other species in addition to a species' own carrying capacity, the sigma constant makes any particular species unable to reach as high a carrying capacity in the presence of other species. This makes sense in a bounded world and harkens back to Hubbell's [6] idea that everyone takes up "space" of some kind even if it is

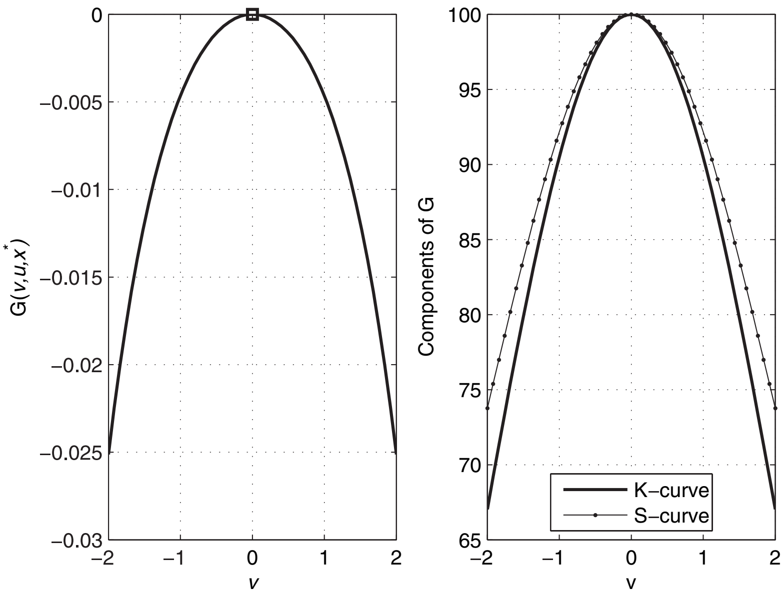


Figure 6: With $\alpha_{\text{comp}} = 0.5$, the model once again yields an ESS coalition of one.

not directly related to the evolutionarily important competitive interactions among species within a community.

If we now set $\alpha_{\text{comp}} = 0.5$, keeping $\sigma_k = \sqrt{5}$, the evolutionary local minimum we previously had with this solution becomes a local maximum, as illustrated in the left panel of Figure 6, resulting in an ESS coalition of one with

$$u_1 = 0, \quad x_1^* = 66.67.$$

The S-curve is now computed using $\alpha_{\text{comp}} + \sum \alpha\left(v, u_j\right) x_j^*$. We see from the right panel that the nonzero α_{comp} term keeps the S-curve above the K-curve. Furthermore, because of the α_{comp} term, a single species cannot utilize all of the evolutionarily important resource,[3] resulting in a smaller equilibrium population. In essence, another resource (e.g., space) constrains a single species before its population reaches carrying capacity on the evolutionarily important resource (e.g., nitrogen).

A fixed competitive cost opens up opportunities for coexistence. For example, keeping $\alpha_{\text{comp}} = 0.5$, increasing $\sigma_k = \sqrt{8}$, and allowing only one species to evolve, we again end up with a landscape similar to Figure 5. However, if a second

[3]Such a resource affects the evolution of a strategy.

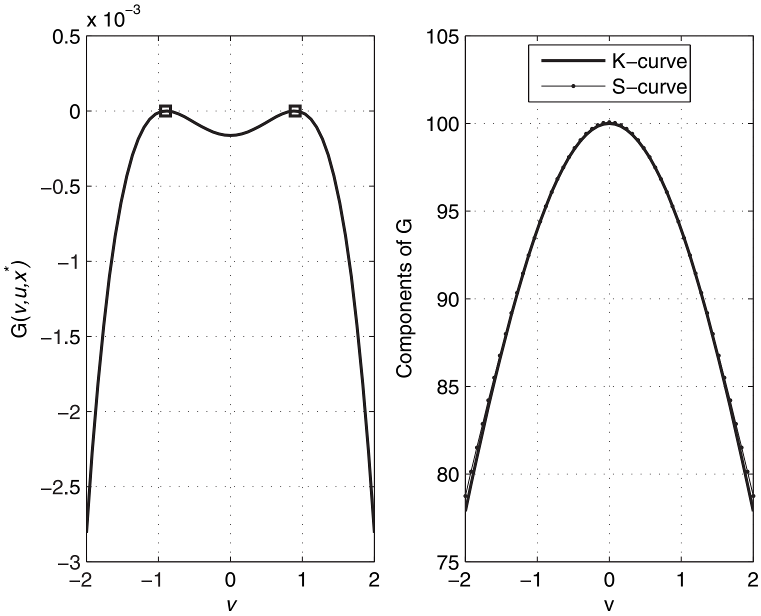


Figure 7: An ESS coalition of two.

species is introduced, we obtain an ESS coalition of two with

$$\mathbf{u} = [\ -0.9013\ \ 0.9013\ ]$$
$$\mathbf{x}^* = [\ 35.65\ \ 35.65\ ] \quad \text{sum}(\mathbf{x}^*) = 71.30.$$

In this case the relative abundance of each species is equal, as might be expected from the symmetry of the model. The results shown in Figure 7 illustrate the fact that the ESS solution is a result of very similar shapes of the K-curve and the S-curve. As a result the adaptive landscape is flattened in the neighborhood of the solutions. Note that the total biomass has now increased.

Increasing $\sigma_k = \sqrt{15}$ results in the following ESS coalition of three:

$$\mathbf{u} = [\ -2.012\ \ 0.000\ \ 2.012\ ]$$
$$\mathbf{x}^* = [\ 27.14\ \ 26.75\ \ 27.14\ ] \quad \text{sum}\left(\mathbf{x}^*\right) = 81.03$$

with a further increase in total biomass.

As σ_k increases the G-function gets flatter and flatter due to the very small differences in the K-curve and S-curve (See Figure 8). This results in relative abundance distributions that have small downward stair-steps paired two species at a time. This results from the competitive symmetry inherent in the model.

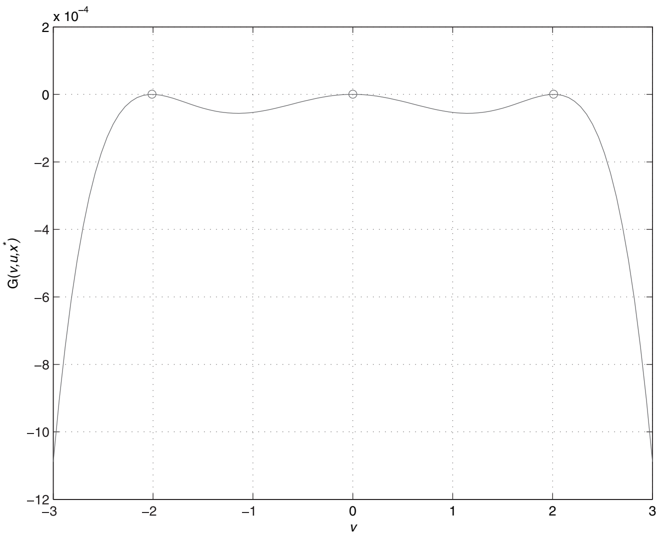


Figure 8: Increasing σ_k increases the number of strategies in the coalition.

5 Nonsymmetric Competition

Frequency-dependent selection results when a particular strategy's payoff depends on all strategies used within the community. It enters the Lotka–Volterra competition model used here through the competition coefficients. As noted by Brown and Vincent [2], when $\alpha_{\text{comp}} = 0$ and the competition function is symmetric about u_j with respect to v as in (4), frequency dependence (but not density dependence) is lost as a factor in determining an ESS and coalitions greater than one do not occur. That is why we used $\alpha_{\text{comp}} > 0$ above. When we introduce competitive asymmetry into our model below, it turns out that we still need to include a positive α_{comp} term in order to allow ESS coalitions greater than one to occur. Borrowing from Brown and Vincent [2], we introduce asymmetry along with a fixed competitive cost into (4) so that

$$\alpha(v, u_j) = \alpha_{\text{comp}} + \exp\left[-\frac{\left(v + \beta - u_j\right)^2}{2\sigma_a^2}\right], \tag{6}$$

where β shifts the value of v required for maximum competitive effect.

Continuing to use $\alpha_{\text{comp}} = 0.5$, but now with $\beta = 2$, Figures 9, 10, and 11 illustrate a coalition of one, two, and three obtained by using $\sigma_k = \sqrt{5}$, $\sqrt{8}$, and $\sqrt{20}$. These results along with several others are summarized in Table 1. All parameters used in the integration runs are the same as above except for σ_k as noted. As σ_k is increased to obtain ESS coalitions of three and beyond, the difference

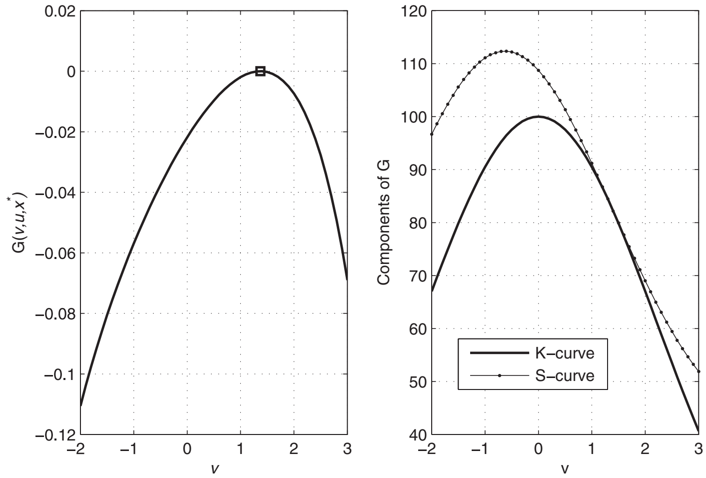


Figure 9: An ESS coalition of one with nonsymmetric α distribution.

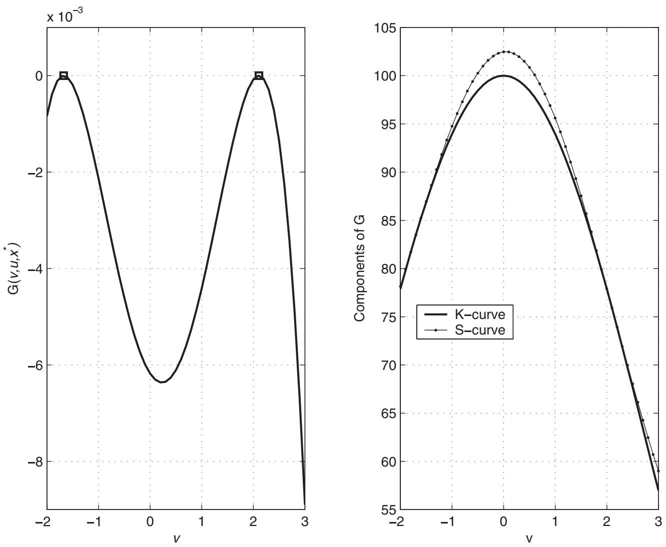


Figure 10: A nonsymmetric ESS coalition of two.

between the K-curve and S-curve become less and less, resulting in a very flat G-function for the larger coalitions.

Figure 12 illustrates the relative abundance plots obtained using the equilibrium population values obtained in Table 1. It is quite clear that the familiar lognormal-like distribution is not obtained. As in the symmetric case, the $K(v)$

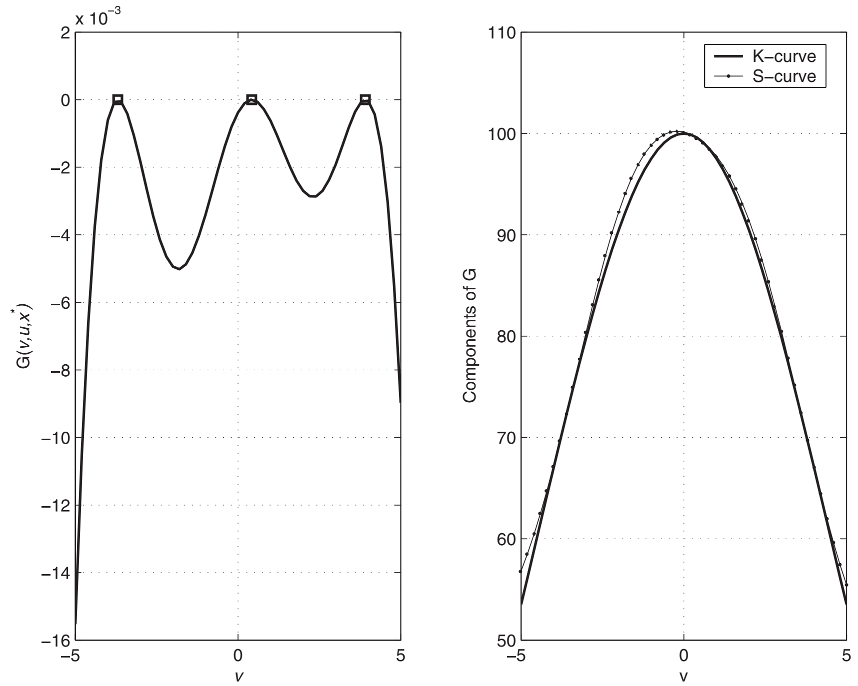


Figure 11: A nonsymmetric ESS coalition of three.

Table 1: Nonsymmetric model. $K_m = 100$, $r = 0.25$, $\sigma_\alpha = 2$, $\alpha_{\text{comp}} = 0.5$, $\beta = 2$.

σ_k	$\frac{x^*}{u^*}$	sum($\mathbf{x}^*$)
$\sqrt{5}$	$\frac{x^*=74.90}{u^*=1.370}$	74.90
$\sqrt{8}$	$\frac{x^*=[6.635 \;\; 65.35]}{u^*=[-1.674 \;\; 2.109]}$	71.99
$\sqrt{20}$	$\frac{x^*=[3.626 \;\; 43.47 \;\; 39.45]}{u^*=[-3.693 \;\; 0.4215 \;\; 3.912]}$	86.55
$\sqrt{30}$	$\frac{x^*=[21.41 \;\; 42.09 \;\; 31.78]}{u^*=[-2.431 \;\; 1.296 \;\; 4.826]}$	95.28
$\sqrt{40}$	$\frac{x^*=[0.4390 \;\; 30.72 \;\; 39.92 \;\; 27.35]}{u^*=[-5.857 \;\; -1.482 \;\; 2.157 \;\; 5.695]}$	98.43
$\sqrt{50}$	$\frac{x^*=[10.30 \;\; 32.55 \;\; 36.76 \;\; 24.34]}{u^*=[-4.922 \;\; -0.9046 \;\; 2.714 \;\; 6.268]}$	104.0
$\sqrt{60}$	$\frac{x^*=[17.15 \;\; 33.45 \;\; 34.40 \;\; 22.14]}{u^*=[-4.189 \;\; -0.3181 \;\; 3.277 \;\; 6.838]}$	107.1
$\sqrt{70}$	$\frac{x^*=[22.17 \;\; 33.85 \;\; 32.55 \;\; 20.47]}{u^*=[-3.545 \;\; 0.2556 \;\; 3.836 \;\; 7.402]}$	109.0
$\sqrt{80}$	$\frac{x^*=[4.182 \;\; 24.61 \;\; 33.18 \;\; 30.68 \;\; 19.08]}{u^*=[-7.339 \;\; -3.034 \;\; 0.7103 \;\; 4.276 \;\; 7.844]}$	111.7

function has only one hump, and species must seek out the small excesses in resources that are available between the K-curve and S-curve as the curves bend around this single hump. Thus, for ESS coalitions greater than one, we do not obtain large differences in equilibrium population size between adjacent species when

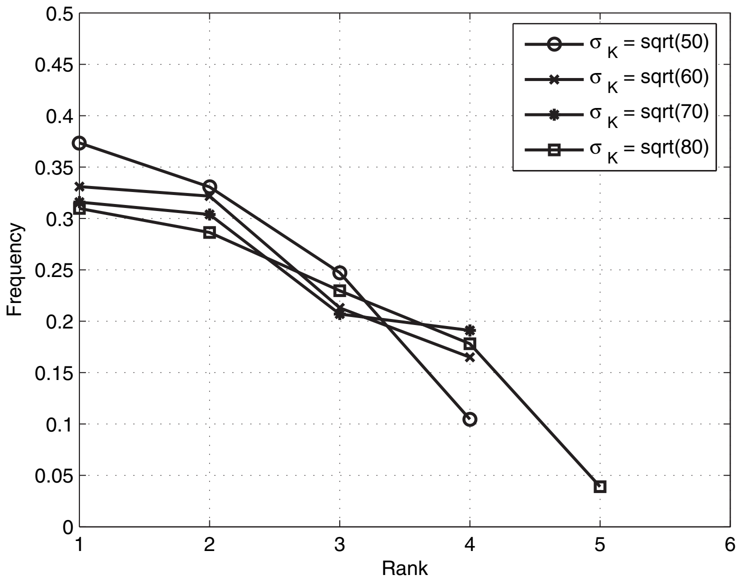


Figure 12: Relative abundance curves for the nonsymmetric case.

ranked by size. However, we have eliminated the "step-like" drop in species abundance that is obtained in the symmetric case. In summary, a frequency-dependent Lotka–Volterra competition model, whether competition is symmetric or not, does not in itself create a lognormal-like curve.

6 Multi-humped carrying capacity

By introducing multi-humps to the $K(v)$ function an ESS coalition no longer needs to seek out small differences between the K-curve and the S-curve. We create obvious resource niches for the species to occupy by adding additional humps to the K function. For example,

$$K(v) = K_{m_1} \exp\left[-\frac{v^2}{2\sigma_k^2}\right] + K_{m_2} \exp\left[-\frac{(v-5)^2}{2\sigma_k^2}\right] + K_{m_3} \exp\left[-\frac{(v+5)^2}{2\sigma_k^2}\right] \quad (7)$$

introduces three humps in the K function, one at zero and two at ± 5. The introduction of humps puts frequency dependence back into the model even with $\beta = 0$ and $\alpha_{\text{comp}} = 0$. Using the same competition coefficient as given by (6),with the parameter values

$$K_m = [\ 100 \ \ 70 \ \ 60\ ]$$
$$\beta = 0, \alpha_{\text{comp}} = 0, \sigma_k = \sqrt{2}, r = 0.25, \sigma_\alpha = 2,$$

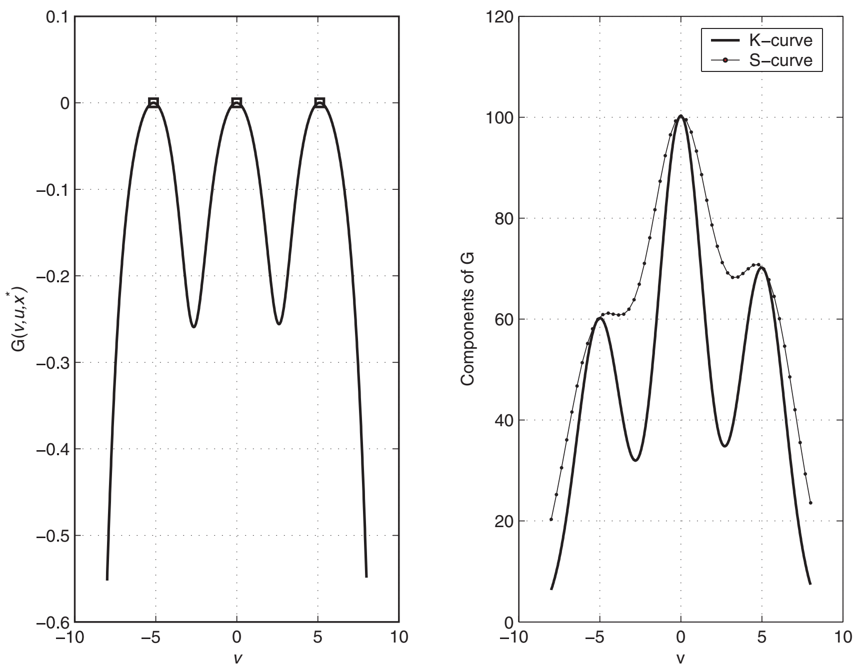


Figure 13: An ESS coalition of three using $\alpha_{\text{comp}} = 0$, symmetric α, and nonsymmetric K.

we obtain an ESS coalition of three

$$\mathbf{u} = \begin{bmatrix} -5.141 & -0.0090 & 5.121 \end{bmatrix}$$
$$\mathbf{x}^* = \begin{bmatrix} 56.28 & 95.68 & 66.32, \end{bmatrix},$$

as illustrated in Figure 13. Here, the ordering of the relative abundance is the same as the ordering of the peak heights. The highest peak has the greatest relative abundance, and the second and third highest peaks have the second and third levels of relative abundance. However, this result is not true in general. For example, introducing five humps

$$K(v) = K_{m_1} \exp\left[-\frac{v^2}{2\sigma_k^2}\right] + K_{m_2} \exp\left[-\frac{(v-2)^2}{2\sigma_k^2}\right] + K_{m_3} \exp\left[-\frac{(v+2)^2}{2\sigma_k^2}\right]$$
$$+ K_{m_2} \exp\left[-\frac{(v+5)^2}{2\sigma_k^2}\right] + K_{m_3} \exp\left[-\frac{(v-5)^2}{2\sigma_k^2}\right] \tag{8}$$

with

$$K_m = [\ 100\ \ 70\ \ 60\ ]$$
$$\beta = 0, \alpha_{\text{comp}} = 0, \sigma_k = 1, r = 0.25, \sigma_\alpha = 2$$

results in an ESS coalition of five

$$\mathbf{u} = [-5.013 \;\; -1.667 \;\; 0.000 \;\; 1.667 \;\; 5.013]$$
$$\mathbf{x}^* = [27.17 \;\; 5.534 \;\; 54.68 \;\; 5.534 \;\; 27.17],$$

as illustrated in Figure 14. The two lowest peaks have the second highest relative abundance. Because of the symmetry in K, we also obtain symmetry in relative abundance so that a relative abundance plot has "step-like" drops.

It is easy to remove any symmetry in K. For Figure 15 the K function is given by

$$K(v) = K_{m_1} \exp\left[-\frac{v^2}{2\sigma_k^2}\right] + K_{m_2} \exp\left[-\frac{(v-2)^2}{2\sigma_k^2}\right] + K_{m_3} \exp\left[-\frac{(v+2)^2}{2\sigma_k^2}\right]$$
$$+ K_{m_4} \exp\left[-\frac{(v-4)^2}{2\sigma_k^2}\right] + K_{m_5} \exp\left[-\frac{(v+5)^2}{2\sigma_k^2}\right]$$
$$+ K_{m_6} \exp\left[-\frac{(v-6)^2}{2\sigma_k^2}\right] \tag{9}$$

with parameter values

$$K_m = [94 \;\; 92 \;\; 91 \;\; 85 \;\; 105 \;\; 80]$$
$$\beta = 0.1, \alpha_{\text{comp}} = 0.5, \sigma_k = \sqrt{0.5}, r = 0.25, \sigma_\alpha = 2.$$

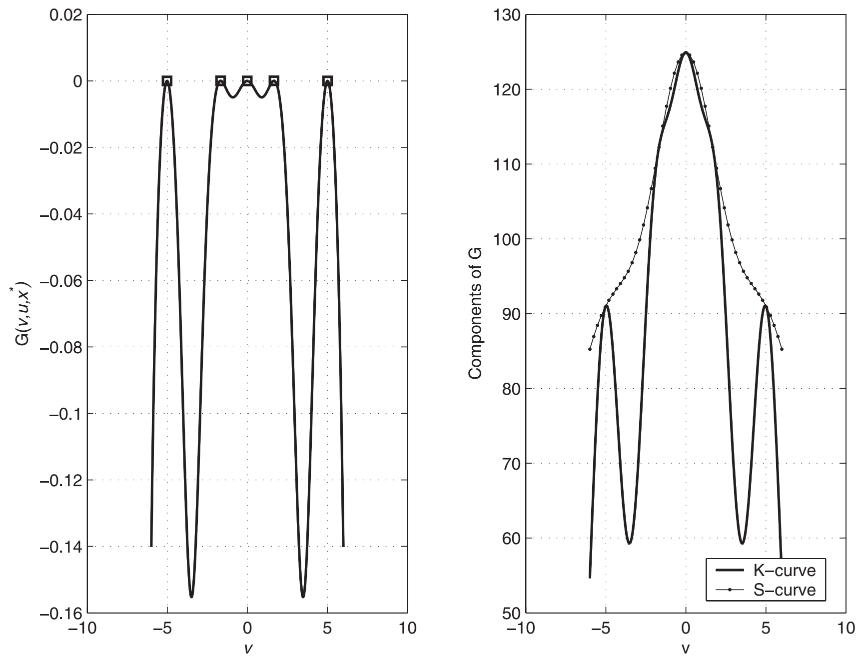


Figure 14: An ESS coalition of five using $\alpha_{\text{comp}} = 0$, symmetric α, and symmetric K.

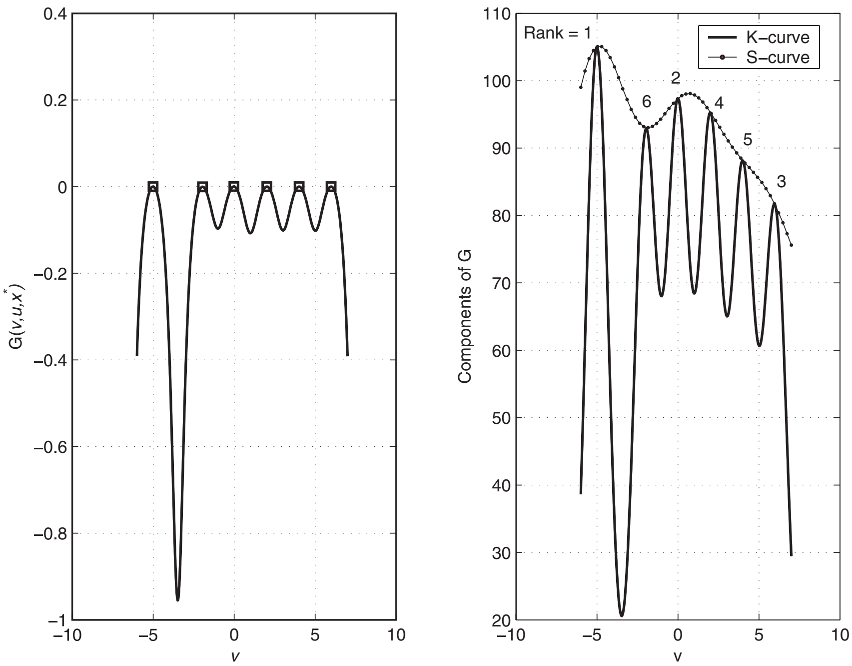


Figure 15: Nonsymmetric niches.

The ESS solution for **u** in this case is an ESS coalition of six located near each of the strategy peaks with equilibrium population values $\mathbf{x}^*$ given by

$$\mathbf{u} = [-5.006 \quad -1.954 \quad -0.0155 \quad 2.0224 \quad 4.015 \quad 5.991]$$
$$\mathbf{x}^* = [45.14 \quad 3.886 \quad 24.41 \quad 16.52 \quad 6.642 \quad 18.09].$$

The K-curve is a sum of several distribution functions such that the ordering of the highest peaks corresponds to the largest values of K_{m_i}. For example, the highest peak corresponds to $K_{m_5} = 105$ (at $v = -5$), the second highest corresponds to $K_{m_1} = 94$ (at $v = 0$), etc. However, we see that the ordering of the relative abundance need not follow this same pattern. For example, the strategy $v = 5.9907$ has the third highest relative abundance yet it is located close to the lowest peak. The reason relative abundance does not necessarily follow the peak heights is the frequency dependence introduced with the nonsymmetric competition term. The size of the relative maxima of the K function alone does not determine the size of the equilibrium populations at the ESS. Figure 16 is the relative abundance plot of $\mathbf{x}^*$ resulting in a curve that is very similar to the field data plotted in Figure 1.

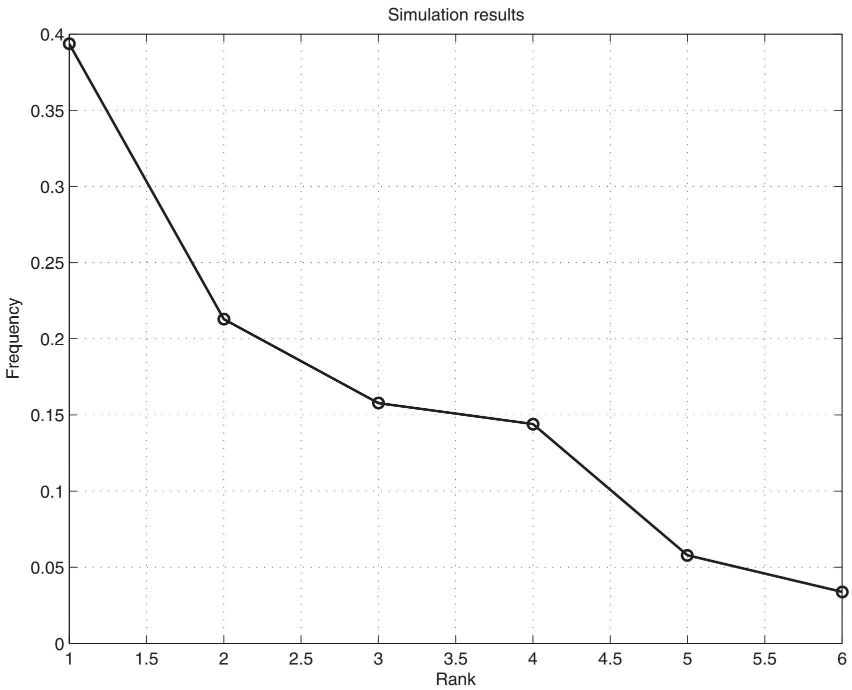


Figure 16: Simulation of relative abundance for six species.

7 Conclusions

The basic question examined here is whether a modified Lotka–Volterra competition model can reproduce relative abundance curves similar to those of natural biological systems. This appears to be the case, as witnessed by the similarity of Figure 16, obtained using this model, to Figures 1 and 2 obtained from the North American Breeding Bird Survey. Under the assumption that a biological system is operating at or near an ESS, the modified Lotka–Volterra competition model can reproduce lognormal-like curves provided that well-defined resource niches, asymmetric competition, and a fixed competitive cost are included in the model. Generally under n niches, there will be n species; however, this need not be the case. Moreover, the ordering of relative abundance of the species does not necessarily follow the ordering of the "richness" of the niches.

A complement to this result is the fact that for an ESS coalition of species existing on a single resource, the relative abundance curve becomes bowed outward, exhibiting a more even distribution of species abundance. Coexistence is possible in this case due to the small differences between the amount of resource available for biomass at a given strategy and the utilization of the resource by the community as a whole. Thus, without distinct niche differentiation based on species traits, coexistence can occur, but community structure is more even than that usually occuring in natural communities.

We began with a form of the Lotka–Volterra competition model that assumes competition to be strongest when strategies are similar. This assumption is realistic, as demonstrated in Fargione et al. [4] for grassland species. However, coexistence between multiple species is not possible for this situation without some modification of the model such as adding a fixed competitive cost. The fixed competitive cost is a spatial constraint that allows multiple species to coexist whilst competing for a single limiting resource with the consequence that a population can never reach carrying capacity on that particular resource. This is not a trade-off in the ability of a species to utilize space versus its ability to utilize, say, nitrogen, but rather is a fixed cost that comes with increasing numbers of competing individuals. This fixed competitive cost also results in increases in community biomass as the number of species increases. This result mirrors the results of large-scale experiments (Tilman et al. [19]) where grassland plots of sixteen species had 2.7 times greater biomass than monocultures. However, the lognormal-like abundance curves do not result solely from the incorporation of a fixed competitive cost.

A major problem with the original Lotka–Volterra competition model under symmetric competition,[4] without a fixed competitive cost, is the fact that frequency-dependent selection[5] vanishes at the ESS (Brown and Vincent [2]). Without frequency-dependent selection, an ESS coalition greater than one would be rare. This is one of the reasons we incorporated asymmetric competition along with a fixed competitive cost into the model. Finally, by introducing multi-humped carrying capacities ("niches") into the model, a dramatic change occurs in the species-abundance distributions. They now bow inward, like that seen in natural communities (Figures 1 and 2). By using the rules of natural selection and incorporating the ecological traits of individuals (the G-function method), our modified Lotka–Volterra competition model is able to make several testable predictions about the abundance of species within a community. First, a fixed competitive cost allows for coexistence on one resource. Second, the productivity of a community should increase with species richness if there is a fixed competitive cost. And finally, although it is possible to obtain relative abundance curves that bow outward when species utilize a single niche, a lognormal-like relative abundance distribution curve is expected when there is nonsymmetric competition in an environment with obvious resource niches.

---

[4]The competition term is symmetric when, for a given player's strategy, the effect of competition is symmetric with respect to the virtual variable v.

[5]Selection will be frequency dependent if the species' payoffs are functions of all the strategies present within the community.

References

[1] Benkman, C. W.: 2003, Divergent selection drives the adaptive radiation of crossbills, *Evolution* **57**(5), 1176–1181.

[2] Brown, J. S. and Vincent, T. L.: 1987, Coevolution as an evolutionary game, *Evolution* **41**, 66–79.

[3] Case, T. J.: 1982, Coevolution in resource-limited competition communities, *Theoretical Population Biology* **21**, 69–91.

[4] Fargione, J., Brown, C. S. and Tilman, D.: 2003, Community assembly and invasion: An experimental test of neutral versus niche processes, *PNAS* **100**(15), 8916–8920.

[5] Fisher, R. A., Corbet, A. S. and Williams, C. B.: 1943, The relation between the number of species and the number of individuals in a random sample of an animal population, *Journal of Animal Ecology* **12**, 42–58.

[6] Hubbell, S. P.: 2001, *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press, Princeton, NJ.

[7] Jones, G. P., Munday, P. L. and Caley, M. J.: 2002, *Coral Reef Fishes, Dynamics and Diversity in a Complex Ecosystem*, Academic Press, New York, chapter Rarity in coral reef fish communities, pp. 81–101.

[8] Kelly, C. K. and Woodward, F. I.: 1996, Ecological correlates of plant range size: Taxonomies and phylogenies in the study of plant commonness and rarity in Great Britain, *Phil. Trans. R. Soc. Lond. B* **351**(1), 1261–1269.

[9] Link, W. A. and Sauer, J.: 1998, Estimating population change from count data: Application to the North American breeding bird survey, *Ecological Applications* **8**(2), 258–268.

[10] Magurran, A. E.: 1988, *Ecological Diversity and its Measurement*, Princeton University Press, Princeton, NJ.

[11] Magurran, A. E. and Henderson, P. A.: 2003, Explaining the excess of rare species in natural species abundance distributions, *Nature* **422**, 714–716.

[12] May, R. M.: 1975a, *General Introduction. In Ecological Stability*, Chapman & Hall, London, pp. 1–14.

[13] May, R. M.: 1975b, *Patterns of Species Abundance and Diversity*, Belknap Press of Harvard University Press, Cambridge, MA, pp. 81–120.

[14] McGill, B. J.: 2003, A test of the unified neutral theory, *Nature* **422**, 881–885.

[15] Murray, B. R., Thrall, P. H., Gill, A. M. and Nicotra, A. B.: 2002, How plant life-history and ecological traits relate to species rarity and commonness at varing spatial scales, *Austral Ecology* **27**, 291–310.

[16] Roughgarden, J.: 1983, *The Theory of Coevolution*, Sinauer, Sunderland, MA, pp. 383–403.

[17] Rummel, J. D. and Roughgarden, J.: 1983, Some differences between invasion-structured and coevolution-structured competitive communities: A preliminary theoretical analysis, *Oikos* **41**, 477–486.

[18] Sugihara, G.: 1980, Minimal community structure: An explanation of species abundance patterns, *Am. Nat.* **116**, 770–787.

[19] Tilman, D., Reich, P., Knops, J., Wedin, D., Mielke, T. and Lehman, C.: 2001, Diversity and productivity in a long term grassland experiment, *Science* **294**, 843–845.

[20] Vincent, T. L. and Brown, J. S.: 2005, *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics*, Cambridge University Press, Cambridge.

[21] Vincent, T. L., Cohen, Y. and Brown, J. S.: 1993, Evolution via strategy dynamics, *Theoretical Population Biology* **44**, 149–176.

[22] Whittaker, R. H.: 1970, *Communities and Ecosystems*, Macmillan, New York.

Foraging Under Competition: Evolutionarily Stable Patch-Leaving Strategies with Random Arrival Times. 1. Scramble Competition

Frédéric Hamelin
CNRS and University of Nice Sophia Antipolis–I3S
École Polytechnique de l'Université de Nice Sophia Antipolis
930 Route des Colles, B.P.145
06903 Sophia Antipolis, France
hamelin@polytech.unice.fr

Pierre Bernhard
CNRS and University of Nice Sophia Antipolis–I3S
École Polytechnique de l'Université de Nice Sophia Antipolis
930 Route des Colles, B.P.145
06903 Sophia Antipolis, France
bernhard@polytech.unice.fr

Philippe Nain
INRIA
2004 Route des Lucioles
06902 Sophia Antipolis, France
philippe.nain@sophia.inria.fr

Éric Wajnberg
INRA
400 Route des Chappes, B.P. 167
06903 Sophia Antipolis Cedex, France
wajnberg@antibes.inra.fr

Abstract

Our objective is to determine the evolutionarily stable strategy [14] that is supposed to drive the behavior of foragers competing for a common patchily distributed resource [16]. Compared to [18], the innovation lies in the fact that random arrival times are allowed.

In this first part, we investigate scramble competition: the game still yields simple Charnov-like strategies [4]. Thus we attempt to compute the optimal long-term mean rate γ^* [11] at which resources should be gathered to achieve the maximum expected fitness: the assumed symmetry among foragers allows us to express γ^* as a solution of an implicit equation, independent of the probability distribution of arrival times.

A digression on a simple model of group foraging shows that γ_N^* can be simply computed via the classical graph associated to the marginal value theorem—N is the size of the group. An analytical solution allows us to characterize the decline in efficiency due to group foraging, as opposed to foraging alone: this loss can be relatively low, even in a "bad world," provided that the handling time is relatively long.

Back to the original problem, we then assume that the arrivals on the patch follow a Poisson process. Thus we find an explicit expression of γ^* that makes it possible to perform a numerical computation: Charnov's predictions still hold under scramble competition.

Finally, we show that the distribution of foragers among patches is not homogeneous but biased in favor of bad patches. This result is in agreement with common observation and theoretical knowledge [1] about the concept of ideal free distribution [12,22].

1 Introduction

Behavioral ecology [13] attempts to assert to what extent the natural selection process could have carved animal behavior. This evolutionary approach focuses on optimal strategies in terms of capitalizing on genetic inheritance through generations; as a common currency between survival ability and reproductive success, we shall use the term—Darwinian—*fitness* [15], analogous to the concept of "utility" in economics.

In this respect, optimal foraging theory [20] seeks to investigate the behavior of an animal searching for a valuable resource such as food or a host to parasitize. In many cases, these resources are spread in the environment as distant *patches* of various *qualities*. Moreover, the resource *intake rate* suffers from patch depletion. As a consequence, it is likely advantageous to leave a patch not yet exhausted in order to find a richer one, in spite of an uncertain *travel time*. Hence the need to determine the optimal leaving rule.

In this context, Charnov's marginal value theorem [4] provides a way to gather resources at an optimal long-term mean rate γ^* that gives the best fitness a forager can expect in its environment.

Actually, this famous theoretical model is applied to a lone forager that has a monopoly on the resources it finds; it predicts that each patch should be left when the intake rate on that patch drops below γ^*, independently of either its quality or on the time invested to reach it.

Naturally, the question arises of whether this result holds for foragers competing for a common patchily distributed resource [16], i.e., whether this is an evolutionarily stable strategy [14]. The authors of [18] assume that somehow n foragers have reached a patch simultaneously, and they investigate their evolutionarily stable giving-up strategy. Our innovation lies in the fact that an a priori unlimited number of foragers reaching a patch at random arrival times is allowed.

In Section 2, we develop a mathematical model of the problem at hand and recall Charnov's classical marginal value theorem. In Section 3, we investigate the *scramble competition* case, where the only competition between foragers is in sharing a common resource.

In a companion paper [9], we extend the model to take into account actual *interference*; i.e., a decline of the intake rate due to competition.

2 Model

We consider a population of independent animals foraging freely in an environment containing a patchily distributed resource, assumed to be stationary; i.e., the spatial and qualitative statistical distributions of the patches remain constant over time. In other words, there is no environment-wide depletion but only local depletion; an ad hoc renewal process of the resource is then implicitly assumed, although it might not necessarily be an appropriate modeling shortcut [2,3]. We then focus on a single forager evolving in this environment, among its conspecifics.

2.1 Local Fitness Accumulation

2.1.1 A Lone Forager on an Initially Unexploited Patch

We consider the case of a single forager acquiring some fitness from a patch of resource. We let

- $q \in \mathbb{R}^+$ be the quality of the patch, i.e., the potential fitness it initially offers,
- $p \in \mathbb{R}^+$ be the current state of the patch, i.e., the amount of fitness remaining,
- $\rho = p/q \in \Sigma_1 = [0, 1]$ be the fitness remaining on the patch relative to its quality.

Let $f(q, \tau)$ be the fitness gathered in a time τ on a patch of quality q. Our basic assumption is that the intake rate $\dot{f} = \partial f(q, \tau)/\partial\tau$ is a known function $r(\rho)$ continuous, strictly increasing and concave. In Appendix A.3 we derive such a law from an assumption of random probing on a patch. It yields

$$\dot{f} = r(\rho), \quad f(0) = 0,$$

resulting in

$$q\dot{\rho} = -r(\rho), \quad \rho(0) = 1. \tag{1}$$

We find it convenient to introduce the solution $\phi(t)$ of the differential equation

$$\dot{\phi} = -r(\phi), \quad \phi(0) = 1.$$

Theorem 2.1. *Our model is given by*

$$f(q, \tau) = q\left[1 - \phi\left(\frac{\tau}{q}\right)\right]. \tag{2}$$

It yields: $\forall q$,

- $f(q, 0) = 0$,
- $\tau \mapsto f(q, \tau)$ *is strictly increasing and concave,*
- $\lim_{\tau \to \infty} f(q, \tau) = q$.

2.1.2 A Lone Forager on a Previously Exploited Patch

Assume that the forager reaches a patch that has already been exploited to some extent by a conspecific. The patch is characterized by its initial quality q and its ratio of available resource ρ_0 at arrival time. The dynamics are still (1) initialized at $\rho(0) = \rho_0$, and the fitness gathered is

$$f(q, \rho_0, \tau) = p_0 - p(\tau) = q[\rho_0 - \rho(\tau)].$$

This is depicted on the reduced graph, Figure 1.

2.1.3 Several Foragers on a Patch

Assume that $n \in \mathbb{N}$ identical foragers are on the same patch. Let the sequence of forager arrivals times be $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ and $i \in \{1, 2, \dots, n\}$. By definition, scramble competition lets the intake rate be independent of n, thus

$$\forall i,\ \dot{f}_i = \dot{f} = r(\rho), \quad f_i(\sigma_i) = 0.$$

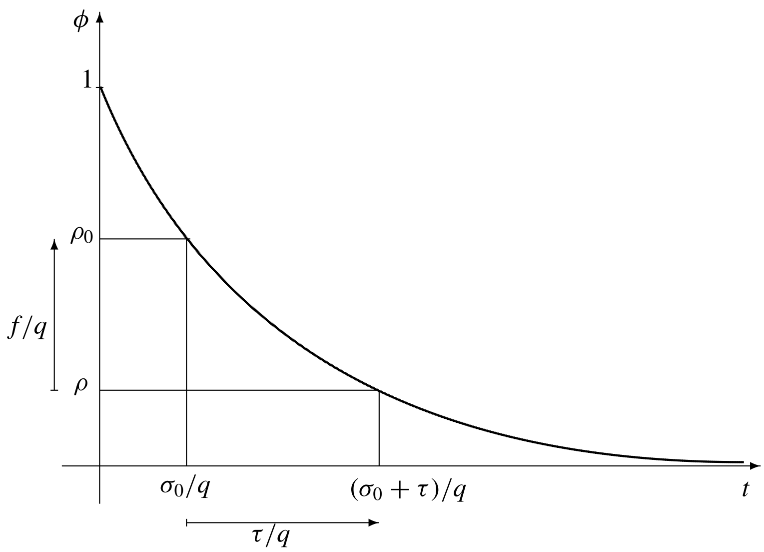


Figure 1: The reduced graph.

Nevertheless, the speed of depletion is multiplied by n:

$$\dot{p} = q\dot{\rho} = -n\dot{f}, \quad \rho(0) = \rho_0.$$

2.2 Global Fitness Accumulation

2.2.1 The Marginal Value Theorem

In order to optimally balance the residence times on the differing patches, a relevant criterion is the average fitness acquired relative to the time invested: assume the quality q of the patch visited is a random variable with cumulative distribution function $\mathcal{Q}(q)$. We allow the residence time to be a random variable, measurable on the sigma algebra generated by q. We also assume that the travel time θ is a random variable of known distribution and let $\bar{\theta} = \mathbb{E}\theta$. It yields

$$\gamma = \frac{\mathbb{E}f(q,\tau)}{\bar{\theta} + \mathbb{E}\tau}. \tag{3}$$

Theorem 2.2. *Charnov's marginal value theorem: the maximizing admissible τ is given as a function of q by the rule*

- *either $\frac{\partial f}{\partial \tau}(q, 0) \leq \gamma^*$ and $\tau^* = 0$,*
- *or τ is such $\frac{\partial f}{\partial \tau}(q, \tau^*)$ that γ^*,*

where γ^* is obtained by placing τ^* in (3).

Proof. Call $\mathrm{D}\gamma$ the Gâteaux derivative of γ in (3). Euler's inequality reads, for any $\delta\tau$ such that $\tau^* + \delta\tau$ is admissible,

$$\mathrm{D}\gamma.\delta\tau = \frac{1}{\bar{\theta} + \mathbb{E}\tau^*} \int_{\mathbb{R}^+} \left[\frac{\partial f}{\partial \tau}(q, \tau^*) - \gamma^*\right] \delta\tau(q)\ \mathrm{d}\mathcal{Q}(q) \leq 0.$$

The increment $\delta\tau$ may have any sign if τ^* is strictly positive, but it must be positive if τ^* is zero. Hence the result. This is a marginal improvement over Charnov's marginal value theorem. □

2.2.2 A Lone Forager Evolving in Our Model

As in the classical model, we consider in this subsection a lone forager which has a monopoly on resource it finds. Under the main modeling assumption of Section 2.1.1, the criterion becomes

$$\gamma = \mathbb{E}\left\{q\left[1 - \phi\left(\frac{\tau(q)}{q}\right)\right]\right\} \Big/ \left[\bar{\theta} + \mathbb{E}q\frac{\tau(q)}{q}\right].$$

Charnov's optimal patch-leaving strategy is to leave when $\dot{f} = \gamma^*$. In our model, the intake rate of a lone forager only depends on ρ, hence an equivalent

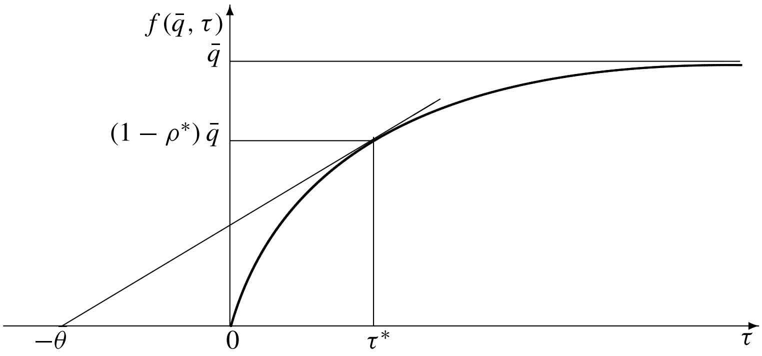


Figure 2: The marginal value theorem.

threshold is $\rho^* = r^{-1}(\gamma^*)$. Note that any unexploited patch should be attacked independently of its quality since, for every q, $(\partial f/\partial \tau)(q, 0) = r(1)$ and $r(1) > \gamma^*$ by construction.

A simple property of our model—see equation (2)—is that $\tau^*(q)/q$ is a constant, say z, $z = \phi^{-1}(\rho^*)$, for any q. Hence the following expression of γ^*, if we let $\bar{q} = \mathbb{E}q$:

$$\gamma^* = \frac{1-\phi(z)}{\bar{\theta}/\bar{q} + z}.$$

Therefore, one can compute the optimal value of ρ^*, or equivalently γ^*, via the well-known graph in Figure 2. One can notice the duality between $\bar{q}$ and $\bar{\theta}$: multiplying the average level of resource by n is equivalent to dividing the average travel time by n.

As a consequence, the patches should be relatively less depleted in a *good world* [6]—with rich and easy to find patches—than in a bad one—with scarce patches offering few resources.

Thus, in our particular case, only $\bar{q}$ is relevant: "it suffices to know $\bar{q}$—rather than $\mathcal{Q}(q)$—to be able to behave optimally." Hence this model stands if the resource is "only" stationary in a weak sense; i.e., if the **means** of the qualitative and spatial[1] statistical distributions of the patches remain constant over time.

2.2.3 An Explicit Formula for ρ^*

We now make use of the particular form of the function r of Appendix A.3: it allows the function $\phi(t)$ to be inverted into

$$\phi^{-1}(\rho) = h(1-\rho) - \alpha \ln(\rho). \tag{4}$$

[1]More precisely, this is the mean travel time which has to remain stationary.

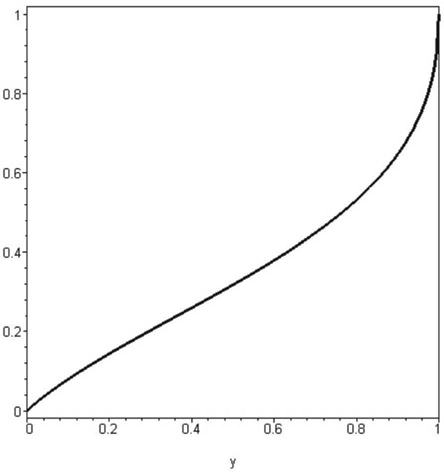


Figure 3: The function $\rho^*(y)$.

It yields an analytical solution, simply by performing an optimization in ρ as $\rho^* = \arg\max_\rho \gamma(\rho)$ with

$$\gamma(\rho) = \frac{1-\rho}{\bar{\theta}/\bar{q} + \phi^{-1}(\rho)}, \quad \rho \in \Sigma_1.$$

Hence

$$\rho^* = -1/W_{-1}(-\mathrm{e}^{-(1+x)}), \tag{5}$$

where $x = \bar{\theta}/(\alpha\bar{q})$ and W_{-1} is the Lambert W function as defined in [7]—this is indeed the "*nonprincipal*" branch of this multi-valued function that contains the solution as $\rho^* \in \Sigma_1 \Rightarrow W \leq -1$.

Thus ρ^* depends on $1 + x$, a sort of inverse duty cycle as $\alpha\bar{q}$ is the time needed to cover an average patch in a systematic way; note that ρ^* does not depend on the handling time h although γ^* does. Let $y = 1/(1 + x) = \alpha\bar{q}/(\alpha\bar{q} + \bar{\theta})$; the function $\rho^*(y)$ is plotted in Figure 3.

As expected, in a bad world the patches should be relatively more depleted than in a good one—high "duty cycle" y—where the forager would be harder to please.

3 Scramble Competition

Scramble competition only takes into account the fact that the resource depletes faster due to simultaneous foraging activities on the patch. As a consequence, the departure of a forager only slows down the depletion. Hence there is no hope to see ρ, or equivalently the intake rate, increase. Moreover, as foragers are assumed to be identical, they surely share the same optimal long-term mean rate γ^* and thus must leave at the same time, independently of their arrival times. Hence adopting commonly the Charnov's patch-leaving strategy given by Theorem 2.2 provides a Nash equilibrium in nonanticipative strategies among the population. As this latter is both strict and symmetric, this is indeed an evolutionarily stable strategy—this is detailed in Appendix B of the second part [9].

3.1 An Attempt to Get an Analytical Expression of γ^*

Let us assume that all foragers apply Charnov's patch-leaving strategy, i.e., they leave when $\dot{f} = \gamma^*$ or equivalently when $\rho = \rho^*$. As a consequence, when a patch is left, it is at a density ρ^* which makes it unusable for any forager. Hence all admissible patches encountered are still unexploited, with $\rho_0 = 1$.

Let t be the time elapsed since the patch was discovered. For a fixed ordered sequence of σ_j's, $j \in \{1, 2, \dots, n\}$, let us introduce a "forager second"—as one speaks of "man month"—$s = S(t, \sigma)$ defined by

$$\dot{s} = j \quad \text{if } \sigma_j \le t < \sigma_{j+1}, \quad s(0) = 0.$$

Equivalently

$$\text{for } t \in (\sigma_j, \sigma_{j+1}), \quad S(t, \sigma) = j(t - \sigma_j) + \sum_{k=1}^{j-1} k(\sigma_{k+1} - \sigma_k). \tag{6}$$

The function $t \mapsto S(t, \sigma)$ is strictly increasing. It therefore has an inverse function denoted $t = S_\sigma^{-1}(s)$, easy to write explicitly in terms of the $s_j = S(\sigma_j, \sigma)$:

$$\text{for } s \in (s_j, s_{j+1}), \quad S_\sigma^{-1}(s) = \frac{1}{j}(s - s_j) + \sum_{k=1}^{j-1} \frac{1}{k}(s_{k+1} - s_k).$$

According to Section 2.1.3, the dynamics of the patch are now

$$\dot{p} = q\dot{\rho} = -jr(\rho), \quad \text{for } t \in (\sigma_j, \sigma_{j+1}).$$

As a consequence, the patch trajectory satisfies

$$\rho(t) = \phi\left(\frac{1}{q} S(t, \sigma)\right).$$

We shall also let t^* be such that $\rho(t^*) = \rho^*$, i.e., to be explicit, if not clearer, $t^* = S_\sigma^{-1} \circ (q\phi^{-1}) \circ r^{-1}(\gamma^*)$.

Let us regroup possible combinations of σ's by the maximum number of foragers reached before they all leave the patch, say $\hat{n}$. When they leave, they have retrieved an amount $\sum_i f_i = q(1 - \rho^*)$ of the resource. By symmetry, the expectation of fitness acquired is, for each of them,

$$\mathbb{E}_\sigma f = \frac{q}{\hat{n}}(1 - \rho^*).$$

Moreover, this is exactly the amount of resource each would have acquired if they all had arrived simultaneously, since in that case they all acquire the same amount of resource.

Let us call a *central trajectory* of order $\hat{n}$ that particular trajectory where all $\hat{n}$ foragers arrived at time zero. We denote with an index $\odot$ the corresponding quantities. Hence, for all $\hat{n}$, $\mathbb{E}_\sigma(f) = f_\odot$.

Now, for a given ordered sequence σ of length $\hat{n}$, the reference forager may have occupied any rank, from 1 to $\hat{n}$. Let ξ be this rank. Call τ_ξ^* its residence time depending on ξ. Note that since they all leave simultaneously,

$$\forall \hat{n},\ \forall \xi \in \{1, \ldots, \hat{n}\}, \quad \tau_\xi^* = \sigma_{\hat{n}} - \sigma_\xi + \tau_{\hat{n}}^*.$$

Again, for reasons of symmetry,

$$\mathbb{E}_\xi \tau_\xi^* = \sigma_{\hat{n}} - \frac{1}{\hat{n}} \sum_{j=1}^{\hat{n}} \sigma_j + \tau_{\hat{n}}^*. \tag{7}$$

Now, τ_n^* is defined by $\phi(S(\sigma_{\hat{n}} + \tau_{\hat{n}}^*, \sigma)/q) = \rho^*$, i.e., according to equation (6):

$$\hat{n}[(\tau_{\hat{n}}^* + \sigma_{\hat{n}}) - \sigma_{\hat{n}}] + \sum_{j=1}^{\hat{n}-1} j(\sigma_{j+1} - \sigma_j) q\phi^{-1}(\rho^*).$$

Note that

$$\sum_{j=1}^{\hat{n}-1} j(\sigma_{j+1} - \sigma_j) = \hat{n}\sigma_{\hat{n}} - \sum_{j=1}^{\hat{n}} \sigma_j.$$

Hence we get

$$\tau_{\hat{n}}^* = \frac{q}{\hat{n}}\phi^{-1}(\rho^*) - \sigma_{\hat{n}} + \frac{1}{\hat{n}} \sum_{j=1}^{\hat{n}} \sigma_j.$$

On the central trajectory of order $\hat{n}$, it holds that $s = \hat{n}t = \hat{n}\tau$, so that

$$\tau_\odot^* = \frac{q}{\hat{n}}\phi^{-1}(\rho^*),$$

so that finally

$$\tau_{\hat{n}}^* = \tau_\odot^* - \sigma_{\hat{n}} + \sum_{j=1}^{\hat{n}} \sigma_j.$$

Place this in (7), and we obtain $\mathbb{E}_\xi \tau_\xi^* \tau_\odot^*$. But this last quantity is independent of σ, so that, for any fixed q and $\hat{n}$,

$$\mathbb{E}_\sigma \tau^* = \tau_\odot^* = \frac{q}{\hat{n}}\phi^{-1}(\rho^*).$$

The random variables q and $\hat{n}$ are surely correlated, as the foragers stay a longer time on better patches, and are thus likely to end up more numerous. Similarly, $\hat{n}$ surely depends on ρ^*; hence we use $\mathbb{E}^*$ to mean that we take the expected value over all patch qualities and sequences of arrival under the optimal scenario. Let then $q^* = \mathbb{E}^*(q/\hat{n})$. We obtain the fixed point equation:

$$r(\rho^*) = \gamma^* = \frac{1-\rho^*}{\bar{\theta}/q^* + \phi^{-1}(\rho^*)}. \tag{8}$$

Yet, it remains a partial result as long as we do not know how to express q^* as a function of ρ^*.

3.1.1 A Digression on a Simple Model of Group Foraging

In this subsection we relax the assumption of independent foragers provided that they are identical—all we need up to now is symmetry among foragers.

Thus let us consider a group of N identical individuals foraging "patch-by-patch"; i.e., the travel times are assumed to be too long to allow the group to cover two patches simultaneously. In this "*information-sharing*" model [8], once a patch is discovered by any member of the group, the others are assumed to join it sequentially. That is, we assume that the group spread itself in a radius [17] that allows every member to benefit from the poorest patch, as a function of the optimal profitability threshold ρ^* computed below. This assumption results in $\hat{n}$ equal to N independently of q and ρ^*. Therefore, the formula (8) is exactly the same as applying Charnov's marginal value theorem for both deterministic patch quality $\bar{q}/N$ and travel time $\bar{\theta}$. One can compute γ^* graphically, as done in Figure 4.

Obviously, foraging in a group is less[2] efficient in term of fitness gathering than foraging alone, if no advantage [5] is taken into account. However, it does not imply that the individual efficiency $\gamma^*(N) =: \gamma^*_N$ is an homogeneous function of

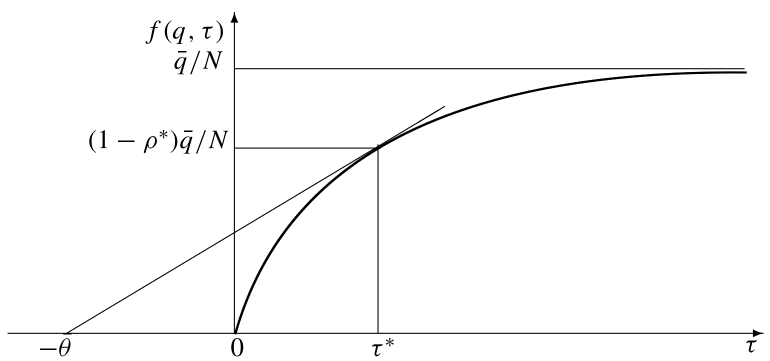


Figure 4: The marginal value theorem.

[2]At best equal, if ever the mean travel time was divided by N while foraging in a group [5].

degree -1; indeed, the relation $\gamma_N^* = \gamma_1^*/N$ would be true if the individuals were acting as if they were alone.

If we make use of the particular form of the function r of Appendix A.3, $N \mapsto \rho^*(N)$ is given by equation (5) with $x = N\bar{\theta}/(\alpha\bar{q})$; as $\gamma^* = r(\rho^*)$, the function $N \mapsto \gamma^*(\rho^*(N))$ is easily obtained. Let $\beta = \alpha/h$, $\mu = \bar{\theta}/(\alpha\bar{q})$ and

$$\Gamma(N) := \gamma_1^*/\gamma_N^* = \left[1 - \beta W_{-1}\left(-\mathrm{e}^{-(1+N\mu)}\right)\right] / \left[1 - \beta W_{-1}\left(-\mathrm{e}^{-(1+\mu)}\right)\right].$$

Let $\kappa = \beta/\left[(1 - \beta W_{-1}\left(-\mathrm{e}^{-(1+\mu)}\right)\right]$ and $\Gamma'(N) := \mathrm{d}\Gamma(N)/\mathrm{d}N$; we obtain

$$\Gamma'(N) = \kappa\mu W_{-1}\left(-\mathrm{e}^{-(1+N\mu)}\right) / \left[1 + W_{-1}\left(-\mathrm{e}^{-(1+N\mu)}\right)\right] > 0.$$

Let $\Gamma''(N) := \mathrm{d}\Gamma(N)/\mathrm{d}N^2$. We have

$$\Gamma''(N) = -\mu^2\kappa W_{-1}\left(-\mathrm{e}^{-(1+N\mu)}\right) \Big/ \left\{\left[1 + W_{-1}\left(-\mathrm{e}^{-(1+N\mu)}\right)\right]^3\right\} < 0.$$

Thus $\Gamma(N)$ is strictly increasing but concave. Therefore, foraging in a group should yield more than only an Nth of what a lone forager would get, provided that the strategy is adapted to the size of the group.

Moreover, it is easy to see that $\lim_{N\to\infty} \Gamma(N) = \infty$, that $\lim_{N\to\infty} \Gamma'(N) = \kappa\mu$ and that $\Gamma''(N)$ increases abruptly in the vicinity of zero. Hence $\Gamma(N)$ can be approximated by an affine function of slope $\kappa\mu$: let $\tilde{\Gamma}(N) := (1 - \kappa\mu) + N\kappa\mu \sim \Gamma(N)$. The "duty cycle" is now $y = 1/(1 + \mu)$. Figure 5 approximately characterizes the decline in individual efficiency resulting from foraging in a group, as opposed to foraging alone. We see that even in a bad world, the loss can be relatively small if the handling time is relatively long.

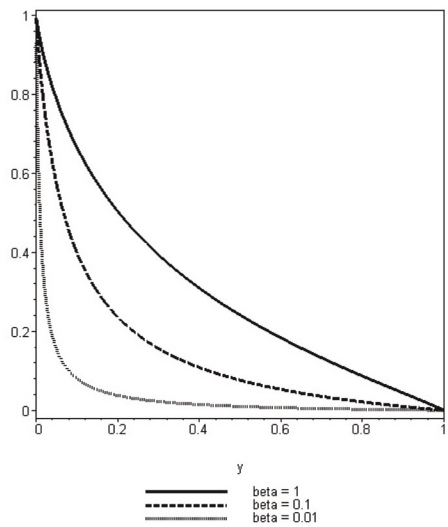


Figure 5: The function $y \mapsto \kappa\mu$.

3.1.2 Back to the Original Problem

As $q^* = \mathbb{E}_q\mathbb{E}^*(q/\hat{n}|q) = \mathbb{E}_q q\mathbb{E}^*(1/\hat{n}|q)$, we shall first consider that q is fixed.

Let ζ_1 be the time a lone forager would stay on a patch of quality q if not disturbed by an intruder: $\zeta_1 := \tau_\odot^*\hat{n} = q\phi^{-1}(\rho^*)$. In order to perform an optimization in ρ as in Section 2.2.3, our purpose is now to compute the function $\zeta_1 \mapsto \mathbb{E}^*(1/\hat{n})$. Let the successive arrival times on a patch be a Poisson process with intensity $\lambda > 0$. This means that the successive interarrival times form a sequence of mutually independent random variables $\{w_n\}$, exponentially distributed with mean $1/\lambda$.

Once a first intruder has arrived, the maximum—in the absence of further intruders—remaining time to deplete the patch up to ρ^* is divided by two as the depletion speed doubles; more generally, after the nth arrival, the maximum remaining residence time is reduced by a factor $(n-1)/n$. Our aim is now to express the cumulative distribution function of $\hat{n}$ in closed form as a function of ζ_1, from which we will deduce $\mathbb{E}^*(1/\hat{n})$.

A way to formulate the problem is the following one: let ζ_n be the remaining effort in "forager seconds" when the nth forager arrives. Clearly

$$\zeta_{n+1} = \zeta_n - nw_n, \quad n \geq 1.$$

Note that the mapping $n \to \zeta_n$ is nonincreasing. Therefore, the random variable $\hat{n}$ is characterized by $\zeta_{\hat{n}+1} \leq 0 < \zeta_{\hat{n}}$.

We have

$$\begin{aligned} P(\hat{n} > M) &= P(\zeta_2 > 0, \ldots, \zeta_{M+1} > 0), \\ &= P(\zeta_1 > w_1 + 2w_2 + \cdots + Mw_M). \end{aligned}$$

This is equivalent to finding the probability distribution of $\sum_{n=1}^{M} nw_n$. As the probability density function of the sum of independent random variables is given by the convolution product of their density functions, one can obtain it by inverting the product of the Laplace–Stieltjes transforms of their probability distributions. This is done in Appendix B and it yields

$$\mathbb{E}^*(1/\hat{n}|q) = 1 - \sum_{l=1}^{\infty}(1 - e^{-\lambda\zeta_1/l})e^{-l}\frac{l^{l-1}}{l!}.$$

Hence

$$q^* = \mathbb{E}^*(q/\hat{n}) = \bar{q} - \sum_{l=1}^{\infty}\left[\left(\bar{q} - \int_0^\infty \mathrm{e}^{-\lambda\zeta_1/l} q\,\mathrm{d}\mathcal{Q}(q)\right)e^{-l}\frac{l^{l-1}}{l!}\right].$$

We now make use of the particular form of $\phi^{-1}(\rho)$ given by equation (4); it yields $\zeta_1 = q[h(1-\rho) - \alpha\ln(\rho)]$.

As the Laplace transform of $q(\mathrm{d}\mathcal{Q}(q)/\mathrm{d}q)$ is the derivative of the Laplace transform of $-\mathrm{d}\mathcal{Q}(q)/\mathrm{d}q$, it yields

$$\int_0^\infty q\mathrm{d}\mathcal{Q}(q)\mathrm{e}^{-\lambda\zeta_1/l} - \mathcal{L}'(\hat{\nu}),$$

with $\hat{\nu} = \lambda[h(1-\rho) - \alpha\ln(\rho)]/l$, where $\mathcal{L}(\nu)$ is the Laplace–Stieltjes transform of q and $\mathcal{L}'(\nu) = \mathrm{d}\mathcal{L}(\nu)/\mathrm{d}\nu$.

Hence

$$q^* = \bar{q} - \sum_{l=1}^{\infty}\left\{[\bar{q} + \mathcal{L}'(\hat{\nu})]e^{-l}\frac{l^{l-1}}{l!}\right\}.$$

Although we now get an explicit expression of $\gamma(\rho)$ as, according to equation (8),

$$\gamma(\rho) = \frac{(1-\rho)}{\bar{\theta}/q^* + \phi^{-1}(\rho)},$$

this expression does not allow us to find an analytical expression for $\rho^* = \arg\max_\rho \gamma(\rho)$.

However, one can perform some numerical computations, as done in Figure 6. We took α as a time unit, $\beta = \alpha$, a unique $q = 200$ units of fitness, $\theta = 50\alpha$ and $L = 100$ for numerical computations, as suggested in Appendix B. At $\lambda \sim 0$, the mean interarrival time is infinite, thus we took $\forall\rho, \mathbb{E}(1/\hat{n}) = 1$. $\lambda = 0.05$ is

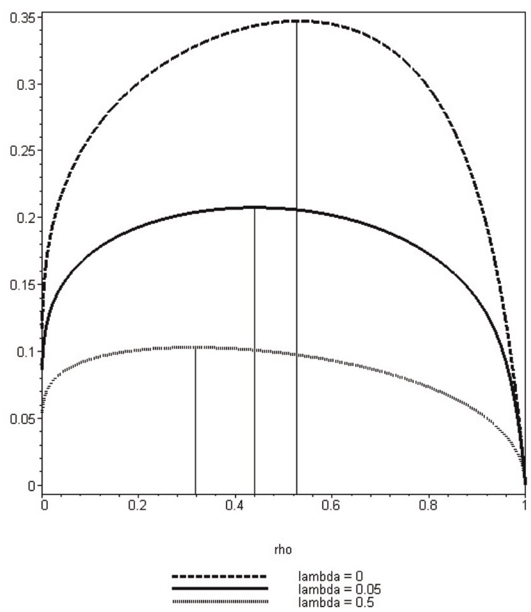


Figure 6: The function $\gamma(\rho)$.

a fair intensity as the mean interarrival time equal to 20α. $\lambda = 0.5$ is an extreme intensity as the mean interarrival time is equal to 2α.

In agreement with Charnov's model, the patches should be more depleted in a bad world—now in terms of the possible presence of competitors.

4 Concluding Remarks

Unavoidably, the consideration of the number of foragers reaching a patch as a function of its quality raises the issue of the relation with another central concept in foraging theory: the *ideal free distribution* [12,22]. It focuses on the distribution that corresponds to a Nash equilibrium among the foragers; i.e., into such a configuration, no one can individually improve its intake rate by moving instantaneously elsewhere. Hence the intake rates of identical foragers should be permanently equalized.

A simple property of our model—see equation (2)—is that a homogeneous and synchronous distribution of foragers yields a permanent equalization of their intake rates; i.e., if the number of foragers on any patch is proportional to patch quality and if they all reach their respective patch at the same time, their intake rates would remain equalized as all patch densities would decrease at the same speed.

Compared to that distribution, the calculations of Appendix B let one compute $\zeta_1 \mapsto \mathbb{E}^*\hat{n}$, the expected maximum number of foragers as a function of patch quality where now ρ^* is fixed and thus ζ_1 is proportional to q:

$$\mathbb{E}^*(\hat{n}) = 1 + \sum_{l=1}^{\infty}(1 - e^{-\lambda\zeta_1/l})e^{-l}\frac{l^{l-1}}{l-1!}.$$

It can be easily shown that the function $\zeta_1 \mapsto \mathbb{E}^*\hat{n}$ is increasing but concave, so good patches seem under matched relative to the "ideal free" distribution mentioned above. This deviation is in agreement with the common observation [21] and previous theoretical results [1] regarding the effect of perturbations such as nonzero travel time—or equivalently the foragers' asynchrony here.

Acknowledgments

This work is part of GDR 2155 "Écologie Comportementale" (CNRS commission 29), and the ESF/BEPAR scientific programme. We thank Minus van Baalen, Carlos Bernstein, Joel S. Brown, Michel de Lara, Jean-Sébastien Pierre, and an anonymous referee for their comments.

Appendix A Modeling Patch Depletion

A.1 Discrete Foraging

We consider in this subsection a situation where the resource comes as a finite number of tokens. We let $q \in \mathbb{N}$—for *quality*—be the initial number of tokens in the unvisited patch.

In our model, a token of resource remains on the patch once exploited, as an empty token. The forager is assumed to search for tokens at random—it is not supposed to search the patch in a systematic way—so that the distribution of depleted resource tokens among the patch will be assumed to be uniform at all times. Thus the forager finds itself more and more often probing a possible resource that turns out to be void. As a result, its efficiency decreases, prompting it to usually leave the patch before it is completely depleted. The decision parameter in the theory of patch use is the time τ that the forager spends on the patch before leaving it, or the *residence time*. We let α be the time it takes to move to a new token and probe it and h, the *handling time*, be the time it takes to actually exploit a token of resource. Let t_k be the time at which the kth valid resource token is found. It is actually exploited at time $t_k + h$. Let p_k be the amount of resource remaining on the patch *after* the kth unit is taken, i.e., $p_k = q - k$—and hence $p_0 = q$. Let also $\rho_k = p_k/q$ be the density of good resource tokens. We seek the law for t_{k+1}.

The forager finds a potential item of resource, possibly already exploited, every α units of time. For a given $t = t_k + h + \ell\alpha$, the event $t_{k+1} = t$ is equivalent to the fact that the items found at times $t_k + h + \alpha$, $t_k + h + 2\alpha, \dots, t_k + h + (\ell - 1)\alpha$ were already exploited, and the one found at time $t_k + h + \ell\alpha$ was not. During that time, ρ does not change, so that, assuming these events are independent—the patch is attacked in a homogeneous fashion—the probability of this event is

$$P_{k,\ell} = (1 - \rho_k)^{\ell-1}\rho_k.$$

Therefore, the expected time t_{k+1} is given by

$$\mathbb{E}(t_{k+1} - t_k - h) = \sum_{\ell=1}^{\infty}(1 - \rho_k)^{\ell-1}\rho_k\ell\alpha = \frac{\alpha}{\rho_k}.$$

Hence

$$\mathbb{E}(t_{k+1} - t_k) = \frac{\alpha + \rho_k h}{\rho_k}. \tag{A.1}$$

Deriving from there the law f, i.e., the expectation of the number of good resource tokens found in a given time τ, is done in Appendix A.2. One computes, for $n \le q$,

$$P_k^n := P\{t_n = k\alpha + (n - 1)h\},$$

and finds that it can be expressed in terms of products a_m^n of combinatorial coefficients

$$a_m^n = (-1)^{n-1}\binom{q-1}{n-1}(-1)^m\binom{n-1}{m},$$

as (equation(A.3))

$$P_k^n = \sum_{m=0}^{n-1} a_m^n \left(\frac{m}{q}\right)^{k-1}.$$

Then, let $k_n = \mathrm{Int}[(\tau - nh)/\alpha]$. The expected harvest is

$$f(q,\tau) = \sum_{n \le q} n P^n_{k_n}.$$

A.2 Combinatorics of Discrete Foraging

We have seen, from equation (A.1), that

$$P\{t_{k+1} - t_k - h\ell\alpha\} =: P_{k,\ell} = (1-\rho_k)^{\ell-1}\rho_k.$$

From there, we compute the full law for the residence time τ_n as follows. Let $P^n_k := P\{t_n = k\alpha + (n-1)h\}$. It is the probability that k attempts were necessary to find n items. It is the probability that $t_0 + (t_1 - t_0 - h) + \cdots + (t_n - t_{n-1} - h) = k\alpha$. The characteristic function of the sum of independent random variables is the product of their characteristic functions. Let therefore

$$\hat{P}_k(z) = \sum_{\ell=1}^{\infty} P_{k,\ell} z^{-\ell} \frac{\rho_k}{z - (1-\rho_k)}.$$

The characteristic function of t_n is therefore

$$\begin{aligned}\hat{P}^n(z) &= \hat{P}_0(z)\hat{P}_1(z)\cdots\hat{P}_{n-1}(z),\\ &= \frac{\rho_0\rho_1\ldots\rho_{n-1}}{[z-(1-\rho_0)][z-(1-\rho_1)]\ldots[z-(1-\rho_{n-1})]}.\end{aligned}$$

If, now, $\rho_0 = 1$ and $\rho_\ell = 1 - \ell/q$, we obtain

$$\hat{P}^n(z)\frac{\left(1-\frac{1}{q}\right)\left(1-\frac{2}{q}\right)\cdots\left(1-\frac{n-1}{q}\right)}{z\left(z-\frac{1}{q}\right)\cdots\left(z-\frac{n-1}{q}\right)}. \tag{A.2}$$

It remains to expand this rational fraction in powers of z^{-1} to compute the probability sought: $P^n_k = P\{t_n = k\alpha + (k-1)h)\}$. This is done through a decomposition in simple elements and expansion of each. If we let

$$\hat{P}^n(z) = \sum_{m=0}^{n-1} \frac{a^n_m}{z - \frac{m}{q}},$$

we obtain, for $n \le q$,

$$a^n_m = (-1)^{n-m-1}\frac{(q-1)!}{(q-n)!\,m!\,(n-m-1)!}(-1)^{n-m-1}\binom{q-1}{n-1}\binom{n-1}{m},$$

and the expansion yields, still for $n \le q$:

$$P^n_k = \sum_{m=0}^{n-1} a^n_m \left(\frac{m}{q}\right)^{k-1}, \tag{A.3}$$

with the convention that $0^0 = 1$, which is useless in practice, since for $k > 1$, the only interesting case, the term $m = 0$ can clearly be omitted.

It can be directly shown that the above formulas enjoy the desired properties that for any fixed $n \le q$, the P_k^n are null if $k < n$, and add up to one:

$$\forall k < n,\ P_k^n = 0, \quad \text{and} \quad \sum_{k=n}^{k=\infty} P_k^n = 1.$$

A.3 Continuous Foraging

Following most of the literature, we shall use a continuous approximation of the above theory, assuming that the resource is, somehow, a continuum: now, $q \in \mathbb{R}^+$. Let us introduce a surface—or volume—resource density D.[3] Two time constants enter into the model:

- α is the time it takes for the forager to explore a unit area that could contain a quantity D of resource, if it were not yet exploited.
- h is the extra time, or *handling time*, it takes to actually retrieve a unit of resource if necessary.

Our hypothesis is that a ratio ρ of the patch area is productive so that an area dæ produces a quantity

$$\mathrm{d}f = \rho D \mathrm{d}æ$$

of resource and the time necessary to gather it is

$$\mathrm{d}t = \alpha \mathrm{d}æ + \rho D h \mathrm{d}æ.$$

Hence we get

$$\dot{f} = \frac{\rho D}{\alpha + \rho D h} := r(\rho).$$

One can relate this equation to Holling's equation [10] by substituting α by the *attack rate*, a parameter giving the amount of resource attacked per unit time, $a = D/\alpha$.

Appendix B Evaluating a Probability Distribution

Let $w_1, \ldots, w_n$ be mutually independent random variables with common probability distribution $P(w_j < x) = 1 - \exp(-\lambda x)$. Define $Y_k = w_1 + 2w_2 + \cdots + kw_k$. The Laplace–Stieltjes transform (LST) of Y_k is given by

$$f_k(s) := \mathbb{E}(e^{-sY_k}) = \prod_{j=1}^{k} \frac{\lambda}{\lambda + js}.$$

[3] In the body of this chapter, we assume that the unit of area chosen is such that $D = 1$ or, equivalently, α is the time required to probe one unit of resource.

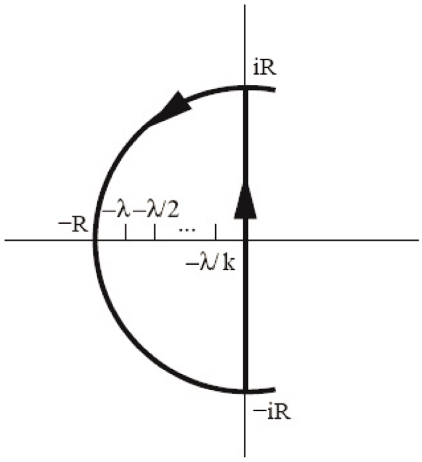


Figure B.1: The contour C_R.

Denote by $g_k(t)$ the density function of Y_k, namely, $g_k(t) = dP(Y_k < t)/dt$. The function $g_k(t)$ may be computed by inverting the LST $f_k(s)$. This gives

$$g_k(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{st} f_k(s) ds,$$

where γ is any real number chosen so that the line $s = \gamma$ lies to the right of all singularities of $f_k(s)$ [19]. The function $f_k(s)$ has only k simple poles, located at points $s = -\lambda/j$ for $j = 1, \ldots, k$. We may therefore take $\gamma = 0$.

The usual way of computing the complex integral $\int_{-i\infty}^{i\infty} e^{st} f_k(s) ds$ is first to consider the complex integral $I(R) := \int_{C_R} e^{st} f_k(s) ds$, where C_R is the contour defined by the half circle in the left complex plane centered at $s = 0$ with radius R, and the line $[-iR, iR]$ on the imaginary axis. R is any real number such that $R > 1/\lambda$ so that all poles of $f_k(s)$ are located inside the contour C_R; see Figure B.1. By applying the residue theorem we see that

$$I(R) = 2\pi i \sum_{l=1}^{k} \text{Residue } (e^{st} f_k(s); s = -\lambda/l).$$

Since the residue of the function $e^{st} f_k(s)$ at $s = -\lambda/l$ is equal to $e^{-\lambda t/l}(\lambda/l) \times \prod_{\substack{j=1 \\ j \neq l}}^{k} l/(l-j)$, we find that

$$I(R) = 2\pi i \sum_{l=1}^{k} e^{-\lambda t/l} \frac{\lambda}{l} \prod_{\substack{j=1 \\ j \neq l}}^{i} \frac{l}{l-j}. \tag{B.1}$$

At this point we have shown that

$$\begin{aligned}
g_k(t) &= \frac{1}{2\pi i} \lim_{R\to\infty} \int_{-iR}^{iR} e^{st} f_k(s) ds, \\
&= \frac{1}{2\pi i} \lim_{R\to\infty} I_R - \frac{1}{2\pi i} \lim_{R\to\infty} \int_{\Gamma_R} e^{st} f_k(s) ds, \\
&= \sum_{l=1}^{k} e^{-\lambda t/l} \frac{\lambda}{l} \prod_{\substack{j=1 \\ j\neq l}}^{k} \frac{l}{l-j} - \frac{1}{2\pi i} \lim_{R\to\infty} \int_{\Gamma_R} e^{st} f_k(s) ds,
\end{aligned}$$

by using (B.1), where $\Gamma_R = C_R - [-iR, iR]$.

One can find constants $K > 0$ and $a > 0$ such that $|f_k(s)| < K/R^a$ when $s = Re^{i\theta}$ for R large enough,[4] so that the integral in the latter equation vanishes as $R \to \infty$ [19, Theorem 7.4].

In summary, the density function $g_k(s)$ of the r.v. Y_k is given by

$$g_k(t) = \sum_{l=1}^{k} e^{-\lambda t/l} \frac{\lambda}{l} \prod_{\substack{j=1 \\ j\neq l}}^{k} \frac{l}{l-j}. \tag{B.2}$$

Let us now come back to the original problem. Define with $\zeta > 0$

$$n = \inf\{k \geq 1 : \zeta - (w_1 + 2w_2 + \cdots + kw_k) \leq 0\},$$

or equivalently

$$n = \inf\{k \geq 1 : \zeta - Y_k \leq 0\}.$$

We are interested in $\mathbb{E}(1/n)$. We have

$$\begin{aligned}
P(n > M) &= P(\zeta - Y_1 > 0, \ldots, \zeta - Y_M > 0), \\
&= P(Y_1 < \zeta, \ldots, Y_M < \zeta), \\
&= P(Y_M < \zeta). \qquad\qquad (B.3)
\end{aligned}$$

Since $P(n = M) = P(n > M - 1) - P(n > M)$ we see from (B.3) that for $M \geq 2$,

$$\begin{aligned}
P(n = M) &= P(Y_{M-1} < \zeta) - P(Y_M < \zeta) \\
&= \int_0^\zeta g_{M-1}(t) dt - \int_0^\zeta g_M(t) dt \qquad\qquad (B.4) \\
&= \sum_{l=1}^{M-1} (1 - e^{-\lambda\zeta/l}) \prod_{\substack{j=1 \\ j\neq l}}^{M-1} \frac{l}{l-j} - \sum_{l=1}^{M} (1 - e^{-\lambda\zeta/l}) \prod_{\substack{j=1 \\ j\neq l}}^{M} \frac{l}{l-j},
\end{aligned}$$

where the latter equality follows from (B.2).

[4]Hint: Always true if $f_k(s) = P(s)/Q(s)$, where P and Q are polynomials and the degree of P is strictly less than the degree of Q.

The right-hand side of (B.4) can be further simplified, to give

$$P(n = M) = \sum_{l=1}^{M} (1 - e^{-\lambda\zeta/l})(-1)^{M-1-l} \frac{M}{(M-l)!} \frac{l^{M-2}}{(l-1)!}, \qquad \text{(B.5)}$$

for $M \geq 2$. It remains to determine $P(n = 1)$. Clearly,

$$P(n = 1) = P(Y_1 > \zeta) = e^{-\lambda\zeta}. \qquad \text{(B.6)}$$

Therefore,

$$\begin{aligned}
\mathbb{E}(1/n) &= \sum_{M=1}^{\infty} \frac{1}{M} P(n = M) \\
&= 1 + \sum_{M=1}^{\infty} \sum_{l=1}^{M} (1 - e^{-\lambda\zeta/l})(-1)^{M-1-l} \frac{1}{(M-l)!} \frac{l^{M-2}}{(l-1)!} \\
&= 1 + \sum_{l=1}^{\infty} (1 - e^{-\lambda\zeta/l}) \frac{1}{(l-1)!} \sum_{M=l}^{\infty} (-1)^{M-1-l} \frac{l^{M-2}}{(M-l)!} \\
&= 1 - \sum_{l=1}^{\infty} (1 - e^{-\lambda\zeta/l}) e^{-l} \frac{l^{l-1}}{l!}. \qquad \text{(B.7)}
\end{aligned}$$

Similarly, we find

$$\mathbb{E}(n) = 1 + \sum_{l=1}^{\infty} (1 - e^{-\lambda\zeta/l}) e^{-l} \frac{l^{l-1}}{l-1!}. \qquad \text{(B.8)}$$

Remark

A way to avoid the calculation of the infinite series in the right-hand side of (B.7), or similarly that of (B.8), is to split the series in two parts: $\sum_{l=1}^{L} (1-e^{-\lambda\zeta/l})e^{-l}l^{l-1}/l!$ and $\sum_{l>L}(1-e^{-\lambda\zeta/l})e^{-l}l^{l-1}/l!$ for some arbitrary—but carefully chosen—integer $L > 1$. The first—finite— series can be evaluated without any problem for moderate values of L, and the second one can be approximated by using Stirling's formula as shown below. Indeed, if we use the standard approximation $l! \sim \sqrt{2\pi l}\, l^l\, e^{-l}$, then it follows that

$$\sum_{l>L} (1 - e^{-\lambda\zeta/l}) e^{-l} \frac{l^{l-1}}{l!} \sim 1 - \frac{1}{\sqrt{2\pi}} \sum_{l>L} (1 - e^{-\lambda\zeta/l}) l^{-3/2}.$$

We can further approximate the infinite series $\sum_{l>L}(1 - e^{-\lambda\zeta/l})l^{-3/2}$ by the integral $\int_L^{\infty} (1 - e^{-\lambda\zeta/x}) x^{-3/2}\, dx$, which gives

$$\sum_{l>L}^{\infty} (1 - e^{-\lambda\zeta/l}) l^{-3/2} \sim \frac{2}{\sqrt{L}} - \sqrt{\frac{\pi}{\lambda\zeta}}\, \text{erf}\left(\sqrt{\frac{\lambda\zeta}{L}}\right),$$

where the error function erf is defined by erf $:= 2/\sqrt{\pi} \int_0^x e^{-t^2} dt$.

References

[1] Bernstein C., Kacelnik A., Krebs J.R.: Individual decisions and the distribution of predators in a patchy environment II: the influence of travel costs and structure of the environment. *Journal of Animal Ecology*, **60**:205–225, 1991.

[2] Brown J.S.: Patch use as an indicator of habitat preference, predation risk and competition. *Behavioral Ecology and Sociobiology*, **22**:37–47, 1988.

[3] Brown J.S., Rosenzweig M.L.: Habitat selection in slowly regenerating environments. *Journal of Theoretical Biology*, **123**:151–171, 1986.

[4] Charnov E.L.: Optimal foraging: the marginal value theorem. *Theoretical Population Biology*, **9**:129–136, 1976.

[5] Clark C.W., Mangel M.: The evolutionary advantages of group foraging. *Theoretical Population Biology*. **30**:45–75, 1986.

[6] Clark C.W., Mangel M.: Dynamic state variable models in Ecology, methods and applications. Oxford Series in Ecology and Evolution. Oxford University Press, New York, USA, 2000.

[7] Corless R.M., Gonnet G.H., Hare D.E.G., Jeffrey D.J., Knuth D.E.: On the Lambert W function. *Advances in Computational Mathematics*, **5**:329–359, 1996.

[8] Giraldeau, L.-A., Beauchamp, G.: Food exploitation: searching for the optimal joining policy. *Trends in Ecology and Evolution*, **14**:102–106, 1999.

[9] Hamelin F., Bernhard P., Shaiju A.J., Wajnberg E.: Foraging under competition: evolutionarily stable patch-leaving strategies with random arrival times. 2. Interference competition. *Annals of Dynamic Games, this volume, Birkhauser*, pp. 349–366, 2007.

[10] Holling C.S.: Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist*, **91**:385–398, 1959.

[11] Houston A.I., McNamara J.M.: Models of adaptive behavior: an approach based on state. Cambridge University Press, Cambridge, UK, 1999.

[12] Kacelnik A., Krebs J.R., Bernstein C.: The ideal free distribution and predator-prey populations. *Trends in Ecology and Evolution*, **7**:50–55, 1992.

[13] Krebs J.R., Davies N.B., editors: Behavioural ecology: an evolutionary approach. Blackwell Science, Oxford, UK, 1997.

[14] Maynard Smith J.: Evolution and the theory of games. Cambridge University Press, Cambridge, UK, 1982.

[15] McNamara J.M., Houston A.I., Collins E.J.: Optimality models in Behavioral Biology. *SIAM Review*. **43**: 413–466, 2001.

[16] Parker G.A., Stuart R.A.: Animal behaviour as a strategy optimizer: evolution of resource assessment strategies and optimal emigration thresholds. *The American Naturalist*, **110**:1055–1076, 1976.

[17] Ruxton G.D., Fraser C., and Broom M.: An evolutionarily stable joining policy for group foragers. *Behavioral Ecology*, **16**:856–864, 2005.

[18] Sjerps M., Haccou P.: Effects of competition on optimal patch leaving: a war of attrition. *Theoretical Population Biology*, **3**:300–318, 1994.

[19] Spiegel M.R.: Shaum's outline of theory and problems of Laplace transforms, Shaum's Outline Series, McGraw-Hill Book Company, New York, USA, 1965.

[20] Stephens D.W., Krebs J.R.: Foraging theory. Monographs in Behavior and Ecology, Princeton University Press, Princeton, New Jersey, USA, 1986.

[21] Sutherland W.J.: From individual behavior to population ecology. Oxford Series in Ecology and Evolution. Oxford University Press, New York, USA, 1996.

[22] Trezenga T.: Building on the ideal free distribution. *Advances in Ecological Research*, **26**:253–302, 1995.

Foraging Under Competition: Evolutionarily Stable Patch-Leaving Strategies with Random Arrival Times 2. Interference Competition

Frédéric Hamelin
CNRS and University of Nice Sophia Antipolis–I3S
École Polytechnique de l'Université de Nice Sophia Antipolis
930 Route des Colles, B.P. 145
06903 Sophia Antipolis, France
hamelin@polytech.unice.fr

Pierre Bernhard
CNRS and University of Nice Sophia Antipolis–I3S
École Polytechnique de l'Université de Nice Sophia Antipolis
930 Route des Colles, B.P. 145
06903 Sophia Antipolis, France
bernhard@polytech.unice.fr

A. J. Shaiju
CNRS and University of Nice Sophia Antipolis–I3S
École Polytechnique de l'Université de Nice Sophia Antipolis
930 Route des Colles, B.P. 145
06903 Sophia Antipolis, France
ajshaiju@gmail.com

Éric Wajnberg
INRA
400 Route des Chappes, B.P. 167
06903 Sophia Antipolis Cedex, France
wajnberg@antibes.inra.fr

Abstract

Our objective is to determine the evolutionarily stable strategy [13] that is supposed to drive the behavior of foragers competing for a common patchily distributed resource [15]. Compared to [17], the innovation lies in the fact that random arrival times are allowed.

In this second part, we add interference to the model: it implies that a "passive" Charnov-like strategy can no longer be optimal. A dynamic programming approach leads to a sequence of wars of attrition [13] with random end times. This game is solved in Appendix A. Under some conditions that prevail in our model, the solution is independent of the probability law of the horizon. As a consequence, the solution of the asynchronous foraging problem investigated here, expressed as a closed loop strategy on the number of foragers, is identical to that of the synchronous problem [17].

Finally, we discuss the biological implications such as a possible connection with the genetic variability in the susceptibility to interference observed in [22].

1 Introduction

As the main concepts and notation are introduced in a companion paper [9], we just summarize them hereafter.

"*Nothing in biology makes sense except in the light of evolution*".[1] In this respect, behavioral ecology interprets animal behavior through an evolutionary approach, via estimating its capacity to get through the natural selection process, and to maximize Darwinian *fitness* [12]—a notion analogous to that of "utility" in economics. Typically, in foraging theory or the art of gathering resources in the environment, fitness is related to the quantity of resource gathered. In many cases, the resource is patchily distributed and the utility function on each patch is strictly increasing, concave and bounded with respect to time. As the *intake rate* decreases with the quantity of resource available on the patch, it is likely advantageous to leave a patch not yet exhausted in order to find a new one, in spite of an uncertain *travel time*. Charnov's marginal value theorem reveals that the optimal giving-up time is when the intake rate is equal to the optimal long-term mean rate γ^*—which, if achieved, gives the best fitness a forager can expect in its environment. This famous theoretical model is actually applied to a lone forager that has a monopoly on resources it finds.

Naturally, the question arises of whether this result holds for foragers competing for a common patchily distributed resource, i.e., whether this is an evolutionarily stable strategy [13]. The authors of [17] assume that somehow n foragers have reached a patch simultaneously, and they investigate the evolutionarily stable giving-up strategy. Our innovation lies in the fact that an a priori unlimited number of foragers reaching a patch at random arrival times is allowed. We shall refer to these situations as, respectively, synchronous and asynchronous foraging.

In the first part [9], we investigated the *scramble competition* case where the only competition between foragers is in sharing a common resource: Charnov's patch-leaving rule remains qualitatively unchanged. In this second part, we extend that model to take into account actual *interference* [18], i.e., the decline of the intake rate due to competition. The complete solution of the new game is obtained in Section 2, and makes use of the solution of a *war of attrition* [13] with random end time, solved in a more general setup in Appendix A.

We freely refer to the concepts of evolutionarily stable strategy (ESS) and replicator dynamics provided by evolutionary game theory. Appendix B gathers

[1]Theodosius Dobzhansky, geneticist, 1900–1975.

some basic facts concerning these topics and their relationship to classical game theory.

2 Interference Competition

In this second part, we assume that beyond sharing the same resource, competition on a patch yields a decline of the intake rate of all the contestants [18]. This effect might even increase with the scarcity of the resource. As a consequence, the departure of a forager surely causes an abrupt rise of the intake rate. It implies that the latter does not only depend on the ratio of available resource but also on the current number of foragers present on the patch. A passive Charnov-like strategy, where the foragers only monitor their own intake rate to decide whether to stay or leave, should no longer be optimal.

Indeed, previous papers [17] reveal that synchronous foragers should trigger a war of attrition, i.e., the foragers should leave at random—but optimally distributed—times, except the lucky one which remains alone on the patch, expected to stay to exhaust the patch up to its profitability threshold.

The question arises as to whether this result holds for asynchronous foragers or to what extent. The doubt mainly arises from the fact that unexpected newcomers can now enter the game.

2.1 Model

Assume that $n \in \mathbb{N}$ identical foragers are on the same patch. Let the sequence of forager arrival times be $\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$ and $i \in \{1, 2, \ldots, n\}$. We let

- $q \in \mathbb{R}^+$ be the quality of the patch, i.e., the potential fitness it initially offers,
- $p \in \mathbb{R}^+$ be the current state of the patch, i.e., the amount of fitness remaining,
- $\rho = p/q \in \Sigma_1 = [0, 1]$ be the fitness remaining on the patch relative to its quality.

Let $m \in \mathbb{R}^+$ be a parameter which quantifies interference intensity among foragers [18]; $m = 0$ corresponds to scramble competition. Let $r(\rho, n, m)$ be a known function such that

- $\forall\, n,\ m, \quad \rho \mapsto r(\rho, n, m)$ is continuous, strictly increasing and concave,
- $\forall\, \rho,\ m, \quad n \mapsto r(\rho, n, m)$ is strictly decreasing if $m > 0$ and invariant otherwise,
- $\forall\, \rho,\ n, \quad m \mapsto r(\rho, n, m)$ is strictly decreasing if $n > 1$ and invariant otherwise.

Our basic assumption is that the fitness gathered by forager i is given by the differential equation

$$\forall i,\ \dot{f_i} = \dot{f} = r(\rho, n, m), \quad f_i(\sigma_i) = 0,$$

and

$$\dot{p} = q\dot{\rho} = -n\dot{f}, \quad \rho(0) = \rho_0.$$

Let the fitness accumulated by forager i after a residence time τ_i be $f_i(\tau_i, \tau_{-i}, \sigma_i)$ where τ_{-i} stands for the set $\{\tau_j\}$, $j \neq i$, which surely impacts f_i.

Following [12, 17], we use an equivalent criterion to that of [9] which is the effective fitness compared to the optimal (Nash) average one for a given residence time:

$$J_i(\tau_i, \tau_{-i}, \sigma_i) = f_i(\tau_i, \tau_{-i}, \sigma_i) - (\bar{\theta} + \tau_i)\gamma^*,$$

where $\bar{\theta}$ is the mean travel time. Note that by definition γ^* is such that the maximum expected J is zero.

2.2 The Game

A priori, we cannot exhibit any Nash equilibrium in pure strategies; hence the need to deal with mixed strategies, say $P_i, i \in \{1, 2, \ldots, n\}$ for n foragers. We shall use the subscript $-i$ to mean all players except player i.

So our criterion becomes the following generating function:

$$\mathcal{G}_i(P_i, P_{-i}, \sigma_i) = \mathbb{E}^{P_i, P_{-i}}_{\tau_i, \tau_{-i}} J_i(\tau_i, \tau_{-i}, \sigma_i). \tag{1}$$

As a consequence of the above definition of γ^*,

$$\mathbb{E}\mathcal{G}_i(P_i^*, P_{-i}^*, \sigma_i) = 0.$$

Let us define a *stage* as a stochastic period during which the number of foragers n remains constant on the patch; note that in such a stage the intake rate is only affected by ρ. Let the superscript $k \in \mathbb{N}$ denote the number of the stage; $k = 0$ indicates the stage at which the reference forager started the game. As there exists a profitability threshold ρ^*, the patch can not be indefinitely exploited; the total number of stages $K \in \mathbb{N}$ and the total number of players $N \in \mathbb{N}$ are thus finite, but a priori unknown.

We define the state at the beginning of stage k as

$$\chi^k = \begin{pmatrix} \rho^k \\ n^k \end{pmatrix} \in \Sigma_1 \times \mathbb{N}.$$

For each stage, each player commits to a *persistence time* $x_i^k \in \mathbb{R}^+$; i.e., if the stage is not yet finished at that time it quits the game and so its own horizon is $K_i = k$. We find it convenient to let the exceptional—zero-measure—case, where all x_i are equal, end the current stage: it means that all players are invited to play again in order to make the patch surely exhausted once visited.

Let us introduce the stochastic variable:

$$\alpha^k = \begin{cases} 1 & \text{if an arrival ended stage } k \\ -1 & \text{if a departure ended stage } k. \\ 0 & \text{otherwise} \end{cases}$$

It depends on the strategies of the players, but if the arrival times are Markovian, as we shall assume, as well as the strategies, it is a Markov process itself.

Let δ^k be the duration of stage k and

$$\kappa_i^k = \begin{cases} 0 & \text{if } x_i^k = \delta^k \text{ and } \max x_{-i}^k > x_i^k \\ 1 & \text{otherwise} \end{cases},$$

i.e., $\kappa_i^k = 1$ if player i remains in the patch beyond the current stage. This yields the following dynamics:

$$\begin{cases} \rho^{k+1} = \rho^k - \Delta_\rho(\rho^k, n^k, \delta^k) =: \Lambda_\rho(\rho^k, n^k, \delta^k) \\ n^{k+1} = n^k + \alpha^k \end{cases},$$

with $\Delta_\rho(\rho, n, \delta)$ a known function that can be derived from the dynamic model of Section 2.1, and which enjoys the following properties:

- $\forall\, \rho, n,\ \Delta_\rho(\rho, n, 0) = 0$,
- $\forall\, \rho, n,\ \delta \mapsto \Delta_\rho(\rho, n, \delta)$ is increasing and concave,
- $\forall\, \rho, n,\ \lim_{\delta\to\infty} \Delta_\rho(\rho, n, \delta) = \rho$.

Each criterion can be expressed as

$$\mathcal{G}_i = \mathbb{E}\left\{ \sum_{k=0}^{K_i} \mathcal{L}(\chi^k, \delta^k) \right\},$$

with

$$\mathcal{L}(\chi, \delta) = \mathcal{L}(\rho, n, \delta) = \frac{q}{n}\Delta_\rho(\rho, n, \delta) - \gamma^*\delta.$$

Previous assumptions made on Δ_ρ yield

- $\forall\, \rho, n,\ \mathcal{L}(\rho, n, 0) = 0$,
- $\forall\, \rho, n,\ \delta \mapsto \mathcal{L}(\rho, n, \delta)$ is concave,
- $\forall\, \rho, n,\ \lim_{\delta\to\infty} \mathcal{L}(\rho, n, \delta) = -\infty$.

To solve the corresponding dynamic game problem via dynamic programming, we introduce the function $V_i^k(\chi)$ which is the optimal expected total future reward for entering stage k in the state χ. We get the following functional equation of dynamic programming:

$$V_i^k(\chi^k) = \mathbb{E}^*\left[\mathcal{L}(\chi^k, \delta^k) + \kappa_i^k V_i^{k+1}(\chi^{k+1})\right] \quad \forall k \le K_i, \tag{2}$$

where $\mathbb{E}^*$ means that we look for a set of strategies which yield a Nash equilibrium at each stage. As the game is surely stationary, V_i does not depend on the stage number k and (2) becomes the following implicit equation:

$$V_i(\rho, n) = \mathbb{E}^*\left[\mathcal{L}(\rho, n, \delta) + \kappa_i V_i(\Lambda_\rho(\rho, n, \delta), n + \alpha)\right] \quad \forall \rho > \rho^*.$$

As a consequence, it suffices to solve the game restricted to one stage to obtain the Nash-optimal strategy in the closed loop. Furthermore, this is surely a war of attrition with a stochastic end time as defined in Appendix A. Indeed the one-stage game can be stated as follows. Let $V_i(\Lambda_\rho(\rho, n, \delta), n) =: \mathcal{V}_i(\delta, n)$ and thus, the game has a utility function

$$U_i(x_i, x_{-i}, \delta) = \mathcal{L}(n, \delta) + \begin{cases} 0 & \text{if } x_i = \delta \text{ and } \max x_{-i} > x_i \\ \mathcal{V}_i(\delta, n) & \text{if } x_i = \delta \text{ and } \max x_{-i} = x_i \\ \mathcal{V}_i(\delta, n+1) & \text{if } \delta < \min\{x_i, x_{-i}\} \\ \mathcal{V}_i(\delta, n-1) & \text{otherwise} \end{cases}.$$

Let $\check{x}$ be such that $\Lambda_\rho(\rho, n, \check{x}) := \rho^*$; it is the time after which a forager, even alone, has no incentive to stay on the patch, i.e., $\mathcal{V}_i(\check{x}, \cdot) = 0$.

Let then $\hat{x} = \arg\max_x \mathcal{L}(n, x)$; both $\hat{x}$ and $\check{x}$ depend on ρ and n.

As a consequence, $\forall n, \forall x > \hat{x}, \mathcal{L}'(n, x) < 0$. Moreover, if there is no departure, the $\mathcal{L}$ function of the next stage is still decreasing. Thus its $\hat{x}$ is zero, and according to Appendix A, its value is zero. Hence if $\delta \in [\hat{x}, \check{x}]$, $\mathcal{V}_i(\delta, n) = \mathcal{V}_i(\delta, n+1) = 0$.

We show in Appendix A that the value of the game is, as in the classical war of attrition, equal to $\mathcal{L}(\hat{x}, n)$. As a consequence,

$$\mathcal{V}_i(x, n-1) = \max_y \mathcal{L}(\Lambda_\rho(\rho, n, x), n-1, y) =: \mathsf{V}(x, n).$$

We therefore obtain the following result.

Theorem 2.1. *The Nash equilibrium of the game (1) is*

$$P^*(x, n) = \begin{cases} 0 & \forall x < \hat{x} \\ 1 - \mathrm{e}^{-\frac{1}{n-1}\int_{\hat{x}}^{x} h(y,n)dy} & \forall x \in [\hat{x}, \check{x}], \\ 1 & \forall x \geq \check{x} \end{cases}$$

with

$$h(x, n) = -\frac{\mathcal{L}'(x, n)}{\mathsf{V}(x, n)}.$$

Hence the solution of the asynchronous foraging problem investigated here, expressed as a closed loop strategy on the number of foragers, is identical to the synchronous problem of [17].

3 Concluding Remarks

3.1 How Does a War of Attrition Influence the Residence Time?

A question that is not addressed by the model is: Does interference, thus a war of attrition, imply that multiple foragers should stay longer on a patch than a lone forager? We cannot answer in a general way.

It is an established fact [17] that a war of attrition causes the forager to stay longer than the "Charnov time." Yet, this Charnov time itself, here $\hat{x}$, depends in a complex fashion on the detailed interference model.

In this respect, the article [7] does not invalidate the theoretical model; on the contrary, this paper seems to corroborate the model of [17], as the larger the number of animals on the patch, the larger is their tendency to leave. First, part of the contestants leave the patch almost immediately; this can be connected to the "$n - K$" of [17]. Then the remaining contestants leave sequentially, as in [17].

3.2 On a Possible Connection with Population Genetics

Up to now, we have focused on mixed strategies in their classical sense: a random strategy x distributed according to a probability density function $p(x)$. Let $p^*(x)$ equalize the opponent's payoff on its spectrum as in a solution of a Nash game.

Note that in a war of attrition, the value of the game is the reward which would have been earned without entering the game. Nevertheless, the Nash solution requires one to play; the question that arises then is: Why should I play if my expected gain is not greater than my guaranteed value? In the context of evolutionary game theory, the answer makes sense: "to prevent the proliferation of any mutant that would alternatively stay longer on the patch." That is, the mutant is equivalent to a cheater in a population commonly and conventionally adopting a simple Charnov-like strategy: by breaking off the convention, it would obtain more fitness and would consequently invade. Note that, in return, adopting such an ESS has no extra cost as the value of the game remains the same.

Evolutionary game theory provides another viewpoint to implement mixed strategies. Instead of considering a monomorphic population playing a common random strategy, let us now consider a polymorphic population in which pure strategies are distributed homogeneously according to p^* (see Appendix B). Since p^* is equalizing, all the individuals of the population can expect the same fitness.

In a population involved in "war of attrition" contests, it simply means that distributing a deterministic persistence time to each individual according to p^* is evolutionarily stable. In other words, a variability in terms of individuals' ability to sustain interference would be expected among the population. Indeed, in this model, interference is taken as a perturbation, not as a decision variable like in a *hawk-dove* contest [5, 6, 13]; interference affects all the contestants equally.

Interestingly, the authors of [22] observed "*the existence of a significant intra-population genetic variability in the susceptibility of females to interference*," acting on the "*time they are willing to invest*." Moreover, there was no significant genetic variability in terms of aggressiveness (unpublished data). Thus these intra-specific interactions seem to be governed by a war of attrition game rather than a hawk-dove one.

However, the connection with these *emigration-threshold genotypes* [15] seems somewhat premature as the stability of the replicator dynamics [10] in a continuous strategy space—as is the case for a war of attrition—is still under investigation [3, 4].

Acknowledgments

This work is part of GDR 2155 "Écologie Comportementale" (CNRS commission 29), and the ESF/BEPAR scientific programme. We thank Minus van Baalen, Carlos Bernstein, Joel S. Brown, Michel de Lara, Jean-Sébastien Pierre, and an anonymous referee for their comments.

Appendix A A War of Attrition with a Stochastic End Time

We consider the following nonzero-sum game:

- n players.
- Player i chooses $x_i \in \mathbb{R}^+$.
- ϵ, the end time, is a positive stochastic variable independent of player decisions.
- The criterion of player i is as follows, where x_{-i} stands for $\{x_j\}$, $j \neq i$:

$$U_i(x_i, x_{-i}, \epsilon) = \begin{cases} L_i(x_i) & \text{if } x_i \leq \min\{x_{-i}, \epsilon\} \text{ and } \max x_{-i} > x_i \\ D_i(x_i) & \text{if } x_i \leq \min\{x_{-i}, \epsilon\} \text{ and } \max x_{-i} = x_i \\ E_i(\epsilon) & \text{if } \epsilon < \min\{x_i, x_{-i}\} \\ W_i(\min x_{-i}) & \text{otherwise} \end{cases}.$$

The hypotheses are: $\forall i$,

- $\exists\,!\,\hat{x} = \arg\max_x L_i(x)$.
- L_i is strictly decreasing for $x > \hat{x}$.
- $W_i(x) > D_i(x) \geq E_i(x) \geq L_i(x)\,\forall x \in [\hat{x}, \check{x})$.
- either $\exists\,\{\check{x} \geq \hat{x} \mid \forall x \geq \check{x}, L_i(x) = W_i(x)\}$,
- otherwise let $\check{x} = \infty$.

We seek a Nash equilibrium, with $P_i(x)$ the cumulative distribution function of player i. We claim the following.

Theorem A.1. *A Nash equilibrium set of strategies must satisfy the following properties:*

- *The Nash-optimal probability density function is continuous over* $[\hat{x}, \check{x})$ *and zero elsewhere but may exhibit a Dirac weight at* $\check{x}$.
- *Let*
$$h_i(x) = -\left\{\frac{P'_\epsilon(x)}{1 - P_\epsilon(x)}\frac{E_i(x) - L_i(x)}{W_i(x) - L_i(x)} + \frac{L'_i(x)}{W_i(x) - L_i(x)}\right\},$$

- $$H_i^*(x) = 1 - \mathrm{e}^{-\int_{\hat{x}}^{x} h_i(y)\,\mathrm{d}y} \quad \forall x \in [\hat{x}, \check{x}],$$
- *and*
$$\mathcal{H}_i(x) := \frac{\prod_{k=1}^{n}[1 - H_k^*(x)]^{\frac{1}{n-1}}}{1 - H_i}.$$
- *The unique Nash-optimal strategy is* $\forall i$*,*
$$P_i^*(x) = \begin{cases} 0 & \forall x < \hat{x} \\ 1 - \mathcal{H}_i(x) & \forall x \in [\hat{x}, \check{x}). \\ 1 & \forall x \geq \check{x} \end{cases}$$

Proof. The hypotheses made clearly show that everyone shares a common *spectrum*, i.e., mixed strategy support, $[\hat{x}, \check{x}]$. Let now P_i, H_i and P_ϵ be the cumulative distribution functions of respectively x_i, $\min x_{-i}$ and ϵ. The generating function is then

$$G_i(x, H_i, P_\epsilon) = \int_{y \in [\hat{x}, \check{x}]} \int_{z \in [\hat{x}, \infty)} U_i(x, y, z)\mathrm{d}P_\epsilon(z)\mathrm{d}H_i(y),$$

$$G_i(x, H_i, P_\epsilon) = \int_{y \in [\hat{x}, x)} \left[\int_{z \in [\hat{x}, y)} E_i(z)\mathrm{d}P_\epsilon(z) + \int_{z \in [y, \infty)} W_i(y)\mathrm{d}P_\epsilon(z)\right] \mathrm{d}H_i(y)$$
$$+ \int_{y \in [x, \check{x}]} \left[\int_{z \in [\hat{x}, x)} E_i(z)\mathrm{d}P_\epsilon(z) + \int_{z \in [x, \infty)} L_i(x)\mathrm{d}P_\epsilon(z)\right] \mathrm{d}H_i(y).$$

As the optimal strategy is equalizing on the opponents' spectrum, in any open set Ω in $[\hat{x}, \check{x})$, one must have

$$\frac{\partial}{\partial x} G_i(x, H_i^*, P_\epsilon) = 0 \quad \forall x \in \Omega.$$

Differentiating $G_i(x, H_i, P_\epsilon)$ yields

$$0 = [E_i(x) - L_i(x)][1 - H_i^*(x)]P_\epsilon'(x)$$
$$+ [1 - P_\epsilon(x)]\left\{L_i'(x)[1 - H_i^*(x)] - [W_i(x) - L_i(x)]H_i^{*\prime}(x)\right\}.$$

Hence
$$H_i^*(x) = 1 - \mathrm{e}^{-\int_{\hat{x}}^{x} h_i(y)\,\mathrm{d}y} \quad \forall x \in [\hat{x}, \check{x}],$$
with
$$h_i(x) = -\left\{\frac{P_\epsilon'(x)}{1 - P_\epsilon(x)} \frac{E_i(x) - L_i(x)}{W_i(x) - L_i(x)} + \frac{L_i'(x)}{W_i(x) - L_i(x)}\right\}.$$
Hence the Nash-optimal strategies are given by

$$\forall i, 1 - H_i^*(x) = \prod_{j \neq i}[1 - P_j^*(x)],$$

where the H_i's are known.

This implies

$$\prod_i [1 - H_i^*(x)] = \prod_i [1 - P_i^*(x)]^{n-1}.$$

Therefore,

$$P_i^*(x) = 1 - \frac{\prod_{k=1}^n [1 - H_k^*(x)]^{\frac{1}{n-1}}}{1 - H_i} =: 1 - \mathcal{H}_i(x) \quad \forall x \in [\hat{x}, \check{x}).$$

Hence we have the unique Nash equilibrium such that

$$\forall i,\ P_i^*(x) = \begin{cases} 0 & \forall x < \hat{x} \\ 1 - \mathcal{H}_i(x) & \forall x \in [\hat{x}, \check{x}) \\ 1 & \forall x \geq \check{x} \end{cases}.$$

An atom of probability takes place on $\check{x}$. Indeed, a Nash equilibrium requires $G_i(x, H_i^*, P_\epsilon) = G_i^* \, \forall x \in [\hat{x}, \check{x})$, where G_i^* is the value of the game. Up to now, we implicitly assumed that H_i was continuous in $[\hat{x}, \check{x})$. Indeed, let $\tilde{x} \in [\hat{x}, \check{x}]$ and suppose that this is a point of discontinuity of amplitude j—for "jump." Per convention, P_i is cadlag. If $\tilde{x} < \check{x}$, $\lim_{x \downarrow \tilde{x}} G_i(x) - G_i(\tilde{x}) = j(1 - P_\epsilon(\check{x}))(W_i(\check{x}) - L_i(\check{x}))$—if the draw is taken into account, in the case where all other foragers have a Dirac at the same $\check{x}$, $L_i(\check{x})$ is replaced by a convex combination of $L_i(\check{x})$ and $D_i(\check{x})$—therefore a Dirac is impossible for any $\tilde{x} < \check{x}$. Moreover, if a jump occurs in H_i at $\check{x}$, $\lim_{x \uparrow \tilde{x}} G_i(x) - G_i(\tilde{x}) = j(1 - P_\epsilon(\check{x}))(L_i(\check{x}) - D_i(\check{x})) = 0$ by the definition of $\check{x}$. Hence a jump is possible on $\check{x}$. To conclude, it is obvious that, from the previous hypotheses on L_i, $\forall x \notin [\hat{x}, \check{x})$, $G_i(x, H_i^*, P_\epsilon) \leq G_i^*$, as $G_i^* = L_i(\hat{x})$.

Hence, if the game is symmetric,

$$P^*(x) = \begin{cases} 0 & \forall x < \hat{x} \\ 1 - \mathrm{e}^{-\frac{1}{n-1}\int_{\hat{x}}^{x} h(y)dy} & \forall x \in [\hat{x}, \check{x}) \\ 1 & \forall x \geq \check{x} \end{cases}.$$

Note that, if $\forall x \in [\hat{x}, \check{x}]$, $P_\epsilon(x) = 0$, the above solution of the war of attrition coincides with the classical solution [1, 2, 8]. □

Appendix B ESS and Classical Game Theory

B.1 Notation and Setup

We consider a compact metric space X as the space of *traits* or *phenotypes* or *pure strategies*. Three cases of interest are

- X is finite (the *finite* case), $X = \{x_1, x_2, \ldots, x_n\}$,
- X is a line segment $[a, b] \subset \mathbb{R}$,
- X is a compact subset of $\mathbb{R}^n$.

We shall use letters x, y for elements of X.

We let $\Delta(X)$ denote the set of probability measures over X. In the finite case, we shall also denote it as Δ_n. We notice that, in the weak topology, $\Delta(X)$ is compact and the mathematical expectation is continuous with respect to the probability law. We shall use letters p, q for elements of $\Delta(X)$.

A population of animals is characterized by the probability $p \in \Delta(X)$ governing the traits of its individuals. There is no need to distinguish whether each individual acts many times, adopting a strategy in $A \subset X$ with probability $p(A)$—the population is then monomorphic and its members are said to use the *mixed strategy* p—or whether each animal behaves in a fixed manner, but in a polymorphic population, where p is the distribution of traits among the population, for any subset $A \subset X$, $p(A)$ is the fraction of the population which has its trait x in A. Then, p also governs the probability that an animal taken randomly in the population behaves a certain way.

We are given a *generating function* $G : X \times \Delta(X) \to \mathbb{R}$ jointly continuous (in the weak topology for its second argument). Its interpretation is that it is the *fitness* gained by an individual with trait x in a population characterized by p.

A case of interest, called hereafter the *linear case*, is when G derives from a function $H : X \times X \to \mathbb{R}$ giving the benefit $H(x, y)$ that an animal with trait x gets when meeting an animal with trait y, according to the expected benefit for trait x:

$$G(x, p) = \int_X H(x, y)\ \mathrm{d}p(y). \tag{B.1}$$

Then G and F below are linear in their second argument. But this is not necessary for many of the results to follow.

The fitness gained by an animal using a mixed strategy q in a population characterized by p is

$$F(q, p) = \int_X G(x, p)\ \mathrm{d}q(x).$$

Note that if $\delta_x \in \Delta(X)$ denotes the Dirac measure at x, $G(x, p) = F(\delta_x, p)$.

The most appealing definition of an ESS is as follows [13].

Definition B.1. The distribution $p \in \Delta(X)$ is said to be an *ESS* if there exists $\varepsilon_0 > 0$ such that for any positive $\varepsilon < \varepsilon_0$,

$$\forall q \neq p^*, \quad F(p, (1-\varepsilon)p + \varepsilon q) > F(q, (1-\varepsilon)p + \varepsilon q).$$

Using only the linearity, it coincides with the original definition of [14].

Theorem B.1. *If F is linear in its second argument, Definition B.1 is equivalent to Definition B.2 below.*

Definition B.2. The distribution $p \in \Delta(X)$ is said to be an ESS if

(I) $\forall q \in \Delta(X), \quad F(q, p) \leq F(p, p)$,
(II) $\forall q \neq p, \quad F(q, p) = F(p, p) \Rightarrow F(q, q) < F(p, q)$.

B.2 Relation to Classical Game Theory

Consider a two-player game between, say, player 1 and player 2. Both choose their action, say, q_1 and q_2, in $\Delta(X)$. Let their respective reward functions, that they seek to maximize, be

$$J_1(q_1, q_2) = F(q_1, q_2),$$
$$J_2(q_1, q_2) = F(q_2, q_1).$$

We have the obvious proposition.

Proposition B.1.

- *Condition B.2 of Definition B.2 is equivalent to the statement that (p, p) is a Nash equilibrium of this game. For that reason, any p satisfying that condition is called a* Nash point.
- *If (p, p) is a strict Nash equilibrium, p is an ESS.*

We immediately have the following, by a theorem due to Von Neumann [21, assertion (17:D) p. 161] in the finite case, and noticed at least since the early 1950s in the infinite case.[2]

Theorem B.2. *Let p be an ESS, then*

(I) $\forall x \in X,\ G(x, p) \le F(p, p)$,
(II) let $N = \{x \in X \mid G(x, p) < F(p, p)\}$, *then* $p(N) = 0$.

A proof completely similar to —but slightly distinct from— the existence proof of the Nash equilibrium lets one state the following result, which applies here.

Theorem B.3. *Let P be a compact space, and let $F : P \times P \to \mathbb{R}$ be a continuous function, concave in its first argument. Then there exists at least one $p \in P$ satisfying condition (I) of Definition B.2.*

B.3 Further Analysis of the Linear Finite Case

B.3.1 Characterization in Terms of the Game Matrix

In the finite linear case, the problem is entirely defined by the matrix $A = (a_{ij})$ with $a_{ij} = H(x_i, x_j)$, as

$$G(x_i, p) = (Ap)_i, \quad F(q, p) = \langle q, Ap\rangle = q^t Ap.$$

We rephrase Theorem B.2 in that context. To do so, introduce the notation **1** to mean a vector—of appropriate dimension—the entries of which are all ones, and

[2] Von Neumann's proof applies to zero-sum games. Its extension to a Nash equilibrium is trivial and can be found, e.g., without claim of novelty, in [11].

the notation for vectors u and v of same dimension $u < v$ to mean that the vector $v - u$ has all its coordinates strictly positive.

We obtain the following more or less classical results.

Theorem B.4. *In the finite linear case, the two conditions of Definition B.2 are respectively equivalent to (I) and (II) below.*

(I) There exists a partition $X = X_1 \cup X_0$, $|X_1| = n_1$, $|X_0| = n_0$, such that, reordering the elements of X in that order and partitioning $\mathbb{R}^n$ accordingly, there exists a vector $p_1 \in \Delta_{n_1}$, a real number α and a vector $h \in \mathbb{R}^{n_0}$ such that

$$p = \begin{pmatrix} p_1 \\ 0 \end{pmatrix}, \quad Ap = \begin{pmatrix} \alpha\mathbf{1} \\ h \end{pmatrix}, \quad h < \alpha\mathbf{1}. \tag{3}$$

(II) Partitioning A accordingly in

$$A = \begin{pmatrix} A_{11} & A_{10} \\ A_{01} & A_{00} \end{pmatrix},$$

$$\forall q_1 \in \Delta_{n_1} \backslash \{p_1\}, \quad \langle q_1 - p_1, A_{11}(q_1 - p_1)\rangle < 0. \tag{4}$$

Note that the vectors $\mathbf{1}$ in the second and third expression of (3) do not have the same dimension. Note also that p_1 may still have some null coordinates.

Proof. For condition (I), this is just a rephrasing of Theorem B.2. Concerning condition (II), the vectors $q \in \Delta(\mathbb{R}^n)$ such that $F(q, p) = F(p, p)$ are all the vectors of the form

$$q = \begin{pmatrix} q_1 \\ 0 \end{pmatrix}, \quad q_1 \in \Delta_{n_1}.$$

As a matter of fact, for all such vectors, $\langle q, Ap\rangle = \alpha$. Thus condition (I) of Definition B.2 says that $\forall q_1 \in \Delta_{n_1} \backslash \{p_1\}$, $\langle q_1 - p_1, A_{11} q_1\rangle < 0$. But we have seen that $\langle q_1 - p_1, A_{11} p_1\rangle = 0$. Therefore, we may subtract that quantity to get (II) above. □

Theorem B.3 says that there always exists at least one solution of equations (3). The question thus is to know whether that solution satisfies condition (II) of the definition. To further discuss that question, let $p_2 \in \mathbb{R}^{n_2}$ be the vector of the nonzero entries of p_1, so that, reordering the elements of X_1 if necessary,

$$p_1 = \begin{pmatrix} p_2 \\ 0 \end{pmatrix}.$$

Let also A_{22} be the corresponding submatrix of A, and for $i = 1, 2$, define $B_i := A_{ii} + A_{ii}^t$ and the $n_i \times (n_i - 1)$-dimensional matrices Q_i obtained by deleting one column from the symmetric projector matrix $P_i := [I - (1/n_i)\mathbf{1}\mathbf{1}^t]$. The condition that the restriction of the quadratic form to the orthogonal subspace to $\mathbf{1}$ be negative definite translates into the following.

Corollary B.1.

- *A necessary condition for a solution of equation (3) to be an ESS is that* $Q_2^t B_2 Q_2 < 0$ *(negative definite).*
- *A sufficient condition is that* $Q_1^t B_1 Q_1 < 0$*, and a fortiori that* $B_1 < 0$*.*

We may note the following fact.

Proposition B.2. *Matrices* B_i*,* $i = 1, 2$ *that satisfy the conditions of Corollary B.1 have at most one nonnegative eigenvalue.*[3]

Another easy corollary is that the number of ESSs is bounded by n. More precisely, we have the following statement.

Corollary B.2. *If there is an ESS in the relative interior of a face, there is no other ESS in that face, and in this statement* Δ_n *is itself an* $n-1$*-dimensional face.—In particular, if there is an ESS in the relative interior of* Δ_n*, it is* the *unique ESS.*

B.3.2 Stability of the Replicator Dynamics

Some authors ([19, 20]) define an ESS—in the finite case—as a stable point p of the replicator dynamics

$$\dot{q}_i = q_i[G(x_i, q) - F(q, q)]. \tag{5}$$

Notice first that a consequence of (5) is that

$$q_i(t) = q_i(0) \exp\left(\int_0^t [G(x_i, q(s)) - F(q(s), q(s)]\, \mathrm{d}s\right)$$

so that if all $q_i(0)$ are non-negative, this is preserved over time. Moreover, one sees that $\sum_i \dot{q}_i = \sum_i q_i G(x_i, q) - (\sum_i q_i) F(q, q) = (1 - \sum_i q_i) F(q, q) = 0$, so that the hyperplane $\{q \mid \sum_i q_i = 1\}$ is invariant. The conclusion of these two remarks is the following:

Proposition B.3. *Under the replicator dynamics,*

- $\Delta(X)$ *is invariant, as well as its interior,*
- *the faces of* $\Delta(X)$ *are invariant as well as their interiors.*

It is known (see, e.g., [16] for a much more detailed analysis) that in the finite linear case, the relationship between these two concepts is as in the next theorem. Note that in the continuous case, the situation is far more complex and still open. In the later case, the evolution equation in $\mathbb{R}^n$ is replaced by one in a measure space, so that the definition of stability depends on the topology used—and the Lyapunov function used here is not continuous in the natural weak topology.

[3] Some authors have mistakenly replaced *at most one* by *exactly one*.

Theorem B.5. *In the finite linear case, every asymptotically stable point of (5) is a Nash point. Every ESS is a locally*[4] *asymptotically stable point of (5), and its attraction basin contains the relative interior of the lowest dimensional face of* $\Delta(X)$ *it lies on.*

Two particular cases of this theorem are as follows.

Corollary B.3. *In the finite linear case:*

- *If an ESS is an interior point of* $\Delta(X)$ *it is globally stable in the interior of* $\Delta(X)$.
- *Every pure strategy, whether an ESS or not, is a rest point of (5). The above theorem implies nothing more for a pure ESS.*

Proof of the theorem. To prove the necessity, assume p is not a Nash point, so that there is an index k such that $p_k = 0$, but $G(x_k, p) > F(p, p)$. Take an initial q with $q_k > 0$. Then it is impossible that $q(t) \to p$, as this would require that $q_k(t) \to 0$, and hence that

$$\int_0^t [G(x_k, q(s)) - F(q(s), q(s))] \, \mathrm{d}s \to -\infty,$$

while in a neighborhood of p the integrand would be positive.

For the sufficiency, restrict the attention to the subspace $\mathbb{R}^{n_2}$ of Corollary B.1, where all coordinates of p are strictly positive, and further to $\Delta := \Delta_{n_2}$. And consider the Lyapunov function

$$V(q) = \sum_i p_i \ln \frac{p_i}{q_i}.$$

It is zero at p. It can be written $V(q) - \sum_i p_i \ln(q_i/p_i)$, and using the fact that $\ln x < x - 1$ as soon as $x \neq 0$, $V(q) > -\sum_i p_i(q_i/p_i - 1) = 0$ as soon as $\mathrm{D} \ni q \neq p$. Thus its restriction to Δ is indeed a valid Lyapunov function. And trivially, on a trajectory,

$$\frac{\mathrm{d}V(q(t))}{\mathrm{d}t} = -\sum_{i=1}^{n_2} p_i[G(x_i, q) - F(q, q)] - F(p, q) + F(q, q)$$

which is by hypothesis negative on Δ_{n_2}. □

As a matter of fact, one can prove more, using the following fact, the proof of which (based upon compactness) we omit.

Definition B.3. A strategy $p \in \Delta_n$ is called *locally superior* if there exists a neighborhood $\mathcal{N}$ of p in Δ_n such that, for any $q \in \mathcal{N}$, $q \neq p$, $F(q, q) < F(p, q)$.

[4] Relative to the face we are referring to.

Theorem B.6. *In the finite linear case, p is an ESS if and only if it is locally superior.*

Corollary B.4. *In the finite linear case, the basin of attraction of an ESS contains a neighborhood in Δ_n of the relative interior of the lowest dimensional face of Δ_n on which that ESS lies.*

References

[1] Bishop D.T., Cannings C.: A generalized war of attrition. *Journal of Theoretical Biology*, **70**:85–124, 1978.

[2] Bishop D.T., Cannings C., Maynard Smith J.: The war of attrition with random rewards. *Journal of Theoretical Biology*, **74**:377–388, 1978.

[3] Cressman R.: Stability of the replicator equation with continuous strategy space. *Mathematical Social Sciences*, **50**:127–147, 2005.

[4] Cressman R., Hofbauer J.: Measure dynamics on a one-dimensional trait space: theoretical foundations for adaptive dynamics. *Theoretical Population Biology*, **67**:47–59, 2005.

[5] Doebeli M., Hauert C.: Models of cooperation based on the prisonner's dilemma and the snowdrift game. *Ecology Letters*. **8**:748–766, 2005.

[6] Dubois F., Giraldeau L.-A.: The forager's dilemma: food sharing and food defense as risk-sensitive foraging options. *The American Naturalist*, **162**:768–779, 2003.

[7] Goubault M., Outreman Y., Poinsot D., Cortesero A.M.: Patch exploitation strategies of parasitic wasps under intraspecific competition. *Behavioral Ecology*, **16**: 693–701, 2005.

[8] Haigh J., Cannings C.: The n-person war of attrition. *Acta Applicandae Mathematicae*, **14**:59–74, 1989.

[9] Hamelin F., Bernhard P., Nain P., Wajnberg E.: Foraging under competition: evolutionarily stable patch-leaving strategies with random arrival times. 1. Scramble competition. *Annals of Dynamics Games, this volume, Birkhauser*, pp. 327–348, 2007.

[10] Hofbauer J., Sigmund K.: Evolutionary games and population dynamics. Cambridge University Press, Cambridge, UK, 1998.

[11] Kuhn H.W.: Contribution to the theory of games. Volume 24 of *Annals in Mathematics Studies*, Princeton University Press, Princeton, New Jersey, USA, 1950.

[12] McNamara J.M., Houston A.I., Collins E.J.: Optimality models in behavioral biology. *SIAM Review*. **43**: 413–466, 2001.

[13] Maynard Smith J.: Evolution and the theory of games. Cambridge University Press, Cambridge, UK, 1982.

[14] Maynard Smith J., Price G.R.: The logic of animal conflict. *Nature*, **246**: 15–18, 1973.

[15] Parker G.A., Stuart. R.A.: Animal behaviour as a strategy optimizer: evolution of resource assessment strategies and optimal emigration thresholds. *The American Naturalist*, **110**:1055–1076, 1976.

[16] Samuelson L.: Evolutionary games and equilibrium selection. MIT Press, Cambridge, Massachusetts, USA, 1997.

[17] Sjerps M., Haccou P.: Effects of competition on optimal patch leaving: a war of attrition. *Theoretical Population Biology*, **3**:300–318, 1994.

[18] Van Der Meer J., Ens B.J.: Models of interference and their consequences for the spatial distribution of ideal and free predators. *Journal of Animal Ecology*, **66**:846–858, 1997.

[19] Vincent T.: Evolutionary games. *Journal of Optimization Theory and Applications*, **46**:605–612, 1985.

[20] Vincent T., Fisher M.E.: Evolutionary stable strategies in differential and differences equations models. *Evolutionary Ecology*, **2**:321–337, 1988.

[21] Von Neumann J., Morgenstern O.: The theory of games and economic behaviour. Princeton University Press, Princeton, New Jersey, USA, 1944.

[22] Wajnberg E., Curty C., Colazza S.: Genetic variation in the mechanisms of direct mutual interference in a parasitic wasp: consequences in terms of patch-time allocation. *Journal of Animal Ecology*, **73**:1179–1189, 2004.

Evolution of Corn Oil Sensitivity in the Flour Beetle

R. C. Rael
Program in Applied Mathematics
University of Arizona
Tucson, AZ 85721, USA
rrael@math.arizona.edu

T. L. Vincent
Department of Aerospace and Mechanical Engineering
University of Arizona
Tucson, AZ 85721-0119, USA
vincent@u.arizona.edu

R. F. Costantino
Department of Ecology and Evolutionary Biology
University of Arizona
Tucson, AZ 85721, USA
rfc@email.arizona.edu

J. M. Cushing
Department of Mathematics, Program in Applies Mathematics
University of Arizona
Tucson, AZ 85721, USA
cushing@math.arizona.edu

Abstract

We explore the persistence of corn oil sensitivity in a population of the flour beetle *Tribolium castaneum* using evolutionary game methods that model population dynamics and changes in the mean strategy of a population over time. The strategy in an evolutionary game is a trait that affects the fitness of the organisms. Corn oil sensitivity represents such a strategy in the flour beetle. We adapt an existing model of the ecological dynamics of *T. castaneum* into an evolutionary game framework to explore the persistence of corn oil sensitivity in the population. The equilibrium allele frequencies resulting from the evolutionary game are evolutionarily stable strategies and compare favorably with those obtained from the experimental data.

1 Introduction

An evolutionarily stable strategy (ESS) is the endpoint of natural selection and is resistant to changes in the frequencies of alleles that affect fitness [8]. When a population reaches a local fitness maximum in an adaptive landscape determined by allele frequencies (adaptive peak), natural selection will preserve the allele frequencies associated with that adaptive peak, as any change in those allele frequencies will result in a lower population mean fitness. Classical population genetics techniques can be used to predict equilibrium allele frequencies in an ESS, but they do not model the population dynamics.

Evolutionary game theory (EGT) can model both genetic and population dynamics simultaneously, allowing us to study an ESS in systems as they are influenced by changing population sizes. The three main elements of an evolutionary game are players, strategies, and payoffs. The players are the individual organisms, the strategies are heritable phenotypes, and the payoffs are relative fitness gains or losses, usually expressed as net reproductive rates of the individuals. Although EGT is usually used to model dynamics in a Darwinian sense, where strategies are phenotypes, we use it here to estimate allele frequencies with genotypic strategies. In this case, the ESSs are allele frequencies—genetic "strategies"—that maximize population mean fitness. Thus, EGT provides a quantitative prediction of an ESS in addition to the population size when this strategy has been reached.

The evolutionary game has two parts. One part involves ecological dynamics as influenced by the strategy that governs fitness; the other part involves the evolutionary process resulting in changes in strategy frequency [8]. The parts are coupled; the ecological dynamics influence the change in strategy frequencies, and changes in strategy frequencies influence payoffs to the players. The ecological dynamics of the flour beetle *Tribolium castaneum* (Herbst) have been extensively studied experimentally and theoretically, and well-established models supported by experimental data have been developed [2][4][5][6]. The strategy dynamics portion of the game is provided by EGT, which models changes in strategy frequencies over time as they are influenced by the ecological dynamics.

In this study, the strategy is a genetically determined phenotype in *T. castaneum* that results from variation in a single gene that controls sensitivity to corn oil. There are two forms of the gene (alleles) that control corn oil sensitivity. Individuals homozygous for the *cos* allele are unable to properly digest the unsaturated fatty acid present in corn oil [1]. This adaptive trait therefore imparts variable rates of fecundity and mortality among flour beetles cultured on a corn oil substrate [3][4][5]. Individuals may have one of three possible genotypes: *cos/cos*, *cos*/+, or +/+, where + is the wild type allele. We define strategy as a numerical representation of genotype where *cos/cos* individuals have strategy 0, *cos*/+ individuals strategy 0.5, and +/+ individuals strategy 1. This allows us to interpret the mean strategy of the population as the frequency of the + allele. A difference equation model modified to reflect the influence of corn oil sensitivity on fitness is used for the population dynamics of *T. castaneum* in this study [6]. We use this model in an evolutionary game setting to explore the persistence of corn oil sensitivity in a flour beetle population and compare the results of this game to experimental data.

2 Model and Data

The experimental data we use in this study are from a demographic and genetic experiment conducted by Desharnais and Costantino [5]. In that study, twenty-two cultures of *T. castaneum* homozygous for the *cos* allele were initiated with identical demographic distributions of 70 small larvae, 20 large larvae, 16 pupae, and 64 adults. Each population was contained in a one-half pint milk bottle with 20 grams of corn oil medium (90% wheat flour, 5% dried brewers yeast, 5% liquid corn oil), and all were kept at constant environmental conditions. Populations were censused biweekly for 80 weeks. After ten weeks, three replicates were randomly assigned to each of three treatments that introduced the + allele into *cos/cos* homozygous populations. The three treatments included the addition of one female adult, three female adults, or three male and three female adults with the $+/+$ genotype. We refer to these as treatments 1, 2, and 3, respectively. In the control treatment, no manipulations were imposed. Gene frequency data were recorded every four weeks for 24 weeks after perturbation. We use gene frequency and population size data averaged over three replicates for each treatment in which + alleles were added, and averaged over four replicates for the control treatment.

The flour beetle goes through three distinct life history stages upon emerging from an egg: larva, pupa, and adult. Individuals spend approximately two weeks in both the larval and pupal stages. An important biological characteristic of these organisms is cannibalism. Adults eat eggs and pupae, and larvae consume eggs [2]. This characteristic imparts density-dependent regulation on the population size [2]. Several fixed-strategy models exist that do a good job of predicting population sizes in this organism, including models that incorporate multiple life history stages Dennis *et al.* [4].

The model we used as the basis for the first part of the evolutionary game is a single-stage version of the LPA (larvae-pupae-adult) model developed by Dennis *et al.* [4]. This is given by

$$x(t+1) = b\exp(-cx)x + (1-\mu)x. \tag{1}$$

In this model, x is the number of adults at time t, b is the rate of recruitment into the adult class (or the probability of an individual surviving to adulthood), c is the cannibalism rate, and μ is the adult death rate. The term $\exp(-cx)$ is the probability that a potential recruit into the adult stage is not cannibalized between times t and $t+1$. We used parameter values derived from experimental data. One time step represents two weeks in the model.

To cast this system in the standard form for an evolutionary game, we first restate the ecological dynamics equation in terms of a fitness function [8]. This part of the evolutionary game is expressed as

$$x_i(t+1) = x_i[1 + H_i(\mathbf{u}, \mathbf{x})],$$

where $H_i(\mathbf{u}, \mathbf{x})$ is the fitness function for population i, which represents the per capita change in population density from one time step to the next [8]. In this equation, $\mathbf{x}$ denotes the population vector, and $\mathbf{u}$ denotes the strategy vector, so that the mean strategy of population i is u_i. The fitness generating function, or G-function provides a convenient way of expressing all of the H_i functions. The G-function is defined as follows [8]:

$$G(v, \mathbf{u}, \mathbf{x})|_{v=u_i} = H_i(\mathbf{u}, \mathbf{x}).$$

The variable v is a "virtual" variable. When v is replaced by some u_i, the G-function gives the fitness function for population i. The model becomes

$$x_i(t+1) = x_i[1 + G(v, \mathbf{u}, \mathbf{x})|_{v=u_i}].$$

We evaluated the evolution of a single population dependent on a scalar strategy, and thus used the following equation for the ecological dynamics:

$$x(t+1) = x[1 + G(v, u, x)|_{v=u}], \tag{2}$$

where the G-function, as obtained from equation (1), is given by

$$G(v, u, x) = b(v)\exp(-cx) - \mu(v).$$

We assume that the rate of cannibalism is unaffected by the strategy of the beetle and use the constant value $c = 0.028$ [5]. Sensitivity to corn oil does, however, affect reproduction and mortality rates, so the model is modified to account for these effects by making the recruitment and death rates (b and μ) functions of strategy v. We note that the G-function depends on the strategy; however, the fitness of beetles using strategy u_i is unaffected by strategies used by other beetles. The dependence of b and μ on the strategy was determined from experimental data. The average values of b and μ are known for each of the three possible genotypes [5][6]. We fit a quadratic equation to the three data points for each parameter as shown in Figure 1 to determine the functions $b(v)$ and $\mu(v)$. These are defined as

$$b(v) = -18v^2 + 21v + 4 \tag{3}$$

$$\mu(v) = 0.1v^2 - 0.13v + 0.11. \tag{4}$$

These are interpreted as the average recruitment and death rates of a population with a + allele frequency of v. When modeled this way, a population with a + allele frequency of approximately 0.583 has the highest reproduction rate, and one with a frequency 0.65 has the lowest mortality.

Next we define the second part of the game, which produces the strategy dynamics. The standard form for the strategy dynamics according to EGT is given by [8]

$$u(t+1) = u + \left[\frac{\sigma^2}{1 + G(v, u, x)|_{v=u}}\right] \left.\frac{\partial G(v, u, x)}{\partial v}\right|_{v=u}. \tag{5}$$

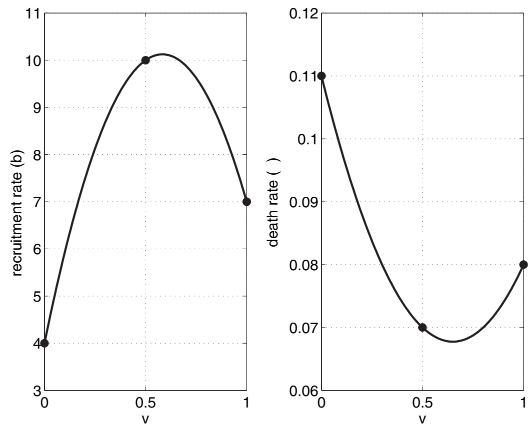


Figure 1: Recruitment (b) and death (μ) rates as functions of strategy (v).

This is derived under the assumption that the rate of change of a trait over time is proportional to the amount of variation present in the population. The strategy at time t is given by u. The parameter σ^2 represents a small variance in strategies from the mean. The data indicate that this parameter ranges from 0.02 to 0.16. We found that the model fit the data well using values of σ^2 in or near this range, including 0.14, 0.2, and 0.3. These are the values we used in the model for comparing to treatments 1, 2, and 3, respectively. Combining the ecological dynamics (2) with the strategy dynamics (5) gives the complete evolutionary game model:

$$\begin{aligned} x(t+1) &= x[1 + G(v,u,x)|_{v=u}] \\ u(t+1) &= \left[u + \frac{\sigma^2}{1 + G(v,u,x)|_{v=u}}\right] \left.\frac{\partial G(v,u,x)}{\partial v}\right|_{v=u}. \end{aligned} \tag{6}$$

In this model, the mean strategy of a population may be interpreted as the frequency of the + allele. Running the complete evolutionary game (6) yields changing population sizes and fitnesses as they are influencing one another. Using this model, the population evolves toward an ESS that in this case maximizes fitness.

3 Results

The EGT model yields results that are reasonably consistent with experimental data. Although the model tended to predict that the population size and strategy reach equilibrium slightly faster than the treatments, they both reached values that are reasonably close to the data. Figures 2 and 3 show comparisons of model results to experimental data for each of the three treatments. The treatments varied by initial allele frequencies. Figure 2 also includes the population dynamics produced by the model in the absence of the + allele along with the data for the control

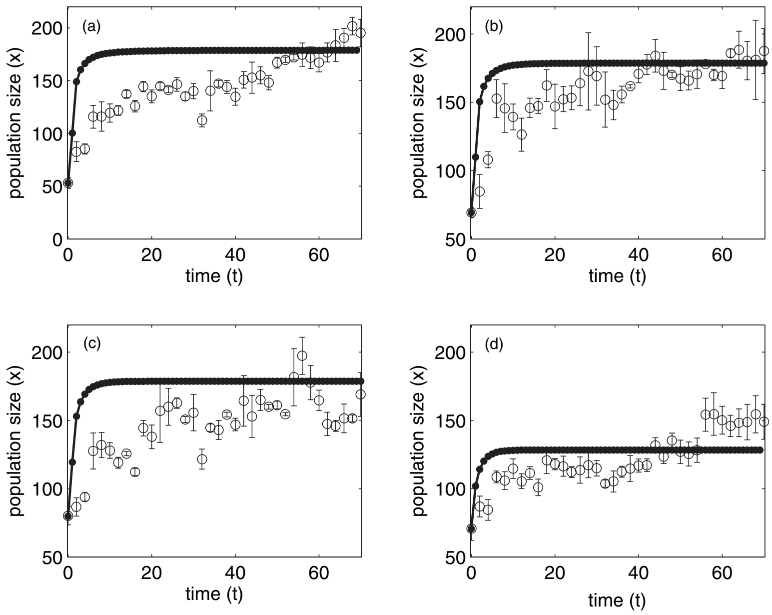


Figure 2: Model results (filled circles) versus experimental data (open circles) with mean standard errors for population dynamics. Treatment 1 (a), treatment 2 (b), treatment 3 (c), control (d). Time represents weeks after the introduction of the + allele, except in the control.

treatment. Figure 3 shows the model results for allele frequency (strategy) along with the mean data over the three treatments. Each data point in these figures is an average over three replicates of the given treatment with the exception of the control treatment data, which are averages over four replicates.

The model is run using the initial values for population size and + allele frequency corresponding to those of the treatments upon addition of + alleles into the population. These values along with those we used for the variance parameter σ^2 are given in Table 1. In addition to the values used for comparing model results to each of the treatments, Table 1 includes the values we used in the model run that we compare to the mean of treatments 1, 2, and 3, and the values we used in the model run for comparison to the control treatment, in which gene frequencies do not change.

For the three genetically segregating populations with initial + allele frequencies of 0.02, 0.04, and 0.08, the model predicts an equilibrium allele frequency of approximately 0.61 with an equilibrium adult population size of approximately 179. In the model, the population size had reached equilibrium at time 70. The homozygous (*cos*/*cos*) control treatment remains genetically unchanged and is forecast by the model to have an equilibrium adult population size of 128. Although all populations had reached equilibrium in the model at time 70, the gene frequencies had not. The population sizes and gene frequencies from the model at the final times for which we have data are listed in Table 2.

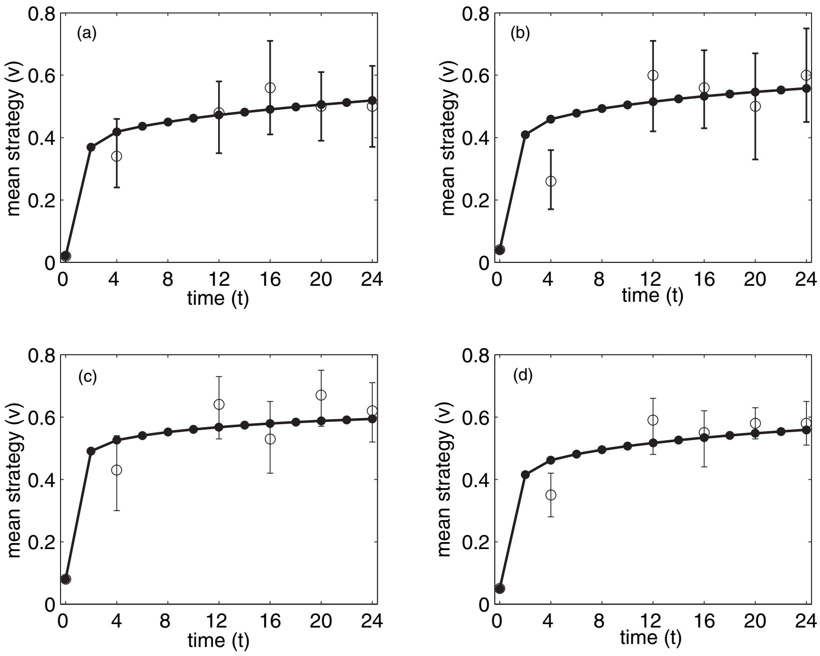


Figure 3: Model results (filled circles) versus experimental data (open circles) with 95% binomial confidence intervals for strategy dynamics. Treatment 1 (a), treatment 2 (b), treatment 3 (c), and mean over these three treatments (d). Time represents weeks after introduction of the + allele. Data in this figure correspond to the first 24 weeks of population data in Figure 2.

Table 1: Parameter and initial values used in the model.

	Treatment				
	Control	1	2	3	Mean
Initial population size	71	53	69	80	64
Initial frequency (+)	0	0.02	0.04	0.08	0.05
σ^2	–	0.14	0.2	0.3	0.2

Table 2: Summary of model results versus data at final data times.

	Treatment				
	Control	1	2	3	Mean
+ Frequency, 24 weeks (model)	0	0.57	0.59	0.61	0.59
+ Frequency, 24 weeks (data)	0	0.50	0.60	0.62	0.58
Pop. size, 70 weeks (model)	128	179	179	179	179
Pop. size, 70 weeks (data)	149	195	188	169	184

4 Discussion

Evolutionary game theory can be applied to the maintenance of polymorphisms within populations [8]. While this is typically done with strategies representing phenotypes, our results show that evolutionary games are potentially useful tools for predicting long-term allele frequencies as well. Using genotype as the adaptive trait, the EGT model produced allele frequencies and population sizes similar to those found in experimental data. These data suggest that there is a stable polymorphism in the population; that is, multiple alleles persist due to a fitness advantage conferred by the heterozygous (*cos*/+) genotype. The population in the evolutionary game reaches an equilibrium due to the population density regulation imposed by the cannibalism in the model [2]. This results in an ESS that in this model also represents the strategy of highest population fitness. This agrees with the experimental data, which support the theoretical prediction that selection is expected to move an ecologically stable system toward an ESS resulting in the maximization of adult numbers [5][6]. This prediction assumes that selection is not frequency dependent. Thus, our model provides a quantitative prediction of an ESS in addition to the population size when the population has reached this strategy. Furthermore, since the + allele frequency at the ESS is less than one, the *cos* allele is still present in the population, which supports the existence of a stable polymorphism at this locus.

Our model produces higher equilibrium population sizes when + alleles are introduced than when no manipulations are imposed, as supported by the data [5][6]. This is an important pattern to note, as it supports the theory that populations have the ability to reach a larger size if they have a higher mean fitness [5]. Because the ESS in this case is the strategy of highest fitness, a population at this ESS reaches the largest population possible as determined by this maximum fitness. The individual fitness conferred by the strategy (sensitivity to corn oil) is not influenced by the strategy of others. For this reason, the G-function we use is solely density dependent.

Genetic variation in a population directly affects the rate of evolution, or change in mean strategy of a population [8]. Since higher amounts of variation in a population are associated with faster evolution, it is expected that introducing different amounts of the + alleles would produce different rates of evolution. The rate of evolution in the model is captured by the parameter σ^2, as strategy dynamics scale with variance [8]. However, there is some difficulty in estimating the variance parameter σ^2 directly from the data. The values of σ^2 that provided good fits to the frequency data were at the upper end of the range of variance in the data and slightly higher. However, values that provided good fits to the data agreed with the initial pattern of variance introduced. That is, treatment 1 had the lowest initial variance, followed by treatment 2, then treatment 3. The values of σ^2 in the model that produced good fits to the data followed this same pattern.

In this model, the mean strategy corresponds to the frequency of the + allele in the population with our numerical convention for representing genotypes as strategies. This is illustrated as follows. If we let P_1, P_2, and P_3 denote the number of individuals with the *cos/cos*, *cos*/+, and +/+ genotypes, respectively, then the frequency of the + allele (p) is given by [7]

$$p = \frac{0.5P_2 + P_3}{\sum_{i=1}^{3} P_i}.$$

This is also the mean strategy of the population given that the strategies of the above genotypes are 0, 0.5, and 1, respectively. This convention therefore allows us to apply EGT in order to determine long-term gene frequencies.

We expect that replacing the model of population dynamics used in the evolutionary game with a multi-stage model will provide a more realistic evolutionary game. We will use the LPA model for *T. castaneum*, which is a difference equation model that includes ecological dynamics equations for the larval and pupal stages of the flour beetle in addition to the adult stage [4]. By incorporating this model into the evolutionary game, demographically detailed population sizes can be predicted, and the allele frequency dynamics will reflect the influence of all life history stages. The game theory model could then be extended further to incorporate multiple interacting flour beetle populations.

While classical genetic models may predict equilibrium allele frequencies based on fitness, the evolutionary game gives a more detailed picture of allele frequency dynamics by incorporating population dynamics. These represent two characteristics of populations that significantly influence one another, ultimately playing a role in determining the ESS. In this study, the theoretical evolutionary game model is connected to experimental data, and we obtain a quantitative estimate for an ESS along with an equilibrium population size for *T. castaneum*. This study shows the utility of EGT models for predicting ESS using genetic information in the modeling of population dynamics.

Acknowledgment

R. R. thanks George Aldridge for critical reading of the manuscript.

References

[1] Costantino, R. F., A. E. Bell, and J. C. Rogler. *Genetic control of lipid metabolism in Tribolium*. Nature 210: 221–222, 1966.

[2] Costantino, R. F., and R. A. Desharnais. *Population Dynamics and the* Tribolium *Model: Genetics and Demography*. Springer-Verlag, Berlin and New York, 1991.

[3] Cushing, J. M., R. F. Costantino, B. Dennis, R. A. Desharnais, and S. M. Henson. *Chaos in Ecology: Experimental Nonlinear Dynamics*, Academic Press, New York, 2003.

[4] Dennis, B., R. Desharnais, J. M. Cushing, and R. F. Costantino. *Nonlinear demographic dynamics: Mathematical models, statistical methods, and biological experiments*. Ecological Monographs. 65(3):261–281, 1995.

[5] Desharnais, R. A., and R. F. Costantino. *Genetic analysis of a population of* Tribolium. *VII. Stability: Response to Genetic and Demographic Perturbations*, Canadian Journal of Genetics and Cytology. 22(4):577–589, 1980.

[6] Moffa, A. M., and R. F. Costantino. *Genetic analysis of a population of* Tribolium*. VI. Polymorphism and demographic equilibrium*. Genetics 87: 785–805, 1977.

[7] Russel, P. J. *Fundamentals of Genetics*, 2nd ed. Addison Wesley Longman, Inc., San Francisco, 2000.

[8] Vincent, T. L., and J. S. Brown, *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics*, Cambridge University Press, Cambridge, 2005.

The Evolution of Gut Modulation and Diet Specialization as a Consumer-Resource Game

Christopher J. Whelan
Illinois Natural History Survey
Midewin National Tallgrass Prairie
30239 South State Highway 53
Wilmington, IL 60481, USA
whelanc@uic.edu

Joel S. Brown
Department of Biological Sciences
University of Illinois at Chicago
M/C 066, 845 West Taylor Street
Chicago, IL 60607, USA

Jason Moll
Department of Biological Sciences
University of Illinois at Chicago
M/C 066, 845 West Taylor Street
Chicago, IL 60607, USA

Abstract

Diet provides an important source of niche partitioning that promotes species coexistence and biodiversity. Often, one species selects for a scarcer but more nutritious food (Thomson gazelle) while another opportunistically consumes low- and high-quality foods indiscriminately (African buffalo). In addition to choosing a diet (selective versus opportunistic), organisms have co-adapted digestion physiologies that vary in size and the throughput rate at which food passes through the gut. We combine these elements into a game of resource competition. We consider a vector-valued strategy with elements of gut size and throughput rate. To the forager, food items now have three properties relating to the value of a particular strategy: profitability (energy gained per unit handling time), richness (energy gained per unit bulk), and ease of digestion (energy gain per unit of passage time). When foraging on foods that differ in profitability, richness, and ease of digestion, adjustment or modulation of gut size and throughput rate leads to digestive-system specialization. Modulation of digestive physiology to a particular food type causes different food types to become antagonistic resources. Adjustment of gut volume and processing thus selects for different degrees of diet specialization or opportunism, and thus may promote niche diversification. This in turn sets the stage for disruptive or divergent selection and may promote sympatric speciation.

1 Introduction

Feeding involves both external and internal processing. Critical external processes include search and pre-consumption handling and encompass the typical domains of foraging ecology. Internal processes include digestion and absorption, the domain of digestive physiology. While foraging and digestion are clearly coordinated processes, foraging ecologists and digestive physiologists typically operate in relative isolation from each other [19, 20]. Holling [4] developed a model of feeding rates, known as the disc equation. This equation gives feeding rate as influenced by random search, prey handling, and prey abundance. We present an extension of Holling's disc equation that incorporates passage rate of food through the gut as an integral component of total food handling time [19]. The model thus integrates both ecological and physiological mechanisms.

Aspects of feeding vary during the annual cycle of many foragers. For instance, diet selection may vary as resource abundances vary seasonally. Food intake rates may vary in response to ecological (e.g., competition, predation risk) and/or physiological (e.g., reproduction, thermoregulation) demands. Individual foragers of many species adaptively modify both ecological strategies and physiological processing capacities in relation to such changes in diet. In our model, we allow both gut morphology (gut volume) and gut processing time (retention time of food in the gut) to flexibly vary in relation to changes in ecological opportunities or physiological demands. Adjustment of gut volume and processing may select for different degrees of diet specialization or opportunism, and thus may promote niche diversification. We combine these elements into a resource game [18] that offers the animals a vector-valued strategy with elements of gut size and throughput time. To the forager, food items now have three salient properties: profitability (energy gained per unit handling time), richness (energy gained per unit bulk), and ease of digestion (energy gain per unit of passage time). By adding size and passage time to digestion, the model increases the likelihood of species coexistence. The model promotes more extreme niche partitioning than models that do not combine external (searching for and ingesting food items) and internal foraging processes (digestion).

2 The Model

We envision an animal that must: 1) search for food, 2) spend time handling encountered food items as part of consumption, and 3) must spend time digesting the food. We assume that time spent searching for food and handling food are mutually exclusive activities. On the other hand, we assume that time spent digesting food is only partially exclusive. When the animal's gut is almost empty, we assume that almost all time spent digesting can be done concurrently with searching for and handling new food items. When the animal's gut is almost full, we assume that almost all the digestive time is now mutually exclusive of search and handling.

To combine these assumptions into a model of feeding, we begin with a modification of the type II functional response [4], in which we include terms for external (pre-consumption) and internal (post-consumption) handling of food:

$$H = (ay)/\{1 + ay[h + gm(b)]\}, \tag{1}$$

where H is harvest rate, a is encounter probability, and y is resource abundance or density. External handling, h, is identical to that in the original disc equation. Internal handling consists of two variables. The first, g, represents the actual processing of food within the gut, and the second, $m(B)$, represents the proportion of gut handling time that is exclusive of alternative foraging activities. External handling, h, and internal handling, g, have units of (time $\times$ item^{-1}). Internal food processing, g, is determined by the quotient of food bulk per item, b (ml $\times$ item^{-1}) and the volumetric flow rate of food through the gut, V_o (ml $\times$ item $^{-1}$): $g = b/V_o$. But V_o = gut capacity, u_1 (ml), divided by retention or throughput time, u_2 (time) [see 6, 11]. Thus passage time per item is given by $g = (bu_2)/u_1$. Exclusive internal handling time, $m(B)$, can take any monotonically increasing functional form. For simplicity, let $m(B) = B$ (a linear function), the proportion of gut volume occupied by food. Gut fullness, B, is given by the bulk rate of intake (bulk of the resource, b, multiplied by its ingestion or harvest rate, H) and the retention time of food in the gut (the quotient of throughput time, u_2, and gut volume, u_1): $B = (bHu_2)/u_1$. This definition of $m(B)$ allows the exclusivity of internal handling to be a continuous, sliding scale that reflects the extent to which gut volume is filled from food consumption. Substituting g and B into (1) and simplifying yields

$$H = (ay)/\{1 + ay[h + b^2H(u_2^2/u_1^2)]\}. \tag{2}$$

This model can be solved explicitly for H:

$$H = \frac{-(1 + ahy) + [(1 + ahy)^2 + 4a^2b^2(u_2^2/u_1^2)y^2]^2}{2ab^2(u_2^2/u_1^2)y}. \tag{3}$$

This explicit expression obscures the way in which external and internal food handling influence the forager's consumption rate.

Expression (2) has three interesting consequences, and we explore its properties graphically (Figure 1). For instance, we now see the intimate connection between harvest rate and gut processing: we need the harvest rate to specify the gut-processing rate, and we need the gut-processing rate to specify the harvest rate. Harvest rate and gut processing mutually feed back onto each other. Second, (2) shows transparently that both pre- and post-consumption food handling limit harvest rate, but they do so jointly. Third, we see that external handling and internal handling are qualitatively different phenomena. External handling, h, is a constant cost per item consumed that is paid in time—it operates qualitatively like a batch reactor [9] that is full (on) or empty (off). Internal handling, $gB = g(bHu_2)/u_1$, in contrast, is a varying cost paid in time because one component, harvest rate, H, is

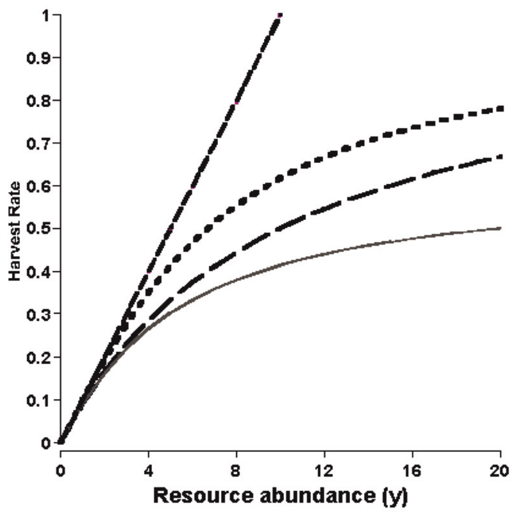


Figure 1: Harvest rate in relation to resource abundance. Straight line occurs when the forager consumes, without handling, all food items encountered (the number of resources harvested equals the encounter probability times resource density). Handling, either internally (Gb – dotted line), externally (h – dashed line), or both ($h + gB$ – solid line), decreases harvest rate.

continuous (see also 6). In other words, internal handling operates as a continuous reactor, such as a plug-flow reactor [9].

3 Digestion as a Resource Game

In the absence of resource dynamics, our functional response model is a simple optimization model. By assuming that natural selection favors the strategy that maximizes the animal's feeding rate, we can find the combination of gut size, u_1, and passage time, u_2, that maximizes equation (3). This is not yet a game.

However, if we consider a population of animals and a pool of resources, then the feeding rates of the animal population can influence the abundance of resources. Furthermore, the abundance of resources influences the animals' feeding rates and their changes in population size. Hence the strategies of other animals influence the abundance of resources, which in turn influences the feeding rate of any focal individual. By introducing a population dynamic on the feeding animals and a dynamic on their resource abundance, we create a consumer-resource game. As a game, we establish a link between the strategies of the animal population and equilibrium resource abundance, y^*. Instead of simple optimization, we now seek an evolutionarily stable strategy (ESS) for the game. In finding the ESS, we rely on the ESS maximum principle and not on the adaptive dynamics approach. Cressman [2] and Garay [3] explore links between these approaches. The nature of our consumer-resource model ensures that the criterion for invasion resistance [3] also produces convergence stability [18].

In our initial model [19], we were interested in determining how fixed characteristics of the forager, h, u_1, and u_2, influence optimal harvest rate and diet selection. We now consider the case in which gut volume, u_1, and gut throughput time, u_2, adjust appropriately to diet quality or quantity [7, 8, 14]. Animals adjust gut volume optimally because guts are costly to maintain (too large a gut wastes energy), but they need sufficient capacity to fuel the forager's energetic and nutrient requirements. Guts need to be big enough, but not too big. We let the cost of maintenance include fixed and variable costs:

$$\gamma = c + \beta k, \tag{4}$$

where γ is the total cost, c is the fixed cost (of all tissue, including the gut), and β is the variable cost of the gut, which increases linearly with gut volume.

Similarly, gut throughput time will be optimally adjusted to diet. Too short a throughput time will result in inadequate breakdown and/or absorption of foods. Too long a throughput time will decrease the forager's harvest rate. We let the forager adjust its throughput rate via its role in energy and nutrient absorption and conversion to forager biomass with Michaelis–Menten kinetics:

$$E = (\alpha E_{\max} u_2)/(\chi + \alpha u_2). \tag{5}$$

In (5) $E_{\max}$ is the maximum rate of conversion of resource to forager biomass, α is the rate of absorption in the intestines, χ is a saturation constant, and u_2 is the throughput time.

Given these relationships, we define a fitness-generating function, or G-function, as a function of net profit, π:

$$G = F(\pi), \tag{6}$$

where F is a monotonically increasing function in π and has the property that when $\pi = 0$, $F = 0$. Net profit, π, is the energy gained from foraging less the cost of maintenance:

$$\pi = EH - \gamma. \tag{7}$$

Next we define ecological dynamics for the forager and its resources. For the forager, let population growth rate be determined by the current population density, x, multiplied by the per capita rate of growth (the fitness-generating function):

$$dx/dt = xG. \tag{8}$$

For the resource, ecological dynamics result from resource renewal less consumption by the forager:

$$dy/dt = r(K - y) - xH. \tag{9}$$

The forager's population will be at equilibrium when $dx/dt = 0$, which occurs if and only if $EH = \gamma$. When the forager's population reaches equilibrium, it will crop the resource density to the minimal level that will sustain it, which we designate as y*. Using the condition that $EH = \gamma$, we can then solve for y^*:

$$y^* = (\gamma E u_1^2)/a[E^2 u_1^2 - \gamma u_2^2 b^2 - \gamma h E u_1^2]. \tag{10}$$

Now, substituting in the fixed and variable costs of gut maintenance for γ, we have

$$y^* = [((c + \beta k) E u_1^2)/a[E^2 u_1^2 - (c + \beta k) u_2^2 b^2 - (c + \beta k) h E u_1^2]. \tag{11}$$

In (11), y^* is now a function of our parameters specifying fixed characteristics of the forager and its resource, and the control variables, u_1 and u_2. We can now specify that the G-function is a function of the forager's own strategies, the strategies used by all other members of the population, at the equilibrium population density of the species:

$$G(\mathbf{v},\mathbf{u},\mathbf{x}) = F(\pi) \tag{12}$$

and

$$F(\pi) = -[1 + ahy^*(u)] + \frac{[1 + ahy^*(u)]^2 + 4a^2b^2(v_2^2/v_1^2)y^*(u)^2]^{1/2}}{2ab^2(v_2^2/v_1^2)y^*(u)}. \tag{13}$$

The first-order necessary condition for the ESS is that $dG/dv = 0$ when $\mathbf{v} = \mathbf{u}^*$, evaluated at y^*. This indicates that the evolutionarily stable gut volume and throughput strategies are those used by all other members of the species at the population density that results in y^*. Note that minimizing y^* is equivalent to maximizing G.

4 Numerical Analysis and Interpretation

We used MATLAB to solve simultaneously for y^*, u_1^*, and u_2^* for 16 foods that represent extreme values of food properties, including all combinations of high and low energetic values of food, $E_{\max}$, food bulk, b, external handling time, h, and rate of absorption, α (see Table 1 in the Appendix). We allowed each of these food properties to vary by a factor of 10. This exercise shows the range of gut system adjustment when the forager consumes foods that differ qualitatively, as might be expected of an animal that consumes food types that may vary in abundance seasonally, that occur in different microhabitats, or that require different search and harvest strategies.

At the ESS for each of the 16 foods, we see a remarkable range of optimal adjustment of gut volume, u_1, and throughput time, u_2. We can draw several interesting conclusions from Table 1. First, the two most unfavorable food types will not sustain the forager. Despite massive increase in gut volume, and reduction of throughput time, the forager cannot live on foods 15 and 16. Second, gut volumes,

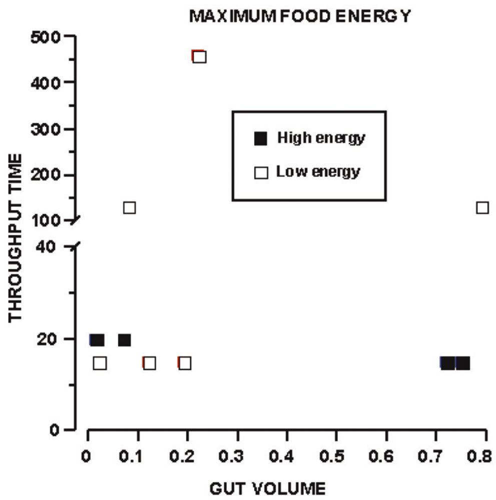


Figure 2: Effect of energetic value of foods (E_{max}) on joint adjustment of gut volume and throughput time. Food types differ in energetic value by a factor of 10. High E_{max} (filled squares) leads to fast throughput times but no consistent effect on gut volume. Low E_{max} (open squares) has no consistent effect on either throughput time or gut volume.

throughput times, and minimal levels of resource vary over a full order of magnitude. Finally, on the 14 foods that will support the forager, we see resource-limited foragers operating on mostly empty guts and generally far from full efficiency.

We can explore the results of the numerical analysis further graphically. In Figures 2–5 we show each of the 14 foods that will sustain the forager within the state space of gut volume and throughput time. In Figures 2–4, the foods are color-coded by a particular food property (e.g., high or low E_{max}, high or low absorption, etc.) to show how that property affects the digestive system.

The high level of energetic value of food, E_{max}, appears to select or lead to fast throughput times, but it has no consistent effect on gut volume (Figure 2). In contrast, at the low level of E_{max}, there was no consistent effect on either gut volume or on throughput time. Food bulk, b, had opposite effects, in that the food types with low bulk (the more favorable property) had no consistent effects on gut volume and throughput times, but foods with high bulk (the less favorable property) consistently led to fast throughput times (Figure 3).

When we combined energetic value and food bulk into richness (E_{max}/b), we find that the richest foods all led to fast throughput times but variable gut volumes, while foods with intermediate and low richness levels had no consistent effects on either gut volume or throughput time (Figure 4).

Foods with high absorption rates (or high ease of digestion), α, tended to lead to fast throughput times (with one exception), but had no consistent effect on gut volume. The exception perhaps traded off fast throughput time with gut volume: this

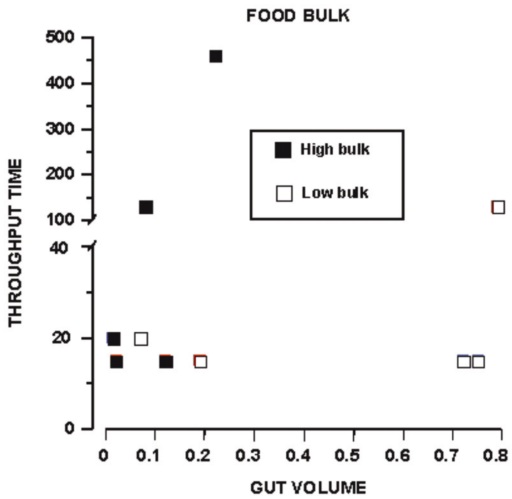


Figure 3: Effect of food bulk (b) on joint adjustment of gut volume and throughput time. Food types differ in food bulk by a factor of 10. Low food bulk (open squares) tends to lead to fast throughput times (with one exception—see text for details), but has no consistent effect on gut volume. High food bulk (filled squares) has no consistent effect on either throughput time or gut volume.

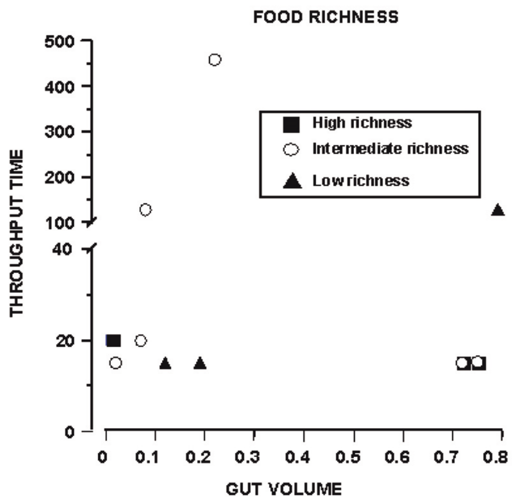


Figure 4: Effect of food richness (E_{max}/b) on joint adjustment of gut volume and throughput time. Foods differ in food richness over two orders of magnitude. Both high (filled triangles) and intermediate food richness (open circles) leads to fast throughput times but has no consistent effect on gut volume. Low food richness (filled diamonds) has no consistent effect on either throughput time or on gut volume.

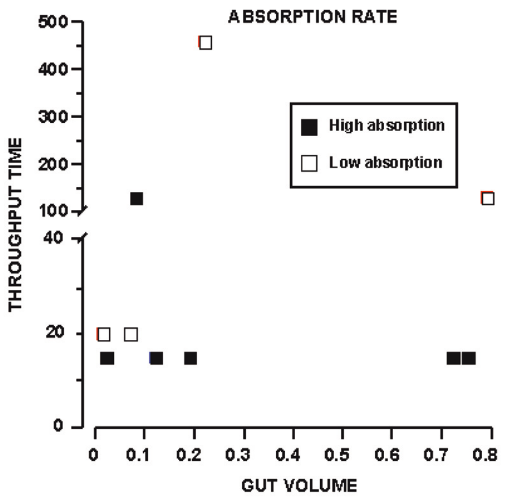


Figure 5: Effect of food absorption rate or ease of digestion (α) on joint adjustment of gut volume and throughput time. Foods differ in rate of absorption by a factor of 10. High rate of absorption (filled squares) leads to fast throughput time, but has no consistent effect on gut volume. Low rate of absorption (open squares) has no consistent effect on either throughput time or on rate of absorption.

food had an intermediate throughput time coupled with very small gut volume (Figure 5). Foods with low absorption rates had no consistent effects on gut volume or throughput time. Foods with long external handling times, h, all led to fast throughput times (Figure 6), while those with short external handling time had no consistent effect on gut volume or throughput time.

When considering all foods together, six natural groupings or syndromes of digestive system function are evident (Figure 7). Interestingly, the digestive system syndrome resulting from the largest number of food types has fast throughput times and small gut volumes. The five foods that lead to this syndrome possess high $E_{\max}$ and low absorption rate. Close by in the state space is a syndrome driven by two foods with high $E_{\max}$ but high absorption rates. This syndrome has fast throughput but considerably larger gut volume. The next most common syndrome consists of fast throughput times but large gut volumes. The four foods leading to this syndrome possess high $E_{\max}$ and high absorption rate. Each of the other three syndromes results from a single food type. Each of these foods has low $E_{\max}$, and either high bulk, low absorption rate, and/or high external handling (all unfavorable food properties). The main point to be drawn from these syndromes is that digestive system modulation in response to qualitatively different diets can lead to substantial differences in gut processing capabilities, and hence, substantial differences in diet specialization.

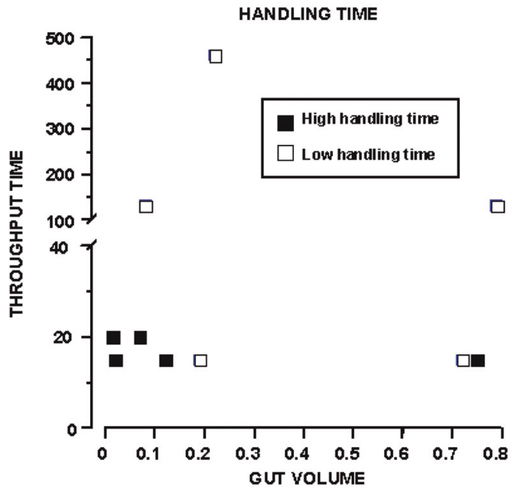


Figure 6: Effect of external food handling (h) on joint adjustment of gut volume and throughput time. Foods differ in external food handling by a factor of 10. High external food handling time (filled squares) leads to fast throughput times, but has no consistent effect on gut volume. Low external food handling time (open squares) has no consistent effect on either throughput time or on gut volume.

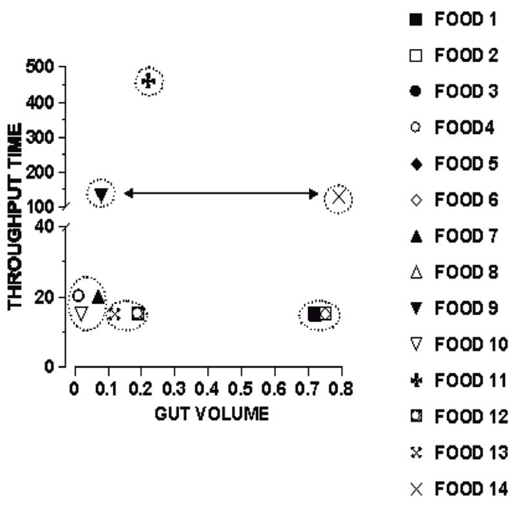


Figure 7: Apparent digestive physiological syndromes resulting from joint adjustment of gut volume and throughput time in response to different food properties. Syndromes are indicated with dotted lines. See text for details.

5 Implications for Speciation and Prospectus for Future Modeling

One of the great controversies of evolutionary biology is sympatric speciation, a process that takes place without geographic separation of the members of the species undergoing differentiation and ultimate reproductive isolation. Models of sympatric speciation require some sort of mechanism that results in disruptive selection on one or more traits [10, 12, 17]. The digestive system modulation and specialization indicated in Figure 7 suggest just such a mechanism. Modulation of digestive physiology to a particular food type causes different food types to become antagonistic resources, in which consumption of both resource types decreases reward or fitness relative to consuming only the food for which the digestive system has modulated to more efficiently process [20]. If some members of a population have modulated to take, say, food type 1, while others have modulated to take food type 2, particularly if the food types occur in different microhabitats or require different search strategies to effectively exploit them, then the stage may indeed be set for disruptive or divergent selection. The diet specialization favored by digestive modulation would tend to reinforce additional disruptive or divergent selection that may result from habitat segregation or behavioral modification related to search and exploitation strategies. This scenario is modeled by Mitchell [12], who found that habitat specialization, stable minima, community invasibility, and sympatric speciation are more likely when individuals are more efficient at converting resources into viable offspring. Gut modulation is obviously a mechanism that enhances resource-use efficiency.

Several purported examples of sympatric specialization involve changes in resource use that are wholly consistent with such a scenario. The apple maggot fly, *Rhagoletis pomonella*, is one example. This fruit fly underwent an apparent host shift from native hawthorn (*Crataegus* spp.) to introduced apple (*Malus pumila*). Other possible examples include *Enchenopa* spp. tree hoppers [22], Lake Victoria (East Africa) haplochromine cichlid fishes [16], *Littorina* spp. periwinkles [5], threespine stickleback (*Gasterosteus aculeatus*) [13], and black-bellied seedcrackers (*Pyrenestes ostrinus*) [15], among others.

The next step is to expand the current model into one in which the forager is confronted with two or more separate resources, a scenario approximating that developed in Brown [1]. In this scenario, the number of control variables for the forager increases to four. These, of course, include the two in this paper: gut volume, u_1, and throughput time, u_2. In addition, the forager, upon encountering an item of the first food type, will have some probability of accepting or rejecting it, u_3. The same applies to encountering an item of the second food type: there will be some probability of accepting or rejecting it, u_4. Foragers may now be behaviorally selective or opportunistic, and via modulation of digestive physiological function, they will possess trade-offs between performance when consuming the two foods. We suspect that one of three outcomes may emerge as the ESS: (1) a single generalist species that forages on the two foods opportunistically; (2) two species that are extreme specialists on either food type 1 or food type 2, respectively; (3) one generalist species that opportunistically forages on both food types and a specialist species that forages selectively on its preferred food type.

Appendix

Table 1: Sixteen foods that differ in energetic value, $E_{\max}$, bulk, b, rate of absorption, α, and external handling time, h, and the optimal gut volume, u_1*, throughput time, u_2*, and minimal level of resource density, $y*$, that result. Effective rate of absorption ("E") and gut fullness are also shown when the forager has reached optimal adjustment of gut volume and throughput time at $y*$. Foods are roughly arrayed from most favorable (top) to least favorable (bottom). Note that two foods (15 and 16) will not sustain the forager. Note also that for each of the first 14 foods, the forager is food limited, but typically operates far from peak rate of absorption (effective $E < 100\%$) with a mostly empty gut.

Food	$E_{\max}$	b	Richness (E/b)	α	h	u_1* (ml)	u_2* (time)	$y*$	"E" (%)	Fullness (%)
1	10	.1	100	.1	1	0.72	15	.003	60	0.063
2	10	.1	100	.1	10	0.75	15	.003	60	0.059
3	10	.1	100	.01	1	0.014	20	.006	16.7	8.5
4	10	.1	100	.01	10	0.014	20	.006	16.7	8.5
5	10	1	10	.1	1	0.72	15	.003	60	0.63
6	10	1	10	.1	10	0.75	15	.003	60	0.59
7	10	1	10	.01	1	0.07	20	.007	16.7	19.2
8	10	1	10	.01	10	0.07	20	.007	16.7	19.1
9	1	.1	10	.1	1	0.08	130	.012	92.8	18.6
10	1	.1	10	.1	10	0.02	15	.017	60	12.3
11	1	.1	10	.01	1	0.22	460	.017	82.1	32.2
12	1	1	1	.1	1	0.19	15	.02	60	15.4
13	1	1	1	.1	10	0.12	15	.02	60	23.2
14	1	1	1	.01	1	0.79	130	.044	56.5	52.3
15	1	.1	10	.01	10	23.0	0	—	—	—
16	1	1	1	.01	10	27.0	0	—	—	—

References

[1] J. S. Brown (1990), Habitat selection as an evolutionary game. *Evolution* **44**, 732–746.

[2] R. Cressman (2004), Coevolution, adaptive dynamics, and the replicator equation for a single species with a continuous trait space. Proceedings, International Society of Dynamic Games, Tucson, Arizona, USA.

[3] J. Garay and L. Eötvös (2004), Adaptive phenotypic change based on ecological stability. Proceedings, International Society of Dynamic Games, Tucson, Arizona, USA.

[4] C. S. Holling (1965), The functional response of predators to prey density and its role in mimicry and population regulation. *Memoirs of the Entomological Society of Canada* **45,** 1–60.

[5] K. Johannesson (2003), Evolution in *Littorina*: ecology matters. *Journal of Sea Research* **49**, 107–117.

[6] P. A. Jumars and C. Martínez del Rio (1999), The tau of continuous feeding on simple foods. *Physiological and Biochemical Zoology* **72**, 633–641.

[7] W. H. Karasov (1996), Digestive plasticity in avian energetics and feeding ecology. Pages 61–84 *in* C. Cary, editor, *Avian Energetics.* Chapman & Hall, New York, New York, USA.

[8] W. H. Karasov and I. D. Hume (1997), The vertebrate gastrointestinal system. Pages 407–480 *in* W. H. Dantzler, editor, *Handbook of Physiology*, Section 13: Comparative Physiology, Vol. 1. Oxford University Press, Oxford, New York, New York, USA.

[9] C. Martínez del Rio, S. Cork, and W. H. Karasov (1994), Engineering and digestion: modeling gut function. Pages 25–53 *in* D. Chivers and P. Langer, editors, *Food and Form and Function of the Mammalian Digestive Tract.* Cambridge University Press, Cambridge, United Kingdom.

[10] J. Maynard Smith (1966), Sympatric speciation. *American Naturalist* **100**, 637–650.

[11] T. J. McWhorters and C. Martínez del Rio (2000), Does gut function limit hummingbird food intake? *Physiological and Biochemical Zoology* **73**, 313–324.

[12] W. A. Mitchell (2006), Adaptive dynamics, resource conversion efficiency, and species diversity. Proceedings, International Society of Dynamic Games, Tucson, Arizona, USA.

[13] H. R. Rundle, L. Nagel, J. Wenrick Boughman, and D. Schluter (2000), Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**, 306–308.

[14] T. B. Smith (1990), Comparative breeding biology of the two bill morphs of the black-bellied seedcracker. *The Auk* **107**, 153–160.

[15] J. M. Starck (2003), Shaping up: how vertebrates adjust their digestive system to changing environmental conditions. *Animal Biology* **53**, 247–257.

[16] E. Verheyen, W. Salzburger, J. Snoeks, and A. Meyer (2003), Origin of the super flock of cichlid fishes from Lake Victoria, East Africa. *Science* **300**, 325–329.

[17] S. Via (2001), Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology and Evolution* **16**, 381–390.

[18] T. L. Vincent and J. S. Brown (2005), *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics*. University of Cambridge Press, Cambridge, United Kingdom.

[19] C. J. Whelan and J. S. Brown (2005), Optimal foraging and gut constraints: reconciling two schools of thought. *Oikos* **110**, 481–496.

[20] C. J. Whelan, J. S. Brown, K. A. Schmidt, B. B. Steele, and M. F. Willson (2000), Linking consumer-resource theory and digestive physiology: application to seasonal diet shifts. *Evolutionary Ecology Research* **2,** 911–934.

[21] C. J. Whelan and K. A. Schmidt (2006), Food acquisition and processing. *In* Stephens, D. W., Ydenberg, R. and Brown, J. S., editors, *Foraging*. University of Chicago Press, Chicago, IL, USA.

[22] T. K. Wood, K. J. Tilmon, A. B. Shantz, C. K. Harris, and J. Pesek (1999), The role of host-plant fidelity in initiating insect race formation. *Evolutionary Ecology Research* **1**, 317–332.

PART V
Applications of Dynamic Games to Economics

Time-Consistent Fair Water Sharing Agreements

Rodney Beard
School of Economics and the ARC Centre for Complex Systems
The University of Queensland
St. Lucia 4072, Queensland, Australia
beard.rodney@yahoo.com.au

Stuart McDonald
School of Economics and the ARC Centre for Complex Systems
The University of Queensland
St. Lucia 4072, Queensland, Australia

Abstract

Scarcity of water has become a major issue facing many nations around the world. To improve the efficiency of water usage there has been considerable interest in recent years in trading water. A major issue in trading water rights is the problem of how an allocation system can be designed in perpetuity that also has desirable properties at each point of time. This is an issue of the time consistency of the contract to trade water. In this chapter we develop a model of dynamic recontracting of water rights and study time consistency properties of the resultant contracts using the ideas of Filar and Petrosjan [7].

Key words. Dynamic cooperative games, water sharing agreements, time consistency of contracts.

1 Introduction

Scarcity of water has become a major issue facing many nations around the world. To improve the efficiency of water usage there has been considerable interest in recent years in using market-based instruments to trade water. The introduction of markets for water was first suggested in the literature by Burness and Quirk [2],[3]. Since then the literature has expanded to include a number of studies including Howe, Schurmeier, and Shaw [12], Provencher [19], and Provencher and Burt [20]. The analysis of dynamic aspects of water trading is still relatively new. The main contributions so far are those of Lahmandi-Ayed and Matoussi [13], Freebairn and Quiggin [9], and Firoozi and Merrifield [8].

A major issue in trading water rights is the problem of how an allocation system can be designed in perpetuity that also has desirable properties at each point of time. Water markets basically consist of two types of contracts: perpetual trades and temporary trades. Reconciling these two contract types is an issue of the time consistency of the contract to trade water. Lahmandi-Ayed and Matoussi have examined the impact of water markets on investment in water-saving technology using a theoretical model. They find that water markets have an unclear impact on investment in new technology. One interpretation of their result is that of time inconsistency in water trading and this in turn leads to the focus of this chapter in identifying time-consistency conditions for contracts for water rights.

Freebairn and Quiggin are concerned with the impact of highly variable rainfall on markets for water and develop an analysis of water trading using contingent contracts in a state-contingent asset pricing model. The concern with highly variable rainfall also lies at the center of this chapter. Unlike Freebairn and Quiggin we attempt to reconcile perpetual contracts for water with temporary trading. Freebairn and Quiggin recognize the difference in the nature of these two contracts and attempt to address this by allowing short-selling as a means to capture temporary trades. Nevertheless, they ignore the issue of the compatibility of permanent and temporary contracts which is the focus of our paper.

Time consistency of cooperative games has been discussed extensively in the dynamic games literature; important contributions to the literature are those of Petrosjan [16], Filar and Petrosjan [7], and Zakharov and Dementieva [22]. However, there have only been a few studies that utilize these ideas in an applied context; for example, Petrosjan and Zaccour [17] and Tarashnina [21]. Time consistency of the solution is a requirement that needs to be imposed on contracts for permanent markets in water to be of practical use. We apply, from Filar and Petrosjan [7], the dynamic extension of cooperative games in characteristic function form to the problem of recontracting water rights in the presence of fluctuations in rainfall. The model extends the ideas of Ambec and Sprumont [1], who model the sharing of water resources along a river to a dynamic cooperative game setting. Using the ideas of Filar and Petrosjan [7] we develop a model of dynamic recontracting of water rights and study time-consistency properties of the resultant contracts. We study the evolution of the characteristic function over time when the solution concept employed is that of *downstream incremental distribution*. The downstream incremental distribution is a compromise solution concept due to Ambec and Sprumont [1] that captures the interests of both upstream and downstream users of a river. Implications for contract and market design for water markets are then discussed.

The remainder of the chapter is organised as follows. Section 2 presents the model. Section 3 discusses dynamics, Section 4 the time consistency of water contracts, Section 5 the coalitional τ-value and bargaining for water, and Section 6 drop-out monotonicity. Section 7 presents a numerical example of the computation of the characteristic function over time using an artificial data set. Conclusions are drawn in Section 8.

2 The Model

The ordering of players along the river may be modelled in a number of ways. One way would be via a graph restricted game along the lines of Myerson [15].

Another approach is to view the ordering along the river as restricting coalitions to a priori unions along the lines of Casas-Mendez et al. [4]. In the specific case here it is modelled as a Greenberg and Weber [10] consecutive game.

We begin by following Ambec and Sprumont [1] and consider a river that flows through a number of jurisdictions, where the set of all jurisdictions corresponds to the set of agents $N = \{1, \dots, n\}$. Agents are linearly ordered along a river so that for any two agents $i, j \in N$, agent $i < j$ implies that i is upstream from j. For any two coalitions $S, T \subset N$, if for all $i \in S$ and $j \in T$, $i < j$, then $S < T$. The first member of coalition S is denoted by $\min S$ and the last member by $\max S$. The set of predecessors of agent i is defined as $P_i = \{j \in N : j \leq i\}$, and the set of strict predecessors of agent i is given by $P_i^0 = P_i \setminus \{i\}$.

As the river flows downstream the volume of water flowing through the river increases via inflows from tributaries and run-off from surrounding land. The flow of water at the source is given by $e_1 > 0$. Because land is assumed to be privately owned and divided up amongst agents, this implies that the inflow of water from surrounding land $e_i > 0$ between property $i - 1$ and i can always be attributed to the ith agent. Each agent's utility is assumed to be of the form

$$U_i(x_i, t_i) = b_i(x_i) + t_i, \quad i \in N, \tag{1}$$

where $x_i \in \mathbb{R}_+$ is the ith agents consumption of water, the function $b_i : \mathbb{R}_+ \to \mathbb{R}$ measures the benefit that the ith agent derives from water consumption, and t_i is a net transfer of money to the ith agent, with $t_i < 0$ corresponding to a payment and $t_i > 0$ corresponding to the receipt of money.

Here the set of agents N characterises the ordering of jurisdictions, the n-tuple $e = (e_1, \dots, e_n)$ contains all information about the flow of water to the river from various sources, and the n-tuple $b = (b_1, \dots, b_n)$ provides information about benefits each of the agents derives from water consumption. Hence all information about the river and its flows and the placement of agents along the river and their activities can be expressed completely in terms of the triple (N, e, b).

The consumption plan for the grand coalition N is any n-tuple $x = (x_1, \dots, x_n) \in \mathbb{R}_+^n$. For any nonempty coalition $S \subset N$, $x_S \in \mathbb{R}_+^{|S|}$ denotes the consumption plan for the members of coalition S. An allocation is given by $(x, t) = (x_1, \dots, x_n, t_1, \dots, t_n) \in \mathbb{R}_+^n \times \mathbb{R}^n$ of consumption plan-transfer combinations. An allocation (x, t) is said to be feasible if it satisfies the following constraints:

$$\sum_{i \in N} t_i \leq 0 \tag{2}$$

and

$$\sum_{i \in Pj} (x_i - e_i) \leq 0, \quad \forall j \in N. \tag{3}$$

The first constraint assures that net transfers sum to zero. The second constraint guarantees that consumption is feasible at each point along the river; it is expressed in terms of the successor set of each agent $j \in N$, because the water stream e_i

can only be consumed by the followers of j. The optimal feasible consumption plan $(x^*(N), t^*(N))$ can be found by maximising $\sum_{i\in N} U_i(x_i, t_i)$ subject to the feasibility constraints (2) and (3). The maximal welfare of society, which is to be defined among all of the agents, is then measured by $\sum_{i\in N} b_i(x_i^*(N))$.

Following Greenberg and Weber [10] we now define a consecutive game Γ. First note that a coalition $T \subset N$ is called a consecutive if $k \in T$ whenever $i, j \in T$ and $i < k < j$. For an arbitrary coalition $S \subset N$, let $\mathcal{S}$ denote the coarsest partition of S into its consecutive components. Now consider the unique consumption plan for the coalition S that maximises $\sum_{i\in S} b_i(x_i)$ subject to the constraint

$$\sum_{i\in P_j\cap T} (x_i - e_i) \le 0, \quad \forall j \in T \text{ and } T \in \mathcal{S}$$

and let the secure benefit of S be defined by

$$v(S) = \sum_{i\in S} b_i(x_i^*(S)). \tag{4}$$

Note that $v(S) = \sum_{T\in\mathcal{S}} v(T)$, which implies that any water left over by one of the consecutive components of S cannot be guaranteed for consumption by any other consecutive component of S. The consecutive game Γ is defined in terms of the secure benefit v.

The aspiration welfare of an arbitrary coalition S is the highest welfare that the members of S could achieve in the absence of $N\backslash S$. It is obtained by choosing a consumption plan $x^{**}(S)$ that maximises $\sum_{i\in S} b_i(x_i)$ subject to

$$\sum_{i\in P_j\cap S} x_i \le \sum_{i\in P_j} e_i, \quad \forall\, j \in S.$$

The aspiration welfare of coalition S is then expressed by

$$w(S) = \sum_{i\in S} b_i(x_i^{**}(S)).$$

In terms of the individual, in the absence of all other agents, the aspiration welfare of the ith agent is expressed by $w(i) = b_i(\sum_{j\in P_i} e_j)$, with the aspiration welfare distribution given by the n-tuple $(w(1), \dots, w(n))$. Note that for $n \ge 2$, $(w(1), \dots, w(n))$ is not feasible.

Associated with each feasible allocation (x, t) is a welfare distribution given by the n-tuple $(z_1, \dots, z_n) \in \mathbb{R}^n$, where each of the z_i is the utility image $b_i(x_i) + t_i$ for the ith agent. The welfare distribution $(z_1, \dots, z_n)$ for an arbitrary feasible allocation (x, t) is said to satisfy the core lower bound for the consecutive game v if

$$\sum_{i\in S} z_i \ge v(S), \quad \forall\, S \subset N.$$

The welfare distribution $(z_1, \ldots, z_n)$ associated with the feasible allocation (x, t) is said to satisfy the aspiration upper bound if

$$\sum_{i \in S} z_i \leq w(S), \quad \forall\, S \subset N.$$

Note that this implies that the aspiration welfare for any two complementary coalitions S and $N \backslash S$ is incompatible, i.e., $w(S) + w(N \backslash S) > v(N)$.

An immediate consequence of this is that for every agent $i \in N$, $v(P_i) = w(P_i)$ for the predecessor set P_i. This implies that for all agents $i \in N$, the core lower bound and the aspiration upper bound of P_i are equal. This leads to the concept of downstream incremental distribution $z^* = (z_1^*, \ldots, z_n^*) \in \mathbb{R}^n$, which is used by Ambec and Sprumont [1] to capture the marginal value of each agent's contribution to the coalition of upstream agents. More formally, the downstream incremental distribution z_i^* for each agent $i \in N$ can be defined as follows:

$$z_i^* = v(P_i) - v(P_i^0) \text{ or } z_i^* = w(P_i) - w(P_i^0);$$

both are equivalent.

The downstream incremental distribution is an example of a compromise solution concept. Ambec and Sprumont show that the downstream incremental distribution is the unique distribution satisfying the core lower bounds and the aspiration upper bounds. They note that it represents a compromise between the legal principles of absolute territorial integrity and unlimited territorial integrity. As Ambec and Sprumont have shown, the most important property of the downstream incremental distribution z^* is that it is the only distribution that satisfies both core lower bound and the aspiration upper bounds of the consecutive game, i.e., for any coalition $S \subset N$

$$v(S) \leq \sum_{i \in S} z_i^* \leq w(S).$$

3 Discrete Time Dynamics

We now extend the consecutive game considered by Ambec and Sprumont and embed it within a dynamic game along the lines of Filar and Petrosjan [7]. The way to think about this is that there are two types of dynamics occurring here. The first involves the flow of water downstream and is modelled as a consecutive game. The second involves the recharging of the river through periodic rainfall events leading to the beginning of a new consecutive game at each stage.

We consider a sequence of consecutive games Γ_k, where $k = 0, 1, \ldots, K$. Associated with each of the consecutive games Γ_k and its secure benefit v_k is a solution concept $C(v_k)$. This solution concept can be any solution concept from static cooperative game theory, e.g., the Shapley value or the von Neumann–Morgenstern solution. Following Filar and Petrosjan [7], the game dynamics are given by

$$v_{k+1} = f(v_k, \alpha_k), \; k = 0, \ldots, K-1 \qquad (5)$$
$$v_0 = v,$$

where $\alpha_k \in C(v_k)$ is an imputation distribution procedure and v is the characteristic function of the Ambec and Sprumont consecutive game. Note that equation (5) introduces a K-stage imputation sequence $\alpha = \{\alpha_k\}_{k=1}^{K-1}$ which determines the trajectory of $\{v_k\}_{k=0}^{K}$.

The dynamics of the characteristic function for the downstream incremental distribution z_k associated with each stage game $v_k, k = 1, \ldots, K$ are given by the following recursive scheme:

$$v_{k+1} = v_k + c(v_k, z_k) \sum_{i \in N} z^*_{i\,k}, \qquad (6)$$

or if the transfer function $c_i(v_k, z^*_{i\,k})$ depends on i, then

$$v_{k+1} = v_k + \sum_{i \in N} c_i(v_k, z^*_{i\,k}) z^*_{i\,k} \qquad (7)$$

and

$$c_i(v_k, z^*_{i\,k}) = \begin{cases} -t_i^*, & \text{if } \dfrac{|S|}{N} \Delta(k) > \displaystyle\sum_{i \in S} z^*_{ik} \\ t_i^*, & \text{if } \dfrac{|S|}{N} \Delta(k) \leq \displaystyle\sum_{i \in S} z^*_{ik} \end{cases} \qquad (8)$$

with $v = \sum_{i \in S} b_i(x_i^*(S))$ and $z_i^* = v(P_i) - v(P_i^0)$, where $x_i^*(S)$ is obtained by maximising $\sum_{i \in S} b_i(x_i)$ subject to the constraint

$$\sum_{i \in P_j \cap T} (x_i - e_i) \leq 0, \quad \forall j \in T \text{ and } T \in \mathcal{S},$$

and $\Delta(k) = v_k(N) - \sum_{i \in N} v_k(\{i\})$. This system may be solved numerically by first solving the intertemporal programming problem (equations (2) and (3)) to find the sequence of transfers between agents. This is done in Section 7.

4 Time Consistency of Water Contracts

An important question in negotiating water contracts is how to allocate these contracts in perpetuity. This is essentially a question of how to design time-consistent contracts. Following Filar and Petrosjan [7], we consider a sequence of consecutive games Γ_k and their associated secure benefits $v_k, k = 1, \ldots, K$. The aggregate cooperative game G that is based on this sequence of games can be defined by the following secure benefit:

$$\bar{v}(S) = \sum_{k=0}^{K} v_k(S), \quad \forall\, S \subset N,$$

and we let the aspiration welfare function for G be given by

$$\bar{w}(S) = \sum_{k=0}^{K} w_k(S), \quad \forall\, S \subset N.$$

It is possible to construct the summed solution set $\bar{C}$ for the aggregate game G by summing across the solution sets $C(v_k)$ for every stage game Γ_k,

$$\bar{C} = C(v_0) \oplus C(v_1) \oplus \cdots \oplus C(v_K).$$

Here $\oplus$ is the set addition operator, i.e., for any two sets A and B,

$$A \oplus B = \{a + b;\, a \in A \text{ and b} \in B\}.$$

Let z_k^* be the downstream incremental distribution of each stage game Γ_k, $k = 0, \ldots, K$, and $\bar{z}^* = \sum_{k=0}^{K} z_k^*$ be the summed downstream incremental distribution associated with G. This leads to the following result.

Lemma 4.1. *Suppose that* $\bar{z}^* = (\bar{z}^*_{1\,k}, \ldots, \bar{z}^*_{n\,k}) \in \mathbb{R}^n$ *is the downstream incremental distribution in G, then*

$$\bar{z}^* = \sum_{k=0}^{K} z_k^*,$$

where $z_k^* = (z^*_{1\,k}, \ldots, z^*_{n\,k})$ *is the downstream incremental distribution in* Γ_k*,* $k = 0, \ldots, K$.

Proof. Ambec and Sprumont have shown that for each consecutive stage game Γ_k, the downstream incremental distribution is the only distribution that satisfies

$$\sum_{i \in S} z^*_{i\,k} \geq v_k(S), \quad S \subset N$$

and

$$\sum_{i \in S} z^*_{i\,k} \leq w_k(S), \quad S \subset N.$$

We can see that when we sum across stage games

$$\sum_{k=0}^{K} \sum_{i \in S} z^*_{i\,k} = \sum_{i \in S} \sum_{k=0}^{K} z^*_{i\,k} \geq \sum_{k=0}^{K} v_k(S) = \bar{v}(S), \quad S \subset N$$

and

$$\sum_{k=0}^{K}\sum_{i\in S} z^*_{i\,k} = \sum_{i\in S}\sum_{k=0}^{K} z^*_{i\,k} \leq \sum_{k=0}^{K} w_k(S) = \bar{w}(S), \quad S \subset N.$$

This implies that

$$\sum_{i\in S} \bar{z}^*_i \geq v(S), \quad S \subset N$$

and

$$\sum_{i\in S} \bar{z}^*_i \leq w(S), \quad S \subset N,$$

where for any $i \in N$

$$\begin{aligned}\bar{z}^*_i &= \bar{v}(P_i) - \bar{v}(P^0_i)\\ &= \sum_{k=0}^{K}(v_k(P_i) - v_k(P^0_i))\\ &= \sum_{j\in P_i}\sum_{k=0}^{K}(b_j(x^*_j(P_i)) - b_j(x^*_j(P^0_i)))\end{aligned}$$

and

$$\begin{aligned}\bar{z}^*_i &= \bar{w}(P_i) - \bar{w}(P^0_i)\\ &= \sum_{k=0}^{K}(w_k(P_i) - w_k(P^0_i))\\ &= \sum_{j\in P_i}\sum_{k=0}^{K}(b_j(x^{**}_j(P_i)) - b_j(x^{**}_i(P^0_i)))\end{aligned}$$

□

We now introduce a definition of time consistency for the aggregate game G. For a fixed sequence $\bar{\alpha} = \{\bar{\alpha}_0, \ldots, \bar{\alpha}_K\}$, where $\bar{\alpha}_k \in C(\bar{v}_k)$, $k = 0, \ldots, K$. Let $v_0, v_1, \ldots, v_m$ be the secure benefits for each stage game Γ_k as defined by the recursive dynamics (5). For each sub-game G_k of G starting from stage k with the initial stage game $\bar{\Gamma}_k$ with secure benefit v_k and aspiration welfare w_k, the secure benefit of G_k is defined by

$$\bar{v}^k(S) = \sum_{l=k}^{K} v_l(S),$$

and the aspiration welfare of G_k is defined by

$$\bar{w}^k(S) = \sum_{l=k}^{K} w_l(S).$$

A solution concept is time consistent if for each $\alpha \in \bar{C}$ there exists $\beta = (\beta_0, \ldots, \beta_m)$, $\beta_k \geq 0$, $k = 0, 1, \ldots, m$, such that $\alpha = \sum_{k=0}^{m} \beta_j$ and

$$\alpha^k = \sum_{j=k}^{m} \beta_j \in \bar{C}^k = \sum_{j=k}^{K} \oplus C(v_j).$$

As pointed out by Filar and Petrosjan this definition of time consistency can have problems in the case of nonunique solution concepts [7, p. 54]. However, both downstream incremental distribution and the coalitional τ-value are unique so we do not need to consider the question of internal time consistency. It will therefore suffice to show that the downstream incremental distribution is time consistent. This result is now provided below.

Proposition 4.1. *The downstream incremental distribution is time consistent.*

Proof. Let $\bar{z}^* = (\bar{z}_1^*, \ldots, \bar{z}_n^*)$ be the downstream incremental distribution for the summed game G. This implies that $\bar{z}^* = \sum_{k=1}^{K} z_k^*$, where $z_k^* = (z_{1k}^*, \ldots, z_{nk}^*)$ is the downstream incremental for the consecutive stage game Γ_k. Let $\beta_k = z_k^*$; then for $\beta_k \geq 0$ and

$$\bar{z}_k^* = \sum_{j=k}^{m} \beta_j = \sum_{j=k}^{K} z_k^* \in \sum_{j=k}^{K} \oplus C(v_j) = \bar{C}^k,$$

which indicates that $\bar{z}^*$ is time consistent. □

This result implies that if downstream incremental distribution were to be implemented as a water rights allocation system, then this compromise solution will be consistent through time. The recharge of the water supply through rainfall events would not require a reallocation of water rights as long as users of water are optimizing their use at each point in time. Furthermore, it provides us with a way of reducing the computational burden that would be required in order to calculate the size of the surplus needed for redistribution. This is because as long as the path of v_k is optimal, which will be the case as long as the transfers are given by an optimal sequence and as long as the initial v_0 is optimal, we need not solve a mathematical programming problem at each time step.

5 The Coalitional τ-value and Bargaining for Water

Another possible candidate solution concept is the coalitional τ-value [4]. This is an extension of the τ-value to games with coalition structures. We now turn to the question of time consistency of the coalitional τ-value. Following Casas-Mendez et al. [4] we now introduce a game with a priori unions or coalition structures. Such a game $\Gamma_S = (N, v, P)$ where $P = \{P_1, \ldots, P_m\}$ is a partition of the player set.

A partition of the players may be defined that corresponds to the linear river of Ambec and Sprumont, so that P_i, P_i^0 defines a partition. Consequently, the consecutive game of Ambec and Sprumont may be formulated as a game with a priori unions or coalition structures. It is also straightforward to write P in terms of predecessor and strict predecessor sets. It is straightforward to extend the model by defining other partitions to branching river networks. So this setting appears to have no disadvantage over that employed by Ambec and Sprumont.

We now introduce the quotient game $v^P \in G(M)$, where $v^P = v(\cup_{k \in L} P_k)$ for all $L \subset M$ and $G(M)$ is the set of unions defined to be $M := \{1, \dots, m\}$ the index set of the partitions P_k, i.e., each $k \in M$ is the set of games defined on unions in M. We define the utopia payoff for each agent $i \in N$ by

$$M_i(v, P) := v(N) - v(N \setminus i).$$

Note that the utopia payoff of the ith player would correspond to the downstream incremental distribution of the nth player. Casas-Mendez also note that the utopia payoff can be expressed as follows:

$$M_i(v, P) = v^P(M) - v^P(M \setminus \{k\}) - v^P_{-i}(M) - v^P_{-i}(M \setminus \{k\}), \quad i \in N.$$

Each player has his minimal rights defined in terms of the following equation:

$$m_i(v, P) := \max_{S \in P(k): i \in S} \left\{ v(S) - \sum_{j \in S \setminus \{i\}} M_j(v, P) \right\}.$$

We now provide the following definition of a quasi-balanced game. A game $v \in G(N)$ is said to be quasi-balanced if it satisfies the following two properties:

(1) $m(v) \leq M(v)$ where $m(v)M$ and $(v) \in \mathcal{R}^N$ are vectors whose components are the minimal rights and utopia payoffs functions as defined above.
(2) $\sum_{i \in N} m_i(v) \leq v(N) \leq \sum_{i \in N} M_i(v)$.

The class of quasi-balanced games with player set N is denoted $QBG(N)$.

We can now introduce quasi-balanced games with a priori unions. A game with a priori unions $(v, P) \in U(N)$ is said to be quasi-balanced if and only if the following three conditions are satisfied:

(1) $v^P \in QBG(M)$,
(2) $m(v, P) \leq M(v, P)$, where these are vectors defined analogously to those above,
(3) $\sum_{i \in P_k} m_i(v, P) \leq \sum_{i \in P_k} \tau_i(v^P) \leq \sum_{i \in P_k} M_i(v, P)$ for all $P_k \in P$.

The class of quasi-balanced games with a priori unions and player set N is denoted $QBU(N)$.

We are now ready to introduce the coalitional τ-value. The coalitional τ-value is a map $\tau : QBU(N) \to R^N$ which assigns to every $(v, P) \in QBU(N)$ the vector $(\tau_i(v, P))_{i \in N}$ such that, for all $P_k \in P$ and all $i \in P_k$,

$$\tau_i(v, P) := m_i(v, P) + \alpha_{k'}(M_i(v, P) - m_i(v, P)),$$

where, for each $k \in M$, $\alpha_{k'}$ is such that $\sum_{i \in P_k} \tau_i(v, P) = \tau_k(v^P)$. The coalitional τ-value is a compromise between the upper and lower vectors defined by the utopia payoff and the minimal right of a player. Compare this to the aspiration upper bounds of Ambec and Sprumont. We now present the following result on the time consistency of the coalitional τ-value.

Proposition 5.1. *The coalitional τ-value is time consistent.*

Proof. Set $\beta_k = \tau^*_{ik}$. Now consider $\alpha = (\alpha_0, \ldots, \alpha_K) \in \bar{C}$ represented as $\alpha = \sum_{k=0}^{K} \alpha_k$, where $\alpha_k \in C(v_k)$. Linearity implies

$$\begin{aligned}\tau_i^* &= m_i(v, P) + \alpha_{k'}(M_i(v, P) - m_i(v, P)) \\ &= (1 - \alpha_{k'})(v(S^*) - \sum_{j \in S \setminus \{i\}} (v(N) - v(N \setminus \{j\}))) \\ &\quad + \alpha_{k'}(v(N) - v(N \setminus \{i\})).\end{aligned} \tag{9}$$

Because we know that $v(S) = \sum_{k=0}^{m} v_k(S)$ (Filar and Petrosjan [7, p. 51]), the right-hand side of τ_i^* may be expanded as follows:

$$\begin{aligned}\tau_i^* &= (1 - \alpha_{'k})(\sum_{k=0} v_k(S^*) - \sum_{k=0}^{m} \sum_{j \in S^* \setminus \{i\}} (v_k(N) - v_k(N \setminus \{i\}))) \\ &\quad + \alpha_{k'} \sum_{k=0}^{m} (v_k(N) - v_k(N \setminus \{i\})) \\ &= \sum_{k=0}^{m} [m_{ik}(v, P) + \alpha_{k'}(M_{ik}(v, P) - m_{ik}(v, P))] \\ &= \sum_{k=0}^{m} \tau_{ik}(v, P).\end{aligned} \tag{10}$$

Letting $\beta_k = \alpha_k$ we can see that $\alpha^k = \sum_{j=k}^{m} \beta_j = \sum_{j=k}^{m} \alpha_j \in \bar{C}^k$. □

By Theorem 3 of Casas-Mendez et al. [4, p. 501] the coalitional τ-value is unique so that we need not consider internal time consistency.

6 Drop-Out Monotonicity

It would seem that both downstream incremental distribution and the coalitional τ-value are time consistent and both possible candidates for perpetual contracts in water rights. How then can we choose between these two solution concepts? Hendrickx [11] has shown that the model of Ambec and Sprumont is related to sequencing games. Sequencing games were first introduced by Curiel, Pederzoli, and Tijs [5] to study the problem of sharing cost savings in queueing.

Hendrickx shows that downstream incremental distribution, which he refers to as the μ-rule, captures the concept of drop-out monotonicity, in sequencing games with regular cost functions. Drop-out monotonicity is a requirement that costs (or values) are not reduced (increased) when an agent drops out of the queue.

We now modify the model of Ambec and Sprumont to make the connection to sequencing games more explicit. A sequencing situation consists of a queue of n players whose positions in the queue are described by a permutation σ of the player set N. For example, $\sigma(i) = j$ denotes that player i holds position j in the queue. A given permutation of the player set corresponds to a particular ordering of players. The predecessor set $P(\sigma, i) := \{j \in N | \sigma(j) < \sigma(i)\}$ is equivalent to the predecessor set of Ambec and Sprumont for a given permutation. A given permutation of the player set corresponds to a particular ordering of players along the river. Each player has associated with him a utility function employed above.

We now introduce the following definition of a sequencing situation modified for water flow. A sequencing situation (for water flow) is an ordered triple (σ, b, x) consisting of where $\sigma \in \Pi_N$ is the set of permutations of the player set N, b is the utility of a player, and x is the water consumed by a player. The amount of water x consumed by an agent located along the river takes the place of the service time in the standard form of the sequencing game. The way to think about this is that both are resources consumed by agents in sequence.

Curiel, Pederzoli, and Tijs [5] define the τ-value of a sequencing situation to be the τ-value of the corresponding sequencing game. Note that a given permutation of the player set defines a particular P_k of the set of a priori unions of the game with a priori unions and that consequently the coalitional τ-value corresponds to the τ-value of a sequencing game for a given permutation of the player set. Sequencing games may be viewed as a special case of a game with a priori unions in which the partition of the player set possesses a sequential order.

The question is whether or not the τ-value is drop-out monotonic. We now formally define drop-out monotonicity following Hendrickx [11] as follows. An allocation rule is called drop-out monotonic if for all (σ, b, x) and all $q \in N$

$$\alpha_j((\sigma, b, x)) \leq \alpha_j((\sigma, b, x)^{-q}), \quad \alpha \in C$$

for all $j \in N\backslash\{q\}$, where $(\sigma, b, x)^{-q} = (N\backslash\{q\}, (x_i)_{i \in N\backslash\{q\}}, (b_i)_{i \in N\backslash\{q\}})$. Note that the inequality is reversed here compared with Hendrickx because we are dealing with utility functions, not cost functions. The following result is now provided.

Proposition 6.1. *The τ-value is not drop-out monotonic and stable.*

Proof. This is a corollary of the fact (shown by Hendrickx [11, p. 42 Theorem 4.4.2]) that downstream incremental distribution is the uniquely stable and drop-out monotonic allocation rule for sequencing games. □

One way to think about drop-out monotonicity is as a coalitional analogue of an individual rationality constraint. If it is satisfied, players would have an incentive to force player q not to participate in the contract. This could be considered anti-social behaviour. Consequently, there is a trade-off between the self-enforcement characteristics of the τ-value (individual rationality) and the fairness element of drop-out monotonicity. It would seem difficult to resolve this tension in the concept.

7 Numerical Solution of Characteristic Function for Downstream Incremental Distribution

In this section we consider the problem of numerically computing the evolution of the characteristic function for downstream incremental distribution over time. For the purpose of illustration we will use 9 periods. For a discrete time problem this can lead to the problem of the characteristic function taking on negative values. This problem has been addressed by the literature on regularization in multistage cooperative games [6]. However, as can be seen in what follows, we do not encounter this problem here.

We consider a river with five agents spread along the river. We will assume a quasi-linear utility function of the form $U(x, t) = \sqrt{x} + t$. Note that the inflows of water into the river are shown in Table 1. Each column represents a different point of inflow along the river and each row a different point in time. Location 5 is furthest upstream and hence has the least accumulated water.

These data were generated by simulation and scaled to allow for increased available water downstream. They are used simply to indicate the procedure employed and should not be considered as specific to any particular river.

We first need to compute the Pareto efficient allocation of water over time. This is done by solving the following nonlinear programming problem:

$$\max_{x_{ik}, t_{ik}} \sum_{k=0}^{9} \sum_{i \in N} \sqrt{x_{ik}} + t_{ik}$$

Table 1: Water flows e_{ik} at each point in time and at each location.

k	1	2	3	4	5
0	78.126432	54.688502	38.281951	26.797366	18.758156
1	80.268229	56.187760	39.331432	27.532002	19.272401
2	35.410701	24.787490	17.351243	12.145870	8.5021093
3	89.630385	62.741270	43.918889	30.743222	21.520255
4	81.151035	56.805724	39.764007	27.83480	19.484363
5	5.4502527	3.8151769	2.670623	1.8694367	1.3086056
6	48.065390	33.645773	23.55204	16.486428	11.540500
7	71.329912	49.930938	34.951656	24.466159	17.126311
8	52.199123	36.539386	25.577570	17.904299	12.533009
9	68.720569	48.104398	33.673078	23.571155	16.499808

subject to

$$\sum_{k=0}^{9}\sum_{i\in N} t_{ik} \leq 0$$

and

$$\sum_{k=0}^{9}\sum_{i\in P_j}(x_{ik} - e_{ik}) \leq 0, \quad j = 1, \dots, 5.$$

The Pareto efficient level of welfare is 282.2127. This welfare needs to be distributed in a fair way between the agents along the river and over time. A water trading contract needs to be designed that achieves such a division of water in a fair way both initially and in perpetuity. We therefore need to compute both initial allocations and, using a time-consistent imputation distribution principle (IDP), compute the evolution of the characteristic function of each coalition over time. The results are presented in Table 2 for water consumption and Table 3 for the transfers between agents.

For each coalition $\{5, 4, 3, 2, 1\}$, $\{5, 4, 3, 2\}$, $\{5, 4, 3\}$, $\{5, 4\}$, $\{5\}$ we need to solve a nonlinear programming problem to compute the initial value of the

Table 2: Pareto efficient allocation of water x^*_{ik} over time.

k	1	2	3	4	5
0	61.0352029	42.7246420	29.9072496	20.9350987	14.6545469
1	61.0352035	42.7246425	29.9072496	20.9350718	14.6545528
2	61.0352029	42.7246420	29.9072493	20.9350718	14.6545528
3	61.0352029	42.7246420	29.9072493	20.9350718	14.6545528
4	61.0352029	42.7246420	29.9072491	20.9350717	14.6545526
5	61.0352029	42.7246420	29.9072491	20.9350717	14.6545526
6	61.0352029	42.7246420	29.9072496	20.9350721	14.6545529
7	61.0352032	42.7246423	29.9072496	20.9350721	14.6545529
8	61.0352032	42.7246423	29.9072496	20.9350721	14.6545529
9	61.0352032	42.7246423	29.9072496	20.9350721	14.6545529

Table 3: Pareto efficient transfers t^*_{ik} between agents over time.

k	1	2	3	4	5
0	0.0013859	−1E-06	−1E-06	−1E-06	−1E-06
1	−1E-06	−6.628E-03	−6.628E-03	−6.628E-03	−6.628E-03
2	−6.628E-03	−6.628E-03	−1E-06	−1E-06	3.723E-03
3	3.723E-03	3.081E-03	3.081E-03	3.081E-03	3.081E-03
4	3.081E-03	3.081E-03	3.081E-03	0.0001422	0.0001422
5	−3.446E-03	−3.446E-03	−3.446E-03	−7.269E-03	−7.269E-03
6	−7.269E-03	−7.269E-03	−7.269E-03	−6.373E-03	−6.373E-03
7	−6.373E-03	−6.373E-03	−6.373E-03	−6.373E-03	−6.373E-03
8	−6.373E-03	−6.373E-03	−6.373E-03	−6.373E-03	−6.373E-03
9	−7.486E-03	−7.486E-03	−7.486E-03	−7.486E-03	−7.486E-03

characteristic function at time $k = 0$. We solve the following nonlinear programming problem for each of these coalitions:

$$x_S^{**} = \arg\max \sum_{i \in S} b_i(x_{i0})$$

subject to

$$\sum_{i \in P_j \cap S} x_{i0} \leq \sum_{i \in P_j} e_{i0} \, \forall j \in S.$$

This solution can now be used to compute the evolution of the characteristic function at each point along the river, i.e., for each sub-coalition or predecessor set of agents, over time. This is done by implementing the recursive scheme discussed in Section 3, beginning with the initial values computed above. The solution of this recursive scheme is shown in Table 4.

Variation in the characteristic function over time is largely dependent on the scale and variation of the river flow at each location and each point in time. If the scale is insufficiently large there will be little change in the characteristic function, because there will be insufficient variation in transfers being made between agents at each point in time and at each location. From the perspective of explaining the endogenous formation of water trading contracts via endogenous coalition formation it would be desirable to have the characteristic function of one coalition overtaking the characteristic function of another coalition at some point in time. However, our results show that the grand coalition would form and that trading water leads to a smoothing of consumption and consequently of the characteristic function over time. In other words, the optimal transfers between agents absorb the fluctuations in available water by pricing in the relative scarcity of water. From this perspective the results are not unusual.

Note that because we have employed Pareto optimal transfers at each point in time we have not encountered problems with negativity of the value function that are typical for discrete time multistage coalitional form games. Consequently, we

Table 4: Evolution of characteristic function v_{ik} over time.

k	$v(N,k)$	$v(\{5,4,3,2\},k)$	$v(\{5,4,3\},k)$	$v(\{5,4\},k)$	$v(\{5\},k)$
0	36.92901	27.09009	18.69493	11.50769	5.331069
1	36.98019	27.09005	18.69489	11.50765	5.331032
2	36.98015	27.08760	18.69244	11.50520	5.328580
3	36.97770	27.08515	18.69240	11.50516	5.329957
4	36.97907	27.08629	18.69354	11.50630	5.331097
5	36.98021	27.08743	18.69468	11.51156	5.33635
6	36.97894	27.08618	18.69340	11.50887	5.333670
7	36.97625	27.08346	18.69072	11.50652	5.331313
8	36.97389	27.08111	18.68836	11.50416	5.328957
9	36.97154	27.07875	18.68600	11.50180	5.326600

have not had to resort to regularization methods in computing the evolution of the characteristic function.

8 Conclusion

In this chapter we have examined time consistency properties of compromise solutions concepts for water contracts along a river. We have found that both downstream incremental distribution and the coalitional τ-value are time consistent but that the desirability of the τ-value is reduced compared to that of downstream incremental distribution because the τ-value of the corresponding sequencing game does not satisfy drop-out monotonicity. Consequently, if players were to source water from somewhere other than the river, i.e., drop out of the queue for water, then other players would benefit.

In choosing between these two solution concepts, downstream incremental distribution is to be preferred as a principle for defining water contracts because it possesses less anti-social properties compared with the τ-value.

We computed an example of the evolution of the characteristic function over time for downstream incremental distribution to illustrate the way in which the method works. The computation of fair and time-consistent contracts for water is quite demanding, involving the solution of multiple nonlinear programming problems. For real riverine systems it would be considerably more demanding than for the example presented here. Nevertheless, time consistency goes some way to reducing the computational burden. Altogether we solved six nonlinear programming problems to solve the contract design problem in the example. If we had not been able to resort to time consistency we would have had to solve 51 separate optimization problems. This is a considerable reduction in computational burden.

Possible extensions of this work include considering additional compromise solution concepts and exploring the conditions under which coalitions reform over time. As pointed out by Ambec and Sprumont [1, pp. 460–461] this model can be easily extended to consider the case of branching rivers and the case of sharing the costs of pollutants such as sediments and agricultural run-off as they flow downriver.

References

[1] Ambec, S. and Y. Sprumont (2002) Sharing a River, *Journal of Economic Theory*, Vol. 107, pp. 453–462.

[2] Burness, H.S. and J.P. Quirk (1979) Appropriative Water Rights and the Efficient Allocation of Resources, *The American Economic Review*, Vol. 69, pp. 25–37.

[3] Burness, H.S. and J.P. Quirk (1980) Economic aspects of appropriative water rights, *Journal of Environmental Economics and Management*, Vol. 7, pp. 372–388.

[4] Casas-Mendez, B., I. Garcia-Jurado, A. van den Nouweland and M. Vazquez-Brage (2003) An extension of the τ-value to games with coalition structures, *European Journal of Operational Research*, Vol. 148, pp. 494–513.

[5] Curiel, I., Pederzoli, G. and S. Tijs (1989) Sequencing Games, *European Journal of Operational Research*, Vol. 40, pp. 344–351.

[6] Dementieva, M. (2004) *Regularization in Multistage Cooperative Games*, PhD Thesis, Jvaskala Studies in Computing No. 42, University of Jyvaskala.

[7] Filar, J.A. and L.A. Petrosjan (2000) Dynamic Cooperative Games, *International Game Theory Review* Vol. 2, No. 1, pp. 47–65.

[8] Firoozi, F. and J. Merrifield (2003) An Optimal Timing Model of Water Reallocation and Reservoir Construction, *European Journal of Operational Research*, Vol. 145, pp. 165–174.

[9] Freebairn, J. and J. Quiggin (2004) Water Rights for Variable Supplies, *Risk and Sustainable Management Group, Murray Darling Program Working Paper:2/M04*, School of Economics, The University of Queensland.

[10] Greenberg, J. and S. Weber (1986) Strong Tiebout Equilibrium under Restricted Preferences Domain, *Journal of Economic Theory*, Vol. 38, pp. 101–117.

[11] Hendrickx, R. (2004) *Cooperation and Allocation*, PhD Thesis, CenTeR, University of Tilburg.

[12] Howe, C.W., D.R. Schurmeier and W. Douglas Shaw, Jr. (1986) Innovative approaches to water allocation: the potential for water markets, *Water Resources Research*, Vol. 22, No. 4, pp. 439–445.

[13] Lahmandi-Ayed, R. and M-S. Matoussi (2002) Long-Term Effects of Water Markets: Investing or Rent-seeking? *CERMSEM Universite de Paris I Serie Bleu 2002.11*, CERMSEM Universite de Paris I.

[14] Meinzen-Dick, R.S. and R. Pradhan (2002) Legal Pluralism and Dynamic Property Rights *CAPRi Working paper No. 22 CGIAR System Wide Program on Collective Action and Property Rights*, International Food Policy Research Institute.

[15] Myerson, R.B. (1977) Graphs and Cooperation in Games, *Mathematics of Operations Research*, Vol. 2 No. 3 pp. 225–230.

[16] Petrosjan, L.A. (2005) Cooperative Differential Games, in: A. Nowak, K. Szajowski (Eds.) *Advances in Dynamic Games: Applications to*

Economics, Fianance, Optimization and Stochastic Control, Annals of Dynamic Games 7, Birkhäuser.

[17] Petrosjan, L.A. and G. Zaccour (2003) Time Consistent Shapley Value Allocation of Pollution Cost Reduction, *Journal of Economic Dynamics and Control*, Vol. 27, pp. 381–398.

[18] Petrosjan, L.A. and N.A. Zenkevich (1996) *Game Theory*, World Scientific, Singapore.

[19] Provencher, B. (1993) A Private Property Rights Regime to Replenish a Groundwater Aquifer, *Land Economics*, Vol. 69, pp. 325–40.

[20] Provencher, B. and O. Burt (1994) A Private Property Rights Regime for the Commons: the case for Groundwater, *American Journal of Agricultural Economics*, Vol. 76, pp. 875–888.

[21] Tarashnina, S. (2002) Time-Consistent Solution of a Cooperative Group Pursuit Game, *International Game Theory Review*, Vol. 4, No. 3, pp. 301–317.

[22] Zakharov, V. and M. Dementieva (2004) Multistage Cooperative Games and the Problem of Time Consistency, *International Game Theory Review*, Vol. 6, No. 1, pp. 156–170.

A Hybrid Noncooperative Game Model for Wireless Communications*

Tansu Alpcan
Deutsche Telekom Laboratories
Technische Universität Berlin
Ernst-Reuter-Platz 7
10587 Berlin, Germany
Tansu.Alpcan@tele.kom.de

Tamer Başar
Coordinated Science Laboratory
University of Illinois
1308 West Main Street
Urbana, IL 61801, USA
tbasar@control.csl.uiuc.edu

Abstract

We investigate a hybrid noncooperative game motivated by the practical problem of joint power control and base station (BS) assignment in code division multiple access (CDMA) wireless data networks. We model the integrated power control and BS assignment problem such that each mobile user's action space includes not only the transmission power level but also the respective BS choice. Users are associated with specific cost functions consisting of a logarithmic user preference function in terms of service levels and convex pricing functions to enhance the overall system performance by limiting interference and preserving battery energy. We study the existence and uniqueness properties of pure strategy Nash equilibrium solutions of the hybrid game, which constitute the operating points for the underlying wireless network. Since this task cannot be accomplished analytically even in the simplest cases due to the nonlinear and complex nature of the cost and reaction functions of mobiles, we conduct the analysis numerically using grid methods and randomized algorithms. Finally, we simulate a dynamic BS assignment and power update scheme, and compare it with "classical" noncooperative power control algorithms in terms of aggregate signal-to-interference ratio levels obtained by users.

Key words. Noncooperative games, Nash equilibrium, hybrid games, wireless networks, power control.

*Research supported by NSF grant ITR 00-85917. Tansu Alpcan was with the Coordinated Science Laboratory, University of Illinios, during the course of this research.

1 Introduction

The primary objective of mobile users in wireless networks is to achieve and maintain a satisfactory level of service, which may be described in terms of signal-to-interference ratio (SIR). The mobiles may vary their uplink transmission power levels and connections to the base stations in order to reach this goal. Since in code division multiple access (CDMA) systems signals of other users can be modeled as interfering noise signals, each mobile degrades the level of service of others as a result of these decisions. Hence, the joint power control and base station (BS) assignment problem [1] leads to a conflict of interest between the individual users. In addition, if users have different preferences for the level of service or have varying SIR requirements, then the power control problem can be posed as one of resource allocation. Furthermore, under a distributed power control regime, the mobiles cannot have detailed information on each other's preferences and actions due to communication constraints inherent to the system. It is therefore appropriate to address the joint power control and BS assignment problem within a noncooperative game theoretic framework, where a Nash equilibrium (NE) [2] provides a relevant solution concept.

Several studies exist in the literature that use noncooperative game theoretic schemes to address the power control problem, including the ones by the authors [3–5]. In all of the previous studies, mobiles were considered to be connected to the closest BS, i.e., the one with the lowest channel attenuation. However, the service or SIR level of a mobile is also affected by the number of mobiles in the vicinity of the same BS. Therefore, the performance of the mobiles can be improved by introducing the choice of BS as an additional discrete decision variable in addition to the positive real power level variable. In a related study, Saraydar *et al*. [6] consider a similar problem. However, they analyze the problem in two separate parts involving BS choice and optimization of the power level, and hence, do not model it as a hybrid game.

Accordingly, we consider in this chapter a hybrid noncooperative game motivated by the practical problem of joint power control and BS assignment in CDMA wireless data networks. This constitutes an extension of the model in our earlier studies [3–5] with an additional degree of freedom in optimization. Specifically, we model the integrated power control and BS assignment problem in such a way that each mobile's action space includes not only the transmission power level but also the choice of the BS. The outcomes of the user actions are reflected in a specific cost structure, and each user is associated with a cost function that is parametrized by user-specific prices. The convex pricing schemes considered aim to enhance the overall system performance by limiting the interference and preserving battery energy. We investigate the existence and uniqueness of pure strategy NE solutions of the hybrid game, which constitute the operating points for the underlying wireless network. We make the inherent assumption that a mobile is connected to a single BS. Furthermore, mixed strategy solutions are not feasible for the system considered due to the prohibitive handoff costs in the implementation. We therefore focus on pure strategy solutions. Unfortunately, conditions for the existence and uniqueness of a pure NE cannot be derived analytically even in the simplest cases due to the nonlinear and complex nature of cost and reaction functions of mobiles. Hence, we investigate solutions of this interesting hybrid game numerically using grid methods and randomized algorithms.

The next section describes the wireless network model adopted. We define the hybrid power control game and the cost function in Section 3. Section 4 discusses the NE solution and contains Subsection 4.1 which describes randomized algorithms for numerical analysis. We present our simulation results in Section 5; we investigate the existence and uniqueness properties of NE solutions in Subsection 5.1 and analyze a power update and BS assignment scheme in Subsection 5.2. The paper concludes with a recap of the results and elucidation of directions for future research in Section 6.

2 The Wireless Network Model

We consider a multicell CDMA wireless network model similar to the ones described in [3–5]. The system consists of a set $\mathcal{B} := \{1, \ldots, N\}$ of base stations (BSs) and a set $\mathcal{M} := \{1, \ldots, M\}$ of users. The number of users on the network is limited through an admission control scheme. The ith mobile transmits with a nonnegative uplink power level of $p_i \leq p_{\max}$, where $p_{\max}$ is an upper bound imposed by physical limitations of the mobiles. Figure 1 depicts a simplified picture of the wireless network model considered. We investigate the case where mobiles are given the freedom of choosing the BS connection in addition to determining their transmission power levels. Hence, each mobile connects to a BS which it chooses from the set of BSs on the network, $\mathcal{B}$.

We define $h_{il} p_i$ as the instantaneous received power level from user i at the lth BS. We assume that a mobile connects to one BS only at any given time. The quantity h_{il} $(0 < h_{il} < 1)$ represents the channel gain between the ith mobile and the lth BS [7]. Ignoring fast time-scale fading (such as Rayleigh fading) and

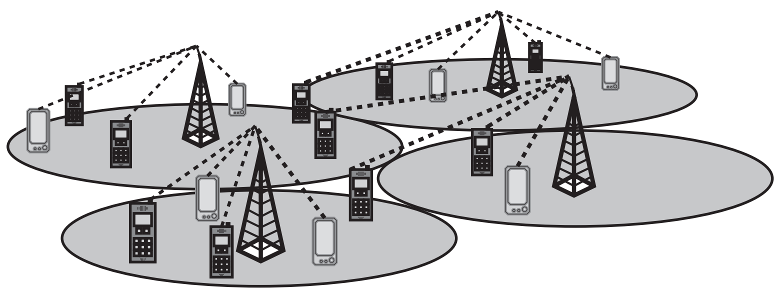


Figure 1: An abstract diagram of the CDMA wireless network model.

shadowing effects in order to simplify the analysis, we model h_{il} in such a way that it depends only on the location of the mobiles with respect to the BSs. Hence, the channel gain h_{il} is given by

$$h_{il} := \left(\frac{0.1}{d_{il}}\right)^2, \tag{1}$$

where d_{il} denotes the (Euclidean) distance of the mobile to the BS, and the loss exponent is chosen as 2, which corresponds to a free space environment [7]. In addition, we assume that the location of a mobile, which is the main factor affecting the channel gain, h, does not change significantly over the time scale of this analysis. This assumption is justified by the fact that the power control algorithm operates with a high frequency.

The level of service a mobile receives is described in terms of SIR [3, 4]. In accordance with the interference model considered, the SIR obtained by mobile i at the base station l is given by

$$\gamma_{il} := \frac{L h_{il} p_i}{\sum_{j \neq i} h_{jl} p_j + \sigma_l^2}. \tag{2}$$

Here, $L := W/R > 1$ is the spreading gain of the CDMA system, where W is the chip rate and R is the data rate of the user. Furthermore, the additional interference at the BS l due to factors other than the transmissions of other mobiles is modeled as a fixed background noise, of variance σ_l^2.

3 The Hybrid Power Control Game

We consider a power control game where each user (mobile) is associated with a specific cost function. Since a user can choose both its power level, which is a continuous variable, and the BS to which it connects, which is discrete in nature, we call this a *hybrid power control game*. The action space, S_i, of the ith user is then defined as

$$S_i = \{(b, p) : b \in \mathcal{B} = \{0, 1, \dots, N\}\ , \text{and}\ \ p \in [0, p_{\max}]\}, \tag{3}$$

and the actions are denoted by $s_i := (b_i, p_i)$. The cost function, J_i, of user i is defined as the difference between the utility function of the user and its pricing function, $J_i = P_i - U_i$. The utility function, $U_i(\mathbf{p}, b_i)$, is chosen as a logarithmic function of the ith user's SIR, which we denote by $\gamma_i(\mathbf{p}, b_i)$. It quantifies approximately the demand or *willingness to pay* of the user for a certain level of service. This utility function can also be interpreted as being proportional to the Shannon capacity [7], if we make the simplifying assumption that the noise plus the interference of all other users constitute an independent Gaussian noise. This means that this part of the utility is simply linear in the throughput that can be achieved (or approached) by user i using an appropriate coding, as a function of its transmission power [3].

The pricing function, $P_i(p_i)$, on the other hand, is imposed by the system to limit the interference created by the mobile, and hence to improve the system performance [6]. At the same time, it can also be interpreted as a cost on the battery usage of the user. It is a convex function of p_i, the power level of the user. Accordingly, the cost function of the ith user is defined as

$$J_i(\mathbf{p}, b_i) = P_i(p_i) - \log(1 + \gamma_i(\mathbf{p}, b_i)), \tag{4}$$

where $\gamma_i(\mathbf{p}, b_i)$ is the ith mobile's SIR level under the given vector of power levels, $\mathbf{p}$, of all mobiles and its BS choice, b_i. One possible pricing function is the linear one, $P_i(p_i) = \lambda_i p_i$, where λ_i is a user-specific parameter. This structure will be used extensively, though not exclusively, in this chapter.

In the hybrid power control game defined, the ith user's optimization problem is to minimize its cost (4), given the sum of power levels of all other users as received at the base stations. Thus, the reaction function of user i is

$$p_i(b_i, \mathbf{p}) = \arg \min_{s_i=(b_i, p_i)} J_i(b_i, \mathbf{p}). \tag{5}$$

Furthermore, if the mobile is connected to the lth BS, i.e., $b_i = l,\ l \in \mathcal{B}$, and P_i is chosen as linear in p_i, $P_i = \lambda_i p_i$, then the optimal power level is given by

$$p_i(b_i = l,\ \mathbf{p}) = \begin{cases} \dfrac{1}{\lambda_i} - \dfrac{1}{Lh_{il}}\left(\displaystyle\sum_{j \neq i} h_{jl} p_j + \sigma_l^2\right), & \text{if } \displaystyle\sum_{j \neq i} h_{jl} p_j < \dfrac{\lambda_i}{Lh_{il}} - \sigma_l^2 \\ 0, & \text{else} \end{cases}. \tag{6}$$

We note that, given a specific BS, the user's optimization problem admits a unique solution (power level) in the linear pricing case, although it may be a boundary solution. The nonnegativity of the power vector ($p_i \geq 0, \forall i$) is an inherent physical constraint of the model. We refer to [3] for details on the boundary analysis and conditions for an inner solution in the linear pricing case and [4] for the analysis of more general convex pricing schemes.

The reaction function (5) is the optimal response of the user to the varying parameters in the model. It depends not only on the user-specific parameters like b_i, λ_i, and h_i but also on the network parameter, L, and total power level received at the BS $l\ (= b_i)$ to which the mobile is connected, $\sum_{j=1}^{M} h_{jl} p_j$. Each BS provides the user the total received power level using the downlink. We refer to [3, 4] for communication constraints and effects. A simplified block diagram of part of the system relevant to the analysis here is shown in Figure 2.

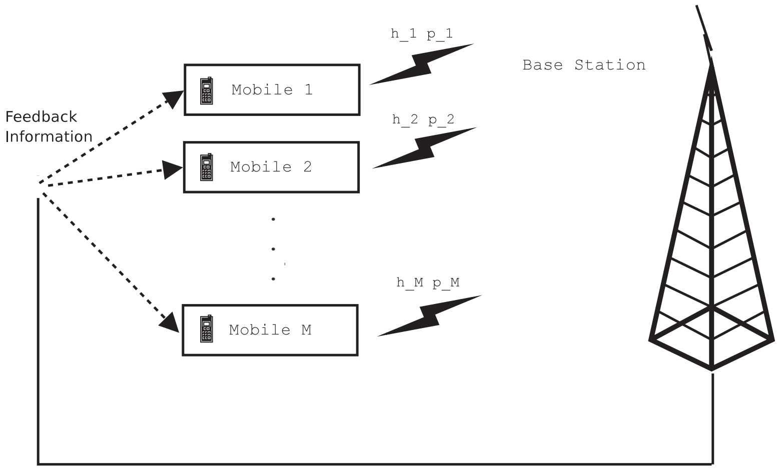


Figure 2: A simplified block diagram of the system.

4 Nash Equilibrium

The Nash equilibrium (NE) in the context of the hybrid power control game is defined as a set of actions, s^*, consisting of the BS choice, b^*, power levels, p^*, and corresponding set of costs, J^*, of users with the property that no user in the system can benefit by modifying its action while the other players keep theirs fixed. Mathematically speaking, the vector of user actions, $\mathbf{s}^* := (\mathbf{b}^*, \mathbf{p}^*)$, is in NE when s_i^* of any ith user is the solution to the following optimization problem given the equilibrium actions of other mobiles, $\mathbf{s}^*_{-i}$:

$$\min_{s_i} J_i(s_i, \mathbf{s}^*_{-i}). \tag{7}$$

We have shown in [3, 4] that once mobiles make the decision on BS connections, the power control game defined admits a unique NE under certain conditions. In the hybrid power control game defined, however, it is not possible to establish the existence or uniqueness of the NE solution analytically. Therefore, we resort to numerical methods, and *randomized algorithms* [8, 9] stand out as a useful tool for numerical analysis [10].

4.1 Randomized Algorithms and Monte Carlo Methods

We utilize randomized algorithms in order to obtain an estimate on the probabilities of existence and uniqueness of NE solutions in the hybrid power control game. Following the approach in [10], these probability estimates are obtained through both *random sampling* and *gridding* techniques on the parameter space. In this case, the parameter space consists of the locations of the mobiles, user-specific parameters such as λ, and system parameters such as L and σ^2. For illustrative purposes, let us call the specific parameter (vector) we are interested in α, while we keep all other parameters fixed.

In the random sampling method, the parameter vector α is chosen to be random with given probability density function f_α, having support sets $\mathcal{S}_\alpha$. We can take, for example, $\mathcal{S}_\alpha$ to be the hyper-rectangular set

$$\mathcal{S}_\alpha = \{\alpha : \alpha_i \in [\alpha_i^-, \alpha_i^+],\ i = 1, 2, \ldots, M\}, \tag{8}$$

and the density function f_α to be uniform on these sets. That is, for $i = 1, 2, \ldots, M$,

$$f_{\alpha_i} = \begin{cases} \dfrac{1}{\alpha_i^+ - \alpha_i^-} & \text{if} \quad \alpha_i \in \left[\alpha_i^-, \alpha_i^+\right] \\ 0 & \text{otherwise} \end{cases}. \tag{9}$$

Then, we generate N_α independent and identically distributed (i.i.d.) vector samples from the set $\mathcal{S}_\alpha$ according to the density function f_α:

$$\alpha^1, \alpha^2, \ldots, \alpha^{N_\alpha}.$$

Another method for generating the sample points of the parameter is to utilize a gridding technique on the support set $\mathcal{S}_\alpha$. In this case, the samples are generated such that they are evenly spaced on the support set. Since this technique provides a nicely distributed sample set, one might think that random sampling is not necessary. However, gridding techniques suffer from a significant drawback called the *curse of dimensionality*. That is, as the dimension of the parameter space increases, the number of samples required to cover the set $\mathcal{B}$ with a uniform grid grows exponentially. Therefore, we resort to random sampling methods when the dimension of the parameter space gets large.

Once the sample points are generated using either random sampling or gridding techniques, we investigate the existence of NE solutions for each sample point or set of parameters. Towards this end, we construct the indicator function

$$\mathcal{I}(\alpha^i) := \begin{cases} 1 & \text{if the game admits a NE} \\ 0 & \text{otherwise} \end{cases}$$

The estimated probability of existence of a NE is readily given by

$$\widehat{p}_{N_\alpha} = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} \mathcal{I}(\alpha^i), \tag{10}$$

which is equivalent to

$$\widehat{p}_{N_\alpha} = \frac{N_{NE}}{N_\alpha},$$

where N_{NE} is the number of vector samples such that the hybrid game admits an NE. A separate indicator function can be defined in a similar manner to compute the probability of having multiple NE solutions.

In order to obtain a "reliable" probabilistic estimate it is important to know how many samples N_α are needed. Let us denote the real probability of existence of a NE by p_α. The Chernoff bound [8] states that for any $\epsilon \in (0,\ 1)$ and $\delta \in (0,\ 1)$ if

$$N_\alpha \geq \frac{1}{2\epsilon^2} \ln\left(\frac{2}{\delta}\right),$$

then, with probability greater than $1-\delta$, we have $\left|\widehat{p}_{N_\alpha} - p_\alpha\right| < \epsilon$. Note that this is a problem-independent explicit bound which can be computed *a priori*. We refer to [8, 9, 11] for a detailed discussion on this issue.

5 Numerical Analysis

We first investigate the existence and uniqueness of Nash equilibrium solutions of the hybrid power control game defined in Section 3 numerically on the wireless network described in Section 2. Then, we simulate a dynamic power update and BS assignment scheme for mobiles.

5.1 Existence and Uniqueness of NE

In order to be able to visualize (some of) the results we start with a one-dimensional (1-D) network model where both mobiles and BSs are located along a line. Although we consider this model only for illustrative purposes, it can also be interpreted in such a way that mobiles staying on a road connect to BSs located at the roadside. In the first simulation, we analyze two mobiles on a 1-D network of two BSs. The system and user parameters are chosen as $L = 128$, $\sigma_1^2 = \sigma_2^2 = 0.1$, and $\lambda_1 = \lambda_2 = 0.1$. For these sets of simulations a linear pricing scheme, where $P = \lambda_i p_i$, is chosen unless otherwise stated. The mobiles are located using a gridding method such that each sample point is evenly spaced with a distance of 0.125. The locations of the two BSs are chosen to be 0.5 and 1.5, respectively. Figure 3 depicts the projected locations of the mobiles and BSs where the game admits a unique or multiple NE. We observe that the game admits a unique NE in all cases except from a single location where both BSs are "equidistant" in terms of SIR levels.

We repeat the analysis above with 3 mobiles instead of 2 on the same network. While the game admits an NE in all of the 4913 sample points, there are 39 points with multiple NEs corresponding to a percentage of 0.79%. Multiple NE solutions occur again at specific (symmetric) locations, as shown in Figure 4. Figure 5, on

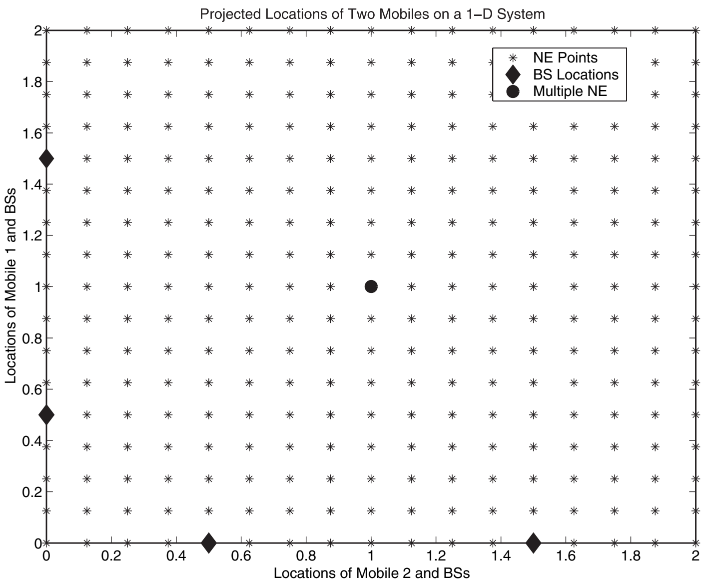


Figure 3: Projected locations of 2 mobiles on a 2BS 1-D network where the game admits an NE.

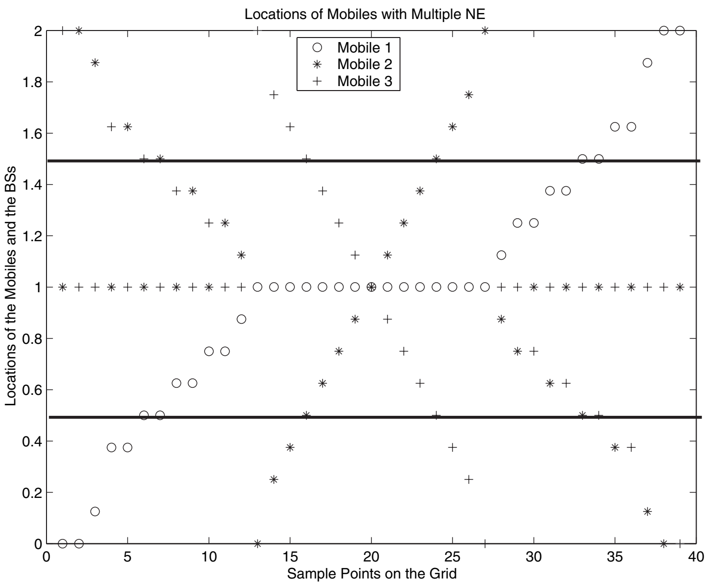


Figure 4: Locations of 3 mobiles on a 2BS 1-D network where the game admits multiple NEs.

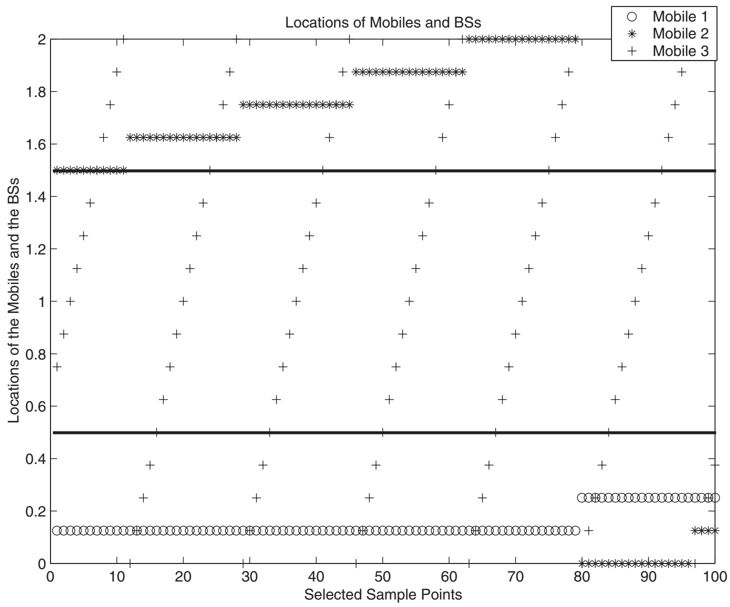


Figure 5: An arbitrary subset of the sample points depicting locations of 3 mobiles on a 2BS 1-D network.

the other hand, depicts an arbitrary subset of 4913 sample locations of mobiles for comparison purposes.

We now investigate the effect of system and user parameters on the NE solutions of the hybrid power control game. We first vary L by choosing its values from the set $\{64,\ 128,\ 192,\ 256\}$. For each value of L, 4913 location points are generated on a grid with a sample distance of 0.25. It is observed that the game admits an NE in all instances. The percentage of samples with multiple NE solutions, on the other hand, is shown in Figure 6. Next, we set $L = 128$ and vary the background noise at the BSs such that $\sigma_1^2,\ \sigma_2^2 \in \{0.1,\ 0.5,\ 1,\ 2\}$. Figure 7 depicts the percentage of multiple NEs, where there is again at least one NE solution in all cases. We finally investigate the effect of pricing parameters $\lambda_1,\ \lambda_2$ while setting $\sigma_1^2 = \sigma_2^2 = 0.1$ and $\lambda_3 = 0.1$. The results are shown in Figure 8. Analyzing the individual cases of the game admitting multiple NE solutions, we conclude that a high percentage of multiple solutions is partly a result of one or more mobiles transmitting with zero NE power due to high prices. In other words, although these cases technically constitute multiple NE solutions, they do not play a significant role in practice. We refer to [3] for a discussion on the relationship between prices and (soft) admission control.

We observe that the simulations conducted using gridding techniques yield somewhat distorted results in terms of multiple NE solutions due to the inherent lattice structure which exhibits specific symmetry properties. Therefore, we repeat

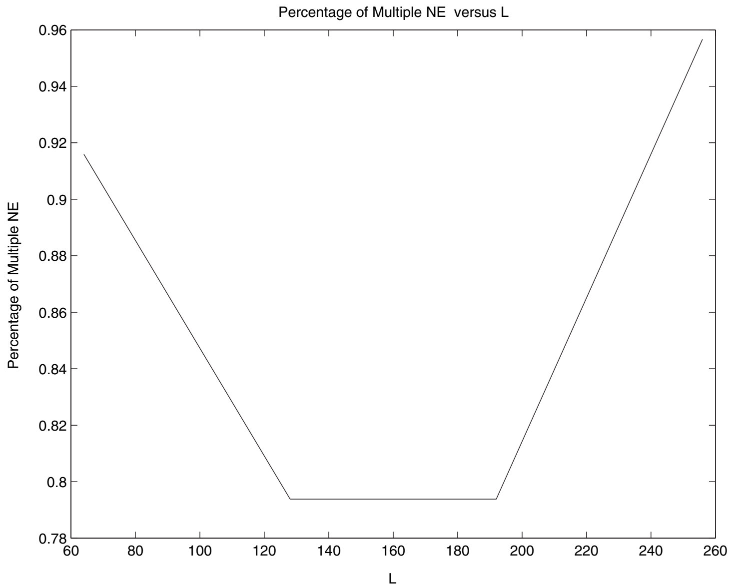


Figure 6: Percentage of mobile locations with multiple NE solutions for different values of L.

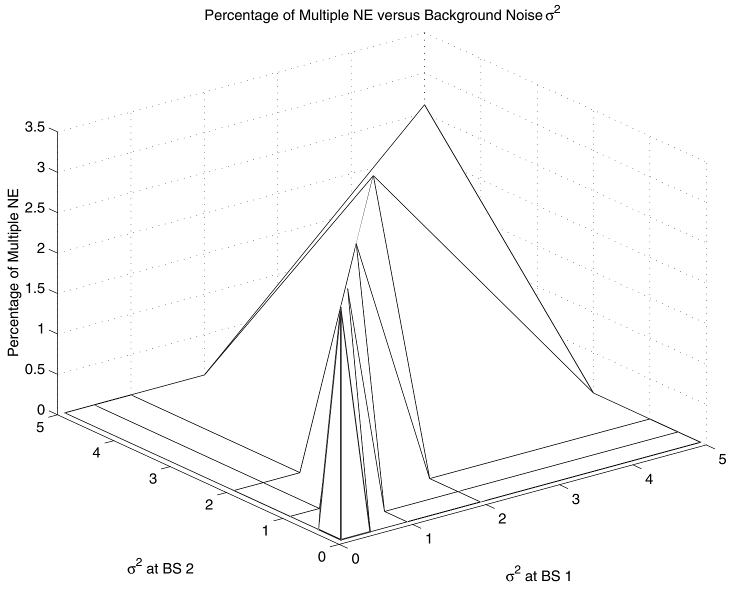


Figure 7: Percentage of mobile locations with multiple NE solutions for different values of σ_1^2 and σ_2^2.

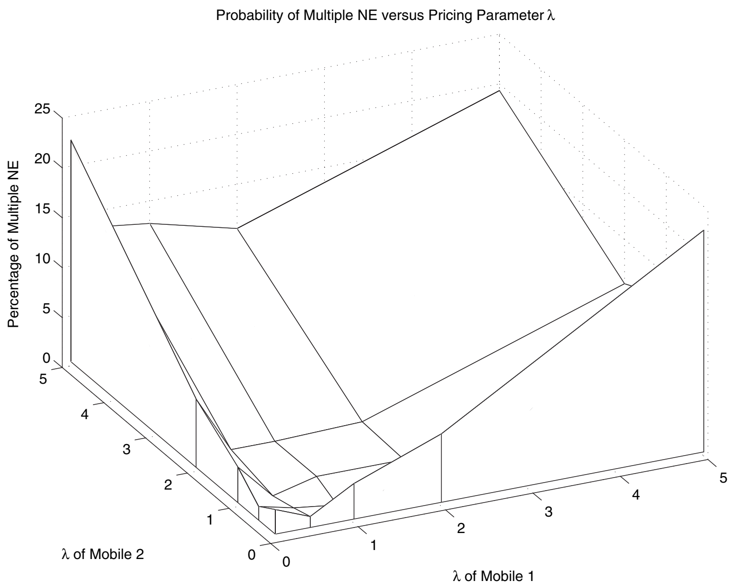


Figure 8: Percentage of mobile locations with multiple NE solutions for different values of λ_1 and λ_2.

the preceding analysis with samples generated using random sampling methods. The locations of the mobiles are chosen randomly with uniform distribution on the support set defined by the boundaries of the network. We repeat the first three simulations with 1000 randomly generated location points. In accordance with the discussion in Section 4.1, 1000 sample points guarantee that our estimates on the percentage of the existence and uniqueness of NE solutions are accurate within $\epsilon = 6\%$ with a probability of at least $1 - \delta = 0.998$. In other words, *Probability* $(|\widehat{p}_N - p| < 0.06) \geq 0.998$, where $\widehat{p}_N$ and p denote respectively the estimated and real probabilities of the analyzed property. In all of these three simulations we observe that the game admits a unique NE with 100% estimated probability. Hence, the probability of having a unique NE solution in 94% of the possible configurations is at least 0.998. This clearly indicates that multiple NE solutions are obtained only at very specific symmetric configurations, which coincide with the lattice structure of the gridding techniques. Figure 9, on the other hand, shows the effect of varying pricing parameters λ_1, λ_2 where $\lambda_3 = 0.1$. We note an overall decrease in the number of instances of the problem with multiple NE solutions due to the random nature of the samples.

Next, we consider a more realistic two-dimensional (2-D) wireless network model. The previous simulations are repeated on the 2-D network again using random sampling. We observe that almost all of the results are comparable with the ones on the 1-D network. Figure 10 depicts the effect of varying pricing

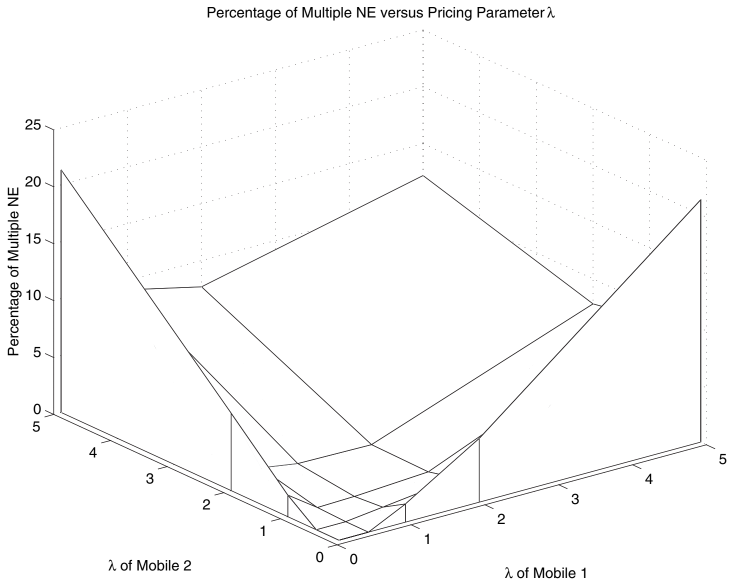


Figure 9: Percentage of mobile locations with multiple NE solutions for different values of λ_1 and λ_2 (random sampling method).

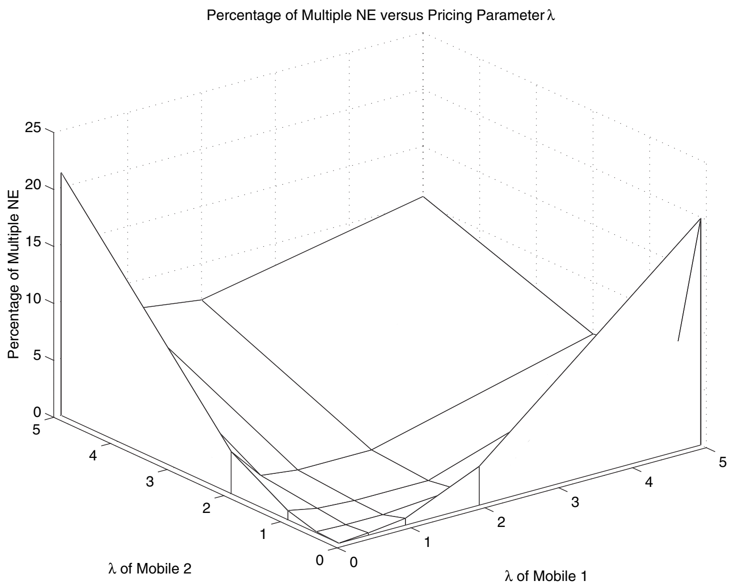


Figure 10: Percentage of mobile locations with multiple NE solutions for different values of λ_1 and λ_2 on a 2-D wireless network.

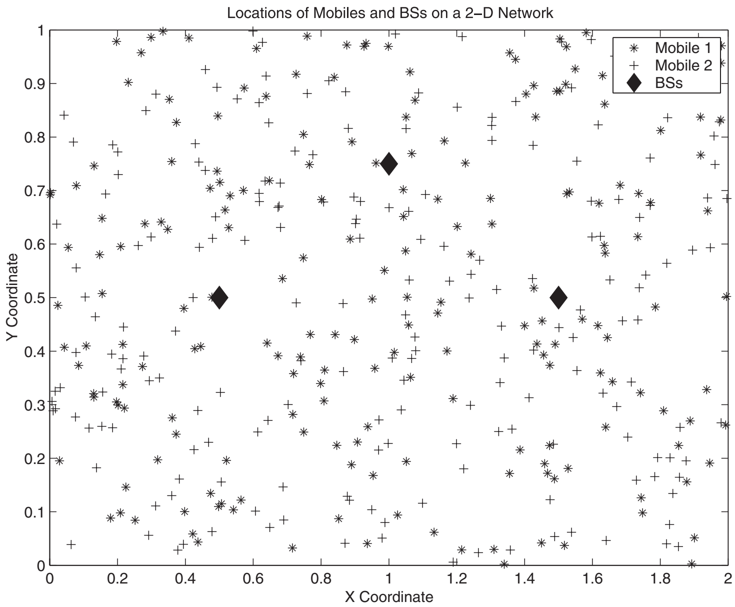


Figure 11: Randomly generated locations of two mobiles on a 2-D wireless network with 3 BSs.

parameters λ_1, λ_2 with $\lambda_3 = 0.1$. In a more realistic simulation, we investigate the existence and uniqueness of NE on a network of 3 BSs with 4 mobiles. The hybrid game admits a unique NE solution in all instances considered corresponding to a probability estimate of 100%. A subset of 1000 randomly generated location points of two of the mobiles and locations of the BSs are depicted in Figure 11.

Finally, we take the pricing function, P_i, in the user costs to be strictly convex, specifically quadratic, $P_i = \lambda_i p_i^2$, where λ_i is again a user-specific parameter. We simulate 3 mobiles on a 2-D wireless network with 2 BSs and study NE solutions for 100 randomly generated location points for varying pricing parameters λ_1, λ_2 where $\lambda_3 = 0.1$. In accordance with the discussion in Section 4.1, 100 sample points guarantee that our estimates on the percentage of the existence of NE solutions are accurate within $\epsilon = 15\%$ with a probability of at least $1 - \delta = 0.98$.

The results shown in Figure 12 are quite different from the ones in the linear pricing case. Instead of having an NE solution in 100% of the cases (and multiple NEs in many of them) we observe that in the majority of the cases there is no pure NE solution. In addition, when it exists the NE solution is observed to be unique in all instances. This result is possibly due to the highly nonlinear nature of the pricing function.

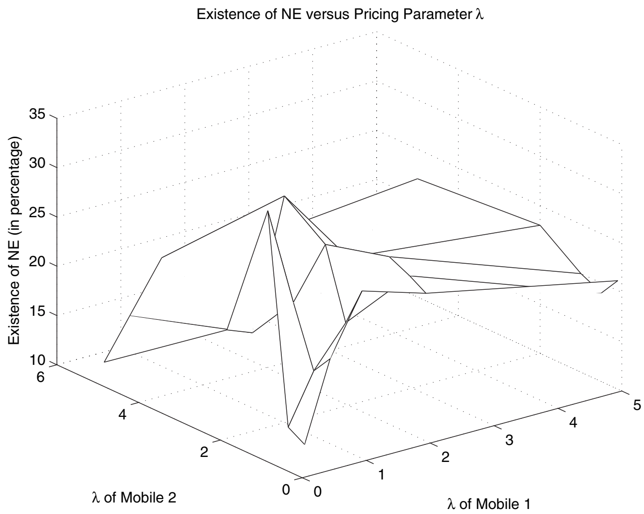


Figure 12: Percentage of mobile locations with NE solutions in the case of quadratic pricing function, $P_i = \lambda_i p_i^2$, for different values of λ_1 and λ_2 on a 2-D wireless network.

5.2 System Dynamics and Convergence

We simulate a joint power update and BS assignment scheme for a 2-D wireless network consisting of 4 BSs and arbitrarily placed 20 mobiles. The system parameters are $L = 128$ and $\sigma_l^2 = 0.1\ \forall l$, and a linear pricing scheme is chosen for this simulation. The locations of BSs and mobiles are shown in Figure 13.

In order to minimize its cost function (7), the ith mobile chooses the BS that maximizes its SIR level and updates its transmission power level using the algorithm (6) such that

$$p_i^{(n+1)} = \frac{1}{\lambda_i} - \frac{1}{L\,h_{il}}\Bigg(\sum_{j \neq i} h_{jl}\, p_j^{(n)} + \sigma_l^2)\Bigg),$$

where n denotes the time, and the BS choice of mobile i is $b_i = l$. In addition, through a projection operation it is ensured that $p_i^{(n)} > 0\ \forall\, i,\ n$. Figures 14 and 15 depict the evolution of the power and SIR levels of the mobiles, respectively. We repeat the same simulation where the mobiles connect only to the nearest BS and update their power levels. The sums of the SIR levels achieved by the users in both cases are compared in Figure 16. Clearly, the additional freedom of BS choice, and hence, the hybrid power control game provides an improvement over "classical" noncooperative power control.

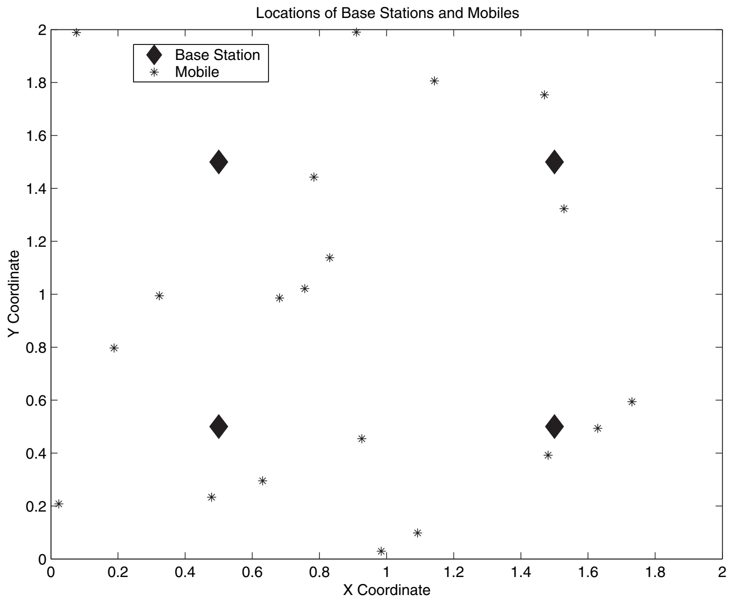


Figure 13: Locations of base stations and mobiles for dynamic simulations.

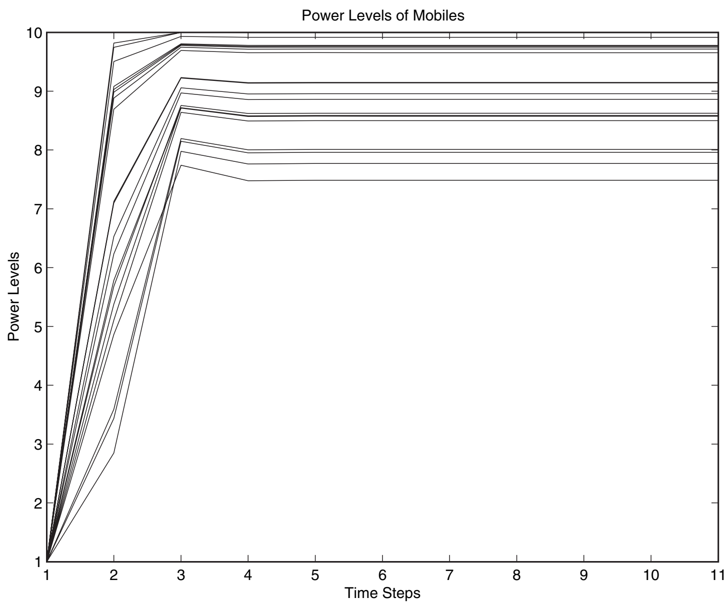


Figure 14: Power levels of mobiles with respect to time.

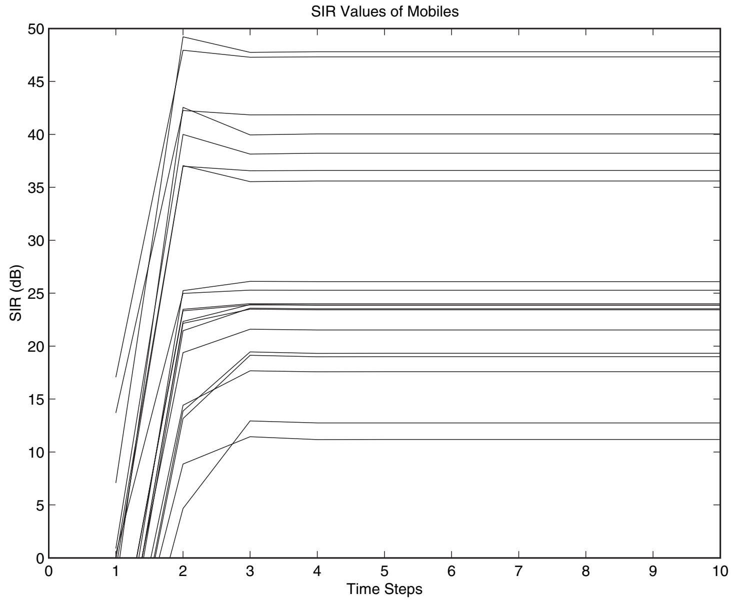


Figure 15: SIR values of mobiles (in dB) with respect to time.

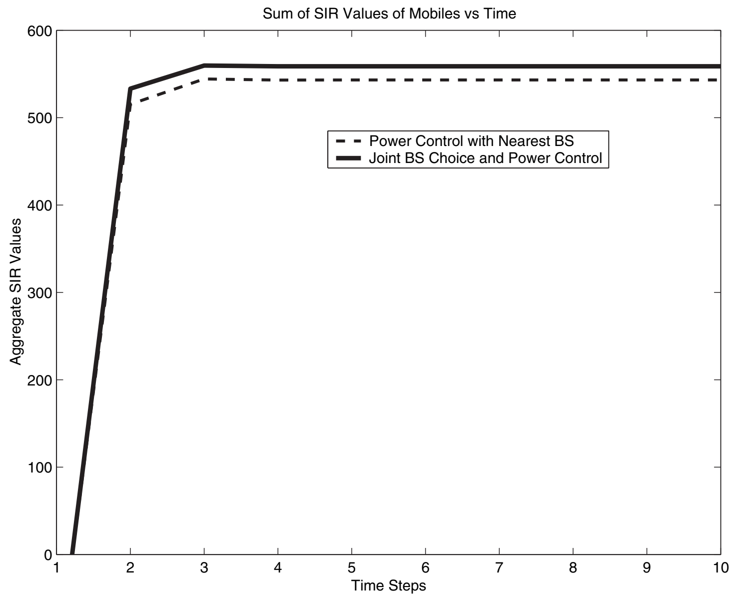


Figure 16: Sums of the SIR values of mobiles (in dB) with respect to time for power control with nearest BS choice and hybrid power control scheme.

6 Conclusions

In this chapter, we have studied a hybrid noncooperative game motivated by the practical problem of joint power control and BS assignment in CDMA wireless data networks, and have extended the noncooperative game theoretic approach in our earlier studies [3–5]. We have developed a hybrid power control game where mobiles are associated with a specific cost structure, and we have investigated the existence and uniqueness of pure NE solutions numerically using randomized algorithms. As part of our numerical analysis we have utilized both gridding techniques and random sampling. The results obtained indicate that the hybrid game admits a unique NE in most of the parameter configurations when a linear pricing scheme is used. We have also encountered multiple NE solutions in specific cases of high prices and symmetric locations of users with respect to the BSs. On the other hand, when a strictly convex pricing scheme is imposed on mobiles we have observed that there exists a pure NE solution only in a minority of the randomly generated configurations. In addition, we have not obtained any multiple NE solutions. Finally, we have simulated a dynamic BS assignment and power update scheme. Simulation results show that the power levels and SIR values of the users converge to their respective equilibrium points. Furthermore, the hybrid power control game has a distinct advantage over "classical" noncooperative power control algorithms in terms of aggregate SIR levels obtained by all users (for most parameter values and configurations). A possible extension of the results here would involve analytical investigation of NE properties as well as convergence characteristics of the joint power update and BS assignment scheme.

References

[1] R. D. Yates and C. Y. Huang, Integrated power control and base station assignment, *IEEE Transactions on Vehicular Technology*, vol. 44, no. 3, pp. 638–644, August 1995.

[2] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. 2nd ed. Philadelphia, PA: SIAM, 1999.

[3] T. Alpcan, T. Başar, R. Srikant, and E. Altman, CDMA uplink power control as a noncooperative game, *Wireless Networks*, vol. 8, pp. 659–669, November 2002.

[4] T. Alpcan and T. Başar, A hybrid systems model for power control in multicell wireless data networks, *Performance Evaluation*, vol. 57, no. 4, pp. 477–495, August 2004.

[5] T. Alpcan, T. Başar, and S. Dey, A power control game based on outage probabilities for multicell wireless data networks, in *Proc. of American Control Conference (ACC) 2004*, Boston, MA, July 2004, pp. 1661–1666.

[6] C. U. Saraydar, N. Mandayam, and D. Goodman, Pricing and power control in a multicell wireless data network, *IEEE Journal on Selected Areas in Communications*, pp. 1883–1892, October 2001.

[7] T. S. Rapaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[8] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems*. London, UK: Springer-Verlag, 2005.

[9] M. Vidyasagar, *A Theory of Learning and Generalization with Applications to Neural Networks and Control Systems*. Berlin, Germany: Springer-Verlag, 1996.

[10] T. Alpcan, T. Başar, and R. Tempo, Randomized algorithms for stability and robustness analysis of high speed communication networks, in *Proc. of IEEE Conference on Control Applications (CCA)*, Istanbul, TURKEY, June 2003, pp. 397–403.

[11] H. Chernoff, A measure of asymptotic efficiency for test of hypothesis based on the sum of observations, *Annals of Mathematical Statistics*, vol. 23, pp. 493–507, 1952.

Incentive-Based Pricing for Network Games with Complete and Incomplete Information

Hongxia Shen
Coordinated Science Laboratory
University of Illinois
1308 West Main Street
Urbana, IL 61801, USA
hshen1@control.csl.uiuc.edu

Tamer Başar
Coordinated Science Laboratory
University of Illinois
1308 West Main Street
Urbana, IL 61801, USA
tbasar@control.csl.uiuc.edu

Abstract

We introduce the concept of nonlinear pricing within the context of our previous Stackelberg network game model and view the Internet Service Provider's (ISP's) policy as an incentive policy, and the underlying game as a reverse Stackelberg game. We study this incentive-design problem under complete information as well as incomplete information. In both cases, we show that the game is not generally *incentive controllable* (that is, there may not exist pricing policies that would lead to attainment of a Pareto-optimal solution), but it is ϵ-incentive controllable (that is, the ISP can induce a behavior on the users that is arbitrarily close to a Pareto-optimal solution). The paper also includes a comparative study of the solutions under linear and nonlinear pricing policies, illustrated by numerical computations.

Key words. Nonlinear pricing, usage-dependent pricing, incentive policy, reverse Stackelberg game, incentive controllability, team solution, incentive-design problem, complete information, incomplete information.

1 Introduction

Recent years have seen a surge of activity in the study of pricing issues in communication networks, and particularly Internet pricing. Pricing has been playing an increasingly important role in controlling congestion and prompting efficient network use, and also in recovering cost and yielding revenue for future network development and expansion. In this chapter, we focus on pricing of the Internet, although the concepts and methods could be applied to other types of communication networks as well. Before we discuss the contributions of this work, we first review some related work on pricing of communication networks. For a more general review of this research area, the reader is referred to [6] and [17].

For a communication network, the resources (which would, for example, be bandwidth in the case of the Internet) that can be shared by the users are generally limited. Thus, one goal of pricing is to achieve efficient network use. In [14], this is posed as maximization of the social welfare (i.e., the sum of all the users' utilities) subject to the capacity constraints. With the Lagrange multipliers associated with the underlying concave program treated as prices, [14] shows that each user's optimization problem given his price becomes decoupled from those of other users, and obtains gradient-based distributed algorithms for the computation of the allocation levels. This work has sparked much interest in modeling a communication network as a resource allocation problem and improving network efficiency, defined as the ratio of the realized social welfare to the maximal social welfare, since the latter may not always be achievable using incentives (or pricing). In particular, [13, 15, 18, 24] assume that users are strategic, instead of being price takers as in [14], and allocate the resources through auctions, as well as analyze the efficiency ratios. The paper [23] assumes that resource providers, service providers, and users are all price takers and introduce an auctioneer for price regulation. An algorithm for the auctioneer is proposed, which finds a welfare maximizing solution in a finite number of iterations. The paper [21] analyzes a similar problem as in [13, 15, 18, 24], but uses a different model and considers a differentiated Quality of Service (QoS). It also introduces an auction-based scheme to maximize the social welfare.

In most of the works referenced above, the network is treated as a social planner such that the objective of pricing is to maximize the social welfare instead of the revenue of resource providers or service providers, although the latter has been investigated in [15, 21, 24]. However, taking into consideration the providers' perspectives is no less important, since lack of this may lead to the outcome that, as stated in [21], "the providers cannot generate sufficient revenue from the network service to sustain the enormous investment required to provide the service." Fortunately, there has been much effort in this direction as well. As in [14], the work [1] also considers a resource allocation problem, but now the prices are set by service providers to maximize their revenue. Efficiency is analyzed with the presence of a monopoly service provider or oligopoly providers, and tight bounds are obtained. Another related work is [12], which studies the interaction between service providers and users, and proposes a market mechanism for bandwidth allocation through double auctions (the resulting allocation is surprisingly efficient as shown in the paper). The impact of pricing on the supply of capacity is also studied in [2], and a mechanism involving long-term contracts between providers and users and a short-term balancing process among users is proposed.

In this paper, we study the relationships of strategic Internet Service Providers (ISPs) and strategic users, and follow the framework of the Stackelberg game model of [4]. In this model, the natural players are the ISPs and the individual users, where the ISPs are the *leaders* and the users are the *followers*. The leaders set the prices for the resources they offer (in this case, bandwidth) and the followers respond by their choices of the amount of bandwidth (or flow) that they are willing to pay for. The payoff for each ISP is the total revenue it collects (minus a fixed cost), and for each user it captures a trade-off between the desire for high bandwidth (or flow) and despise for high congestion cost due to bottleneck links and high payment for flow. Unlike the resource allocation problem in [14] and in some of its following works as previously mentioned, where each user's utility function depends on his own flow, and resources are allocated subject to the capacity constraints, the model here becomes a flow control problem. The presence of congestion cost in the net utility functions of users leads to a natural interaction (coupling) between the decisions of all users using a particular link. This in turn necessitates the modeling of the decision making process among the users (for each fixed pricing policy of the ISPs) as a noncooperative game, with a Nash equilibrium being a natural candidate for a solution. In the case of multiple ISPs, similarly the Nash equilibrium solution concept can be adopted at the higher level of the Stackelberg game.

Our earlier work on this class of problems has looked at *linear*, or *usage-independent*, pricing policies for such network games, and has obtained results on the existence and the uniqueness of an equilibrium, as well as its characterization. This has led to appealing admission policies as well as capacity-expansion schemes when the user population is large. In particular, [4] and [19] have studied the interrelation of a single ISP and multiple users in a single-link network, under uniform pricing and differentiated pricing, respectively. An extension to multiple links has been considered in [5]. In [20], we have extended the single-link model to an environment where the user types may not be known to the ISPs, and studied this network game under incomplete information.

In this chapter, we consider an extension of the earlier model in a significant new direction, where we allow the leader's (let us say there is only one ISP, but still with multiple users) policy to be *nonlinear*, or *usage dependent*, such that it is allowed to depend on the actions of the users. Hence, the ISP's policy may be viewed as an *incentive policy* (see [3, p. 392]). In the previous Stackelberg game with linear pricing, the ISP derives his Stackelberg strategy by "backward induction," that is, he first anticipates the responses of the users to different prices and then accordingly chooses the price that leads to the optimal profit for him (for backward induction, see [7, p. 93]). However, the introduction of nonlinear pricing necessitates a different kind of approach, given the additional functional flexibility: now, the ISP first calculates the action outcome that generates the optimal profit for him, which we refer to as the *team solution*, and then designs the incentive policies to achieve that optimal profit, which we call the *solution of the incentive-design problem*. For this reason, the underlying game is sometimes called a *reverse Stackelberg game*.

Nonlinear pricing of general information goods has been studied before in [22], but this approach has not been applied to communication networks, particularly the Internet. In particular, in some works closely related to ours, such as in [8–11, 16], a service provider either charges the same (unit) price to all the users, or differentiates prices among several service classes or subnetworks, but the users in each class or subnetwork still pay the same (unit) price. The first type of pricing scheme, called *uniform pricing*, can be seen in [10, 11], while the examples of the latter, called *differentiated pricing*, include [8, 9, 11, 16]. Even though sometimes the pricing scheme may be claimed to be "usage based," such as in the one [11], actually the unit price does not depend on a user's usage, and thus the scheme still could be viewed as linear pricing. Obviously, linear pricing is just a special case of nonlinear pricing. Thus, by turning to nonlinear pricing, the ISP takes more control over the users' flows, and consequently could expect an improved profit. On the other hand, a user's payoff cannot deteriorate too much because he always has the choice of not participating, which imposes a restriction on the incentive policies that the ISP can choose. This will be explained in detail in the following sections.

We first provide precise definitions of the terms related to the problem with a single ISP and multiple users, since they will be heavily used here. For a more general context, see [3, pp. 392–396].

Definition 1.1. An *incentive policy*, also referred to as an *incentive function* or simply *incentive*, is defined as the ISP declaring his charge for a user as a function of this user's flow in order to induce a certain amount of flow desired by the ISP.

Definition 1.2. The *team solution* is the action outcome that generates the optimal (maximal) profit for the ISP.

Definition 1.3. The *incentive-design problem* is to find the incentive policy such that the team solution is achieved; this incentive policy is called the *solution of the incentive-design problem*.

Definition 1.4. The underlying Stackelberg game of the incentive-design problem is called a *reverse Stackelberg game*.

Definition 1.5. An incentive-design problem is (linearly or quadratically) *incentive controllable* if the solution of the incentive-design problem exists (which is linear or quadratic, respectively).

We will see later that under some circumstances, the exact team solution may not be achievable by any incentive policy; rather, there exists some incentive policy such that the action outcome can come arbitrarily close to the team solution. In this case, we have the following extended definition of incentive controllability.

Definition 1.6. For an incentive-design problem that is not incentive controllable, if there exists some incentive policy generating an action outcome arbitrarily close to the team solution, then the problem is ϵ*-incentive controllable*, and the corresponding incentive policy and its achieved outcome are ϵ*-team optimal*.

The chapter is organized as follows. We start with the single user case to illustrate the main ideas. We first assume that the ISP knows which type each user belongs to, and hence can deduce the users' (unique) responses to an announced incentive policy. This is called the *complete information game*, and it is dealt with in Section 2. After a brief review of the Stackelberg game model and the solutions developed in [4] and [19], the incentive-design problem is formulated and the team solution is obtained, followed by discussions on incentive controllability. The team solution is also compared with the Stackelberg game solution. Subsequently in Section 3, we extend the study to the *incomplete information game*, where the ISP knows only the probability distribution of the user types, and hence cannot fully deduce the follower responses. Again, we formulate the incentive-design problem and discuss the team solution and incentive controllability. An explicit form of the team solution cannot be obtained in this case; rather, we provide an inductive method to solve for the team solution as well as for the solution of the incentive-design problem. Numerical examples are also provided in this section to illustrate this approach and the results. The paper ends with some discussion on extensions to the multiple user case and also to the multiple ISP case, in Section 4.

2 Complete Information for the Single User Case

2.1 Stackelberg Game Model and Solution Review

First, we review the Stackelberg game model formulated in [4] and generalized in [19], dealing with a link of capacity nc accessed by n users. Here, for the special single user case with $n = 1$ and $c = 1$, the above Stackelberg game model can be stated as follows.

Let x be the flow of the user and p be the price per unit flow charged to him by the ISP. Then, the user's net utility is

$$F_w(x;\, p) = w \log(1 + x) - \frac{1}{1 - x} - px,$$

where w is a positive parameter representing the user's type, which is known to the ISP in this complete information game. Given any fixed price p announced by the leader (ISP), the follower (user) responds with the flow $x(p)$ such that

$$x(p) = \arg \max_{0 \le x < 1} F_w(x;\, p).$$

Then, the ISP's optimal price p^s solves the following optimization problem:

$$p^s = \arg\max\, px(p).$$

Finally, given $w > 1$, the unique Stackelberg game solution, as obtained in [4], is

$$(x^s, p^s) = (x(p^s), p^s) = \left(\frac{w^{\frac{1}{3}} - 1}{w^{\frac{1}{3}} + 1}, \left(w^{\frac{2}{3}} - 1\right)\left(\frac{1}{2}w^{\frac{1}{3}} + \frac{1}{4}\right)\right); \tag{1}$$

if $w \leq 1$, the user always has $x(p) = 0$ for any positive price p.

2.2 Incentive-Design Problem Formulation

In the preceding Stackelberg game, the price per unit flow, p, is usage independent, i.e., it does not depend on the user's flow x; thus, given p, the total charge px is a linear function of x. Now we turn from the above linear pricing to nonlinear pricing. Instead of setting a fixed unit price, the ISP announces the total charge to the user, $r = \gamma(x)$, as a function of his flow, which is not limited to the class of linear functions. Note that if $x = 0$, the ISP cannot charge anything and thus we have $\gamma(0) = 0$. The user's net utility can then be expressed as

$$F_w(x; r) := w \log(1 + x) - \frac{1}{1 - x} - r$$

for $0 < x < 1$, and $F_w(0; r) = -1$ for $x = 0$.

According to Definition 1.2, the team solution is

$$(x^t, r^t) = \arg \max_{0 \leq x < 1, r \geq 0} r, \tag{2}$$

$$\text{s.t. } F_w(x; r) \geq -1, \tag{3}$$

which we assume at this point to exist. The constraint (3) comes from the fact that the user always has the choice of not participating, which guarantees a minimum net utility of -1 for him.

Then, the incentive-design problem, by Definition 1.3, is to find an incentive function, $\gamma : [0, 1) \to \mathcal{R}$, such that

$$\arg \max_{0 \leq x < 1} F_w(x; \gamma(x)) = x^t, \tag{4}$$

$$\gamma(x^t) = r^t, \tag{5}$$

$$\gamma(0) = 0. \tag{6}$$

By Definition 1.5, if there exists a solution to (4) through (6) above, which is then denoted as γ_w^t or just γ^t,[1] we say that the incentive design problem is incentive controllable.

[1]The subscript w of γ is used to emphasize that the ISP's optimal pricing policy depends on w, which is a known parameter to him in this complete information game, and to distinguish this case from the incomplete information game. However, to save notation, we henceforth drop the subscript w of γ, and note that the analysis to follow applies to every fixed type, w.

2.3 Team Solution

For convenience, define

$$Q(x; w) := w \log(1 + x) - \frac{1}{1 - x} + 1.$$

Then, $F_w(x; r) = Q(x; w) - 1 - r$. Hence, (3) is equivalent to $Q(x; w) \geq r$. If $w \leq 1$, $Q(x; w) < 0$ for $0 < x < 1$. Obviously, the only feasible point satisfying (3) is $(x, r) = (0, 0)$. Thus, $(x^t, r^t) = (0, 0)$. In this case, the user will always choose not to participate.

On the other hand, if $w > 1$, then $Q(x; w) > 0$ for $0 < x < x^w_{\max} < 1$, where $x^w_{\max}$ can be obtained by solving $Q(x; w) = 0$. In this case, to maximize r, (3) can be equivalently written as $r = Q(x; w)$. Thus, by (2), $x^t = \arg\max_{0 \leq x < 1} Q(x; w)$ and $r^t = Q(x^t; w)$. Since $Q(x; w)$ is strictly concave in $x \in [0, 1)$, and is increasing at $x = 0$ and decreasing as $x \to 1$, it achieves a unique global maximum at the point where $\frac{\partial Q(x;w)}{\partial x} = 0$, i.e.,

$$x^t = \frac{1 + 2w - \sqrt{1 + 8w}}{2w} =: \alpha(w). \tag{7}$$

To summarize, given any fixed user type, $w > 0$, the team solution is

$$x^t = \alpha(w), \tag{8}$$

$$r^t = Q(\alpha(w); w) = Q(x^t; \alpha^{-1}(x^t)), \tag{9}$$

where $\alpha(w)$ is as defined above for $w > 1$, and $\alpha(w) := 0$ for $0 < w \leq 1$. Note that when we only consider the case $w > 1$, $\alpha : (1, \infty) \to (0, 1)$ is a strictly increasing continuous function, and thus its inverse, $\alpha^{-1} : (0, 1) \to (1, \infty)$, is well defined as

$$\alpha^{-1}(x) := \frac{1 + x}{(1 - x)^2}, 0 < x < 1; \tag{10}$$

while for the case $0 < w \leq 1$, $x^t \equiv 0$ and $\alpha^{-1}(0)$ is not clearly defined, which however does not matter since $r^t = Q(0; \cdot) \equiv 0$.

2.4 Solution of the Incentive-Design Problem

2.4.1 ϵ-Incentive Controllability

With the team solution obtained as above, we next study incentive controllability. Suppose that the problem is incentive controllable, i.e., there exists a γ^t-function that satisfies (4) through (6). Combining (4) and (5), we know that $F_w(x; \gamma^t(x))$ for $0 \leq x < 1$ achieves its maximum $F_w(x^t; r^t) = -1$ at (x^t, r^t). However, from (6), γ^t also needs to go through $(0, 0)$, for which we have $F_w(0; 0) = -1$. In other words, (x^t, r^t) cannot be the unique maximizing point. Since the user is indifferent at these two points, and there is always a possibility that the user may choose his flow to be zero, which will lead to no profit for the ISP, then strictly speaking, the problem is not incentive controllable.

This forces us to look at the extended ϵ-incentive controllability given by Definition 1.6. Actually, by applying the same technique as in Example 7.4 of [3], we can show that the problem is ϵ-incentive controllable and an ϵ-team optimal incentive policy can be found, which leads to an action outcome arbitrarily close to the team solution (x^t, r^t). The method here is as follows: suppose we can find an incentive function satisfying (4) through (6), except for the problem that the solution to (4) may not uniquely be x^t (we denote this function also as γ^t). Then, we revise γ^t by making a small "dip" in the feasible set near (x^t, r^t) to guarantee the uniqueness of the maximizing point; finally, the revised incentive function is ϵ-team optimal and is denoted as $\gamma^{t\epsilon}$. Next, we illustrate this idea for linear incentives, quadratic incentives, and subsequently for more general incentives. Note that this method applies only to the nontrivial case $w > 1$, for which x^t and r^t are positive.

2.4.2 ϵ-Team Optimal Incentive Policies

1) Linear Incentive Functions

A linear function γ passing through $(0, 0)$ and (x^t, r^t) (or an affine function in general, passing through the same points) must be $\gamma(x) = x \cdot r^t/x^t$. Now we need to see whether this linear function satisfies (4) or not. Note that for $0 < a < 1$, $\gamma(ax^t) = ar^t$. Thus, (ax^t, ar^t) is some point along this line between $(0, 0)$ and (x^t, r^t). On the other hand, remember from (9) that $0 = Q(0; w)$ and $r^t = Q(x^t; w)$, where $Q(x; w)$ is strictly concave in x. Hence, $Q(ax^t; w) > ar^t$. As a result,

$$F_w(ax^t; \gamma(ax^t)) = Q(ax^t; w) - 1 - ar^t > -1 = F_w(x^t; r^t).$$

Therefore, (4) cannot be satisfied, and the problem is not linearly incentive controllable or linearly ϵ-incentive controllable. Note that a linear γ corresponds to the constant unit price scheme in [4] and [19], which shows that the classical Stackelberg version of the pricing problem cannot admit a solution that is team optimal or ϵ-team optimal.

2) Quadratic Incentive Functions

Now suppose that $\gamma(x) = a_1 x + a_2 x^2$, which is the most general quadratic function satisfying $\gamma(0) = 0$. For (4) and (5) to hold, we need to have

$$r^t = a_1 x^t + a_2 (x^t)^2, \tag{11}$$

$$\begin{aligned} 0 &= \frac{d}{dx} F_w(x; \gamma(x))\bigg|_{x=x^t} = \frac{\partial Q(x; w)}{\partial x}\bigg|_{x=x^t} - \frac{d\gamma(x)}{dx}\bigg|_{x=x^t} \\ &= \frac{\partial Q(x; w)}{\partial x}\bigg|_{x=x^t} - (a_1 + 2a_2 x^t). \end{aligned} \tag{12}$$

Recall that $\frac{\partial Q(x;w)}{\partial x}\Big|_{x=x^t} = 0$. Thus, from the preceding two equations, we obtain that

$$a_2 = -\frac{r^t}{(x^t)^2} \quad \text{and} \quad a_1 = -2a_2x^t = \frac{2r^t}{x^t}. \tag{13}$$

Note that (12) is necessary but not sufficient for (4) to hold. However, we have the following proposition, whose proof is provided in the Appendix.

Proposition 2.1. *For the quadratic incentive function γ such that $\gamma(x) = a_1x + a_2x^2$ for $0 \leq x < 1$, where a_1 and a_2 are as given in (13), $F_w(x; \gamma(x))$ achieves a unique global maximum in $(0, 1)$ at $x = x^t$.*

Thus, we find a unique quadratic incentive function, $\gamma^t(x) = a_1x + a_2x^2$, that satisfies (4) through (6), with the only problem being that the user might choose not to participate. To resolve this problem, we revise γ^t by making a small "dip" in the feasible set near the team solution to obtain a new incentive function, $\gamma^{t\epsilon}$. For example, we can let $\gamma^{t\epsilon}(x) = \gamma^t(x)$ for $0 \leq x \neq x^t < 1$, and $\gamma^{t\epsilon}(x^t) = \gamma^t(x^t) - \epsilon = r^t - \epsilon$, where ϵ is some small positive number. Then for $\gamma^{t\epsilon}$, x^t becomes the unique optimal solution to (4), where the user receives a net utility of $-1+\epsilon$. Thus, by announcing this incentive policy $\gamma^{t\epsilon}$, the ISP guarantees a profit of $r^t - \epsilon$, which is slightly less than what he gets in the team solution. However, since ϵ can be taken to be arbitrarily small, we finally obtain an ϵ-team optimal incentive policy, $\gamma^{t\epsilon}$.

3) General Incentive Functions

We first provide a graphical illustration for the preceding quadratic incentive policy in Figure 1 (for $w = 100$ as an example).

In the figure, the dashed curve shows $\gamma^t(x) = a_1x + a_2x^2 = 129.5200x - 74.9980x^2$. We can see that it satisfies (5) and (6) by going through $(x^t, r^t) = (0.8635, 55.9196)$ and $(0, 0)$. To see that (4) is also satisfied by γ^t, we depict the curve $F_w(x; r) = -1$, or equivalently, $r = Q(x; w)$. The feasible set for the team solution, defined by (3), is the region below this curve, where the user receives a net utility no less than -1. Clearly, the team solution, (x^t, r^t), is the highest point in this feasible set, where the ISP obtains the largest profit, r^t, while the user's net utility is -1. To make sure that the user chooses the team solution, the ISP's incentive policy γ^t must be above the feasible set, as shown in the figure, such that the user's net utility is no more than -1 along γ^t. Then (4) is satisfied, except that x^t is not the unique solution. To solve that problem, the ISP needs to reduce the charge r^t for x^t by ϵ such that the user gets a net utility of $-1+\epsilon$ at that point, and then the resulting incentive policy, $\gamma^{t\epsilon}$, is ϵ-team optimal.

From the above analysis, by observing the figure it is now obvious why a linear ϵ-team optimal incentive policy does not exist. To summarize, any incentive policy γ^t, not necessarily quadratic, that goes through $(0, 0)$ and (x^t, r^t), and falls above or on $F_w(x; r) = -1$ at all other points, satisfies (4) through (6), except that the user may not stick to the desired flow x^t. Then, by decreasing γ^t a little bit at x^t,

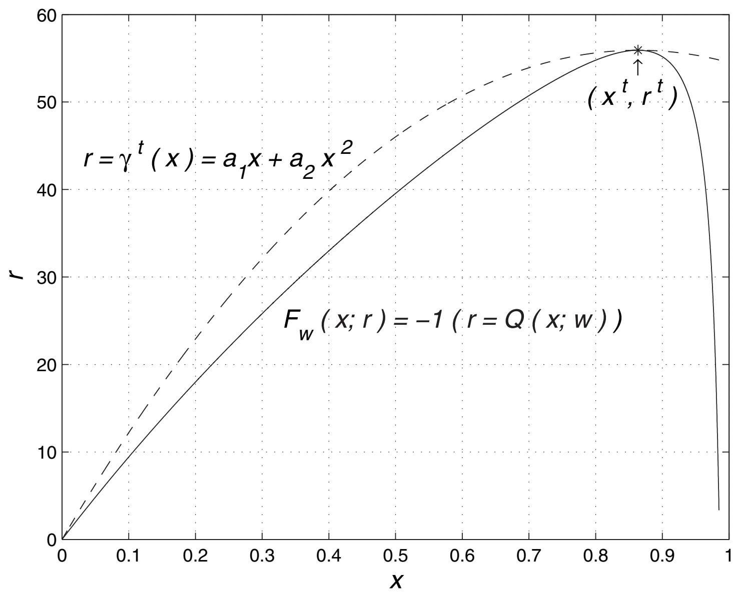


Figure 1: Graphical illustration of incentives for the complete information case ($w = 100$).

we obtain an ϵ-team optimal incentive policy, $\gamma^{t\epsilon}$, and hence the above incentive-design problem is ϵ-incentive controllable.

2.5 Team Solution vs. Stackelberg Game Solution

Now that the problem is ϵ-incentive controllable, we compare the team solution (which is almost achievable by the ISP), given by (8) and (9), with the Stackelberg game solution, given by (1). Note that in (1), p^s is the charge per unit flow; the total profit is

$$r^s = p^s x^s = \left(w^{\frac{1}{3}} - 1\right)^2 \left(\tfrac{1}{2} w^{\frac{1}{3}} + \tfrac{1}{4}\right),$$

given $w > 1$.

In Figure 2, we show comparative results for different values of w. The upper graph compares the total profits for the Stackelberg game and for the reverse Stackelberg game, r^s and r^t, respectively. We can see that the adoption of the ϵ-team optimal incentive policy improves the ISP's profit, which is consistent with our previous conclusion that the classical Stackelberg game cannot admit a team optimal or ϵ-team optimal solution. Furthermore, from the ratio r^t/r^s illustrated in the lower graph, we can see that for w near 1, r^t is almost two times r^s and as w increases to 200, r^t is still around 60 percent higher than r^s.

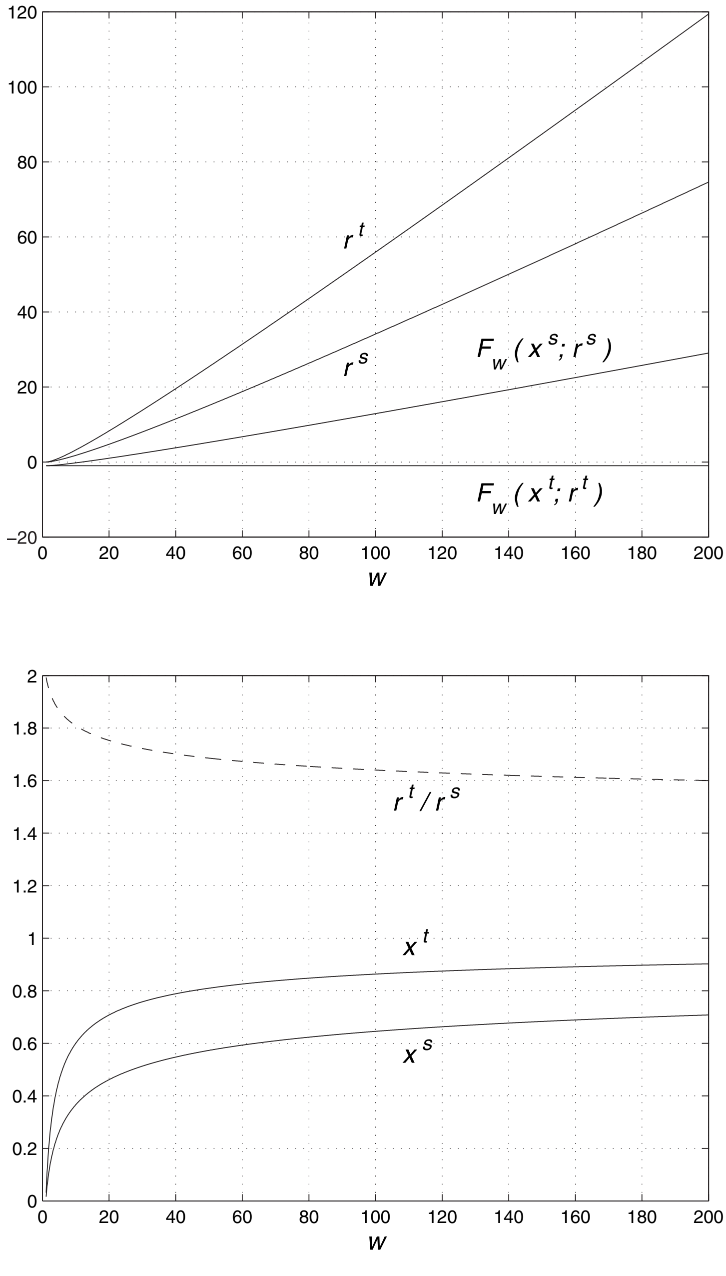


Figure 2: Comparison of the team solution with the Stackelberg game solution for the complete information case.

Also in Figure 2, we compare the user's flows, x^s and x^t, and his net utilities, $F_w(x^s; r^s)$ and $F_w(x^t; r^t)$, for the two games. The results show that compared with the Stackelberg game solution, a larger flow is induced from the user by the ϵ-team optimal incentive policy, though he gets a smaller net utility of $-1+\epsilon$.

3 Incomplete Information for the Single User Case

In this section, we extend our study on the special single user case to the incomplete information game. In this game, the user's type, w, is not revealed to the ISP; rather, the ISP knows only the distribution of the user's type, say (assuming a discrete distribution), $w = w^i$ w.p. $q_i \in (0, 1)$ for $i \in \{1, \ldots, m\}$, where $\sum_{j=1}^{m} q_j = 1$.

3.1 Incentive-Design Problem Formulation

We first formulate the incentive-design problem for the incomplete information game. Remember that if the user's type is w^i, $i \in \{1, \ldots, m\}$, then his net utility is

$$F_{w^i}(x^i; r^i) = w^i \log(1 + x^i) - \frac{1}{1 - x^i} - r^i$$

for $0 < x^i < 1$, and $F_{w^i}(0; r^i) = -1$, where x^i is the user's flow and r^i is the ISP's charge. However, since the user's true type is not known to the ISP, the best policy for him is to maximize the expected profit, and thus the team solution, by Definition 1.2, is

$$\{(x^{it}, r^{it})\}_{i=1}^{m} = \arg \max_{\{0 \le x^i < 1, r^i \ge 0\}_{i=1}^{m}} \left\{ E[r] = \sum_{j=1}^{m} q_j r^j \right\}, \tag{14}$$

$$\text{s.t. } F_{w^i}(x^i; r^i) \ge -1, i \in \{1, \ldots, m\}, \tag{15}$$

$$F_{w^i}(x^i; r^i) \ge F_{w^i}(x^j; r^j), i, j \in \{1, \ldots, m\}, \ i \ne j, \tag{16}$$

where we again assume at this point that a solution exists. The constraint (16) is necessary so that a user with the type w^i will choose (x^{it}, r^{it}) which is the flow-charge pair desired for him.

Then, by Definition 1.3, the incentive-design problem is to find an incentive function, $\gamma : [0, 1) \to \mathcal{R}$, such that

$$\arg \max_{0 \le x < 1} F_{w^i}(x; \gamma(x)) = x^{it}, i \in \{1, \ldots, m\}, \tag{17}$$

$$\gamma(x^{it}) = r^{it}, i \in \{1, \ldots, m\}, \tag{18}$$

$$\gamma(0) = 0. \tag{19}$$

By Definition 1.5, if the solution to (17) through (19) exists, we say that the incentive-design problem is incentive controllable, and denote this solution as γ^t.

3.2 Constraint Reduction

From the analysis for the complete information case, specifically (8) and (9), we know that if a user's type as captured by the parameter w is no more than 1, he will always choose not to participate. Thus, without any loss of generality, we can assume that $w^1 > \cdots > w^m > 1$. To compute the team solution, we first deduce some properties that this solution must satisfy.

Lemma 3.1. *Given $w^i > w^j > 1$, for any $x^i \in [0,1)$, $x^j \in [0,1)$, $r^i \geq 0$, and $r^j \geq 0$ such that*

$$F_{w^i}(x^i; r^i) \geq F_{w^i}(x^j; r^j) \quad \text{and} \quad F_{w^j}(x^j; r^j) \geq F_{w^j}(x^i; r^i) \tag{20}$$

are satisfied, we must have $x^i \geq x^j$.

Proof. Since $w^i > w^j$, (20) implies

$$\begin{aligned} F_{w^i}(x^i; r^i) - F_{w^j}(x^i; r^i) &= (w^i - w^j)\log(1 + x^i) \\ \geq F_{w^i}(x^j; r^j) - F_{w^j}(x^j; r^j) &= (w^i - w^j)\log(1 + x^j). \end{aligned}$$

Since $w^i > w^j$, $w^i - w^j > 0$ and thus $x^i \geq x^j$. □

Since the team solution must satisfy (16), a direct implication of Lemma 3.1 is the following.

Corollary 3.1. *For $w^1 > \cdots > w^m > 1$, $x^{1t} \geq \cdots \geq x^{mt} \geq 0$. Furthermore, for $i < j$, $i, j \in \{1, \ldots, m\}$,*

$$F_{w^i}(x^{it}; r^{it}) \geq F_{w^i}(x^{jt}; r^{jt}) \geq F_{w^j}(x^{jt}; r^{jt}) \geq F_{w^j}(x^{it}; r^{it}).$$

Now we are in a position to prove the following theorem, which plays an important role in the derivation of $\{(x^{it}, r^{it})\}_{i=1}^m$.

Theorem 3.1. *For $w^1 > \cdots > w^m > 1$, the constraints for the team solution $\{(x^{it}, r^{it})\}_{i=1}^m$, (15) and (16), can be simplified as*

$$\begin{aligned} &x^1 \geq \cdots \geq x^m \geq 0; \\ &F_{w^m}(x^m; r^m) = -1, && (21) \\ &F_{w^i}(x^i; r^i) = F_{w^i}(x^{i+1}; r^{i+1}), \quad i \in \{1, \ldots, m-1\}. && (22) \end{aligned}$$

Proof. By Corollary 3.1, $\{(x^{it}, r^{it})\}_{i=1}^m$ must satisfy $x^{1t} \geq \cdots \geq x^{mt} \geq 0$. Now we prove (21) and (22) in several steps.

Step 1: From Corollary 3.1, we know that for any $i < m$, the solution must satisfy

$$F_{w^i}(x^{it}; r^{it}) \geq F_{w^i}(x^{mt}; r^{mt}) \geq F_{w^m}(x^{mt}; r^{mt}).$$

Therefore, the constraint (15) can be reduced to $F_{w^m}(x^m; r^m) \geq -1$. Actually, to maximize the expected profit in (14), we should make the r^i's as large as possible, or equivalently, the $F_{w^i}(x^i; r^i)$'s as small as possible. Therefore, $F_{w^m}(x^{mt}; r^{mt})$ must take the smallest value -1, and (15) finally reduces to (21). Note that taking the equality in (21) has no conflict with the comparative relationships defined by (16).

Step 2: Fix (x^m, r^m) such that $x^m \geq 0$ and $F_{w^m}(x^m; r^m) = -1$. Then (16) can be rewritten as

$$F_{w^i}(x^i; r^i) \geq F_{w^i}(x^m; r^m) \quad \text{and} \quad F_{w^m}(x^i; r^i) \leq F_{w^m}(x^m; r^m), \quad i \in \{1, \dots, m-1\}; \tag{23}$$

$$F_{w^i}(x^i; r^i) \geq F_{w^i}(x^j; r^j), \quad i, j \in \{1, \dots, m-1\},\ i \neq j. \tag{24}$$

In fact, for $i < j$, $i, j \in \{1, \dots, m-1\}$,

$$\begin{aligned} F_{w^i}(x^i; r^i) - F_{w^j}(x^j; r^j) &\geq F_{w^i}(x^j; r^j) - F_{w^j}(x^j; r^j) \\ &\geq F_{w^i}(x^m; r^m) - F_{w^j}(x^m; r^m). \end{aligned}$$

The first inequality comes from (24) and the second inequality holds since $x^j \geq x^m \geq 0$. Therefore, $F_{w^j}(x^j; r^j) \geq F_{w^j}(x^m; r^m)$ implies that $F_{w^i}(x^i; r^i) \geq F_{w^i}(x^m; r^m)$. On the other hand,

$$\begin{aligned} F_{w^m}(x^i; r^i) &= F_{w^j}(x^i; r^i) - (w^j - w^m)\log(1 + x^i) \\ &\leq F_{w^m}(x^j; r^j) = F_{w^j}(x^j; r^j) - (w^j - w^m)\log(1 + x^j), \end{aligned}$$

because of (24) and the fact that $x^i \geq x^j \geq 0$. So, $F_{w^m}(x^j; r^j) \leq F_{w^m}(x^m; r^m)$ implies that $F_{w^m}(x^i; r^i) \leq F_{w^m}(x^m; r^m)$. As a result, (23) can be reduced to

$$F_{w^{m-1}}(x^{m-1}; r^{m-1}) \geq F_{w^{m-1}}(x^m; r^m), \tag{25}$$

$$F_{w^m}(x^{m-1}; r^{m-1}) \leq F_{w^m}(x^m; r^m). \tag{26}$$

Again, to maximize the expected profit, given (x^m, r^m) as fixed, $F_{w^{m-1}}(x^{m-1}; r^{m-1})$ must take the smallest value such that (25) becomes

$$F_{w^{m-1}}(x^{m-1}; r^{m-1}) = F_{w^{m-1}}(x^m; r^m),$$

and then (26) is just equivalent to $x^{m-1} \geq x^m$, which is automatically satisfied. Note that the above equality has no conflict with the comparative relationships defined by (24).

Next, we let $k = m - 1$ and proceed to Step 3.

Step 3: Fix $\{(x^i, r^i)\}_{i=k}^m$ such that $x^k \geq \cdots \geq x^m \geq 0$, $F_{w^m}(x^m; r^m) = -1$, and $F_{w^i}(x^i; r^i) = F_{w^i}(x^{i+1}; r^{i+1})$ for $i \in \{k, \dots, m-1\}$. Rewrite (24) as

$$F_{w^i}(x^i; r^i) \geq F_{w^i}(x^k; r^k) \quad \text{and} \quad F_{w^k}(x^i; r^i) \leq F_{w^k}(x^k; r^k), \quad i \in \{1, \dots, k-1\}; \tag{27}$$

$$F_{w^i}(x^i; r^i) \geq F_{w^i}(x^j; r^j), \quad i, j \in \{1, \dots, k-1\},\ i \neq j. \tag{28}$$

Following a similar argument as in Step 2, (27) can be reduced to

$$F_{w^{k-1}}(x^{k-1}; r^{k-1}) = F_{w^{k-1}}(x^k; r^k)$$

and $x^{k-1} \geq x^k$, which has no conflict with the comparative relationships defined by (28).

Step 4: If $k > 2$, decrease k by 1 and repeat the above Step 3. Note that (24) should be replaced by (28) obtained in the previous iteration.

Finally, by repeating Step 3 for k from $m-1$ to 2, we inductively prove the theorem. □

In the following analysis, the functions Q, α, and α^{-1} are as defined in Section 2.3 for the complete information game.

Corollary 3.2. *For $w^1 > \cdots > w^m > 1$, $\{(x^{it}, r^{it})\}_{i=1}^m$ must satisfy $x^{1t} = \alpha(w^1)$ and $x^{it} \leq \alpha(w^i)$ for $i \in \{2, \ldots, m\}$.*

Proof. We prove this corollary by induction.

First, we must have $x^{mt} \leq \alpha(w^m)$. If this does not hold, i.e., $x^{mt} > \alpha(w^m)$, then we can show that $\{(x^{it}, r^{it})\}_{i=1}^m$ cannot yield the optimal expected profit, which contradicts the definition of the team solution. Actually, we may change $\{(x^i, r^i)\}_{i=1}^m$ from $\{(x^{it}, r^{it})\}_{i=1}^m$ to $\{(x^{i*}, r^{i*})\}_{i=1}^m$ such that $x^{m*} = \alpha(w^m)$, $x^{i*} = x^{it}$ for $i \neq m$ and the constraints given in Theorem 3.1 are satisfied. Obviously, there is no problem with the constraint $x^1 \geq \cdots \geq x^m \geq 0$, since x^m decreases from x^{mt} to $\alpha(w^m)$ and all the other x^i's remain unchanged. We also need to check (21) and (22). Note that (21) is equivalent to $r^m = Q(x^m; w^m)$. Thus, $r^{m*} = Q(\alpha(w^m); w^m) > r^{mt} = Q(x^{mt}; w^m)$, since $Q(x^m; w^m)$ achieves a unique global maximum in $[0, 1)$ at $x^m = \alpha(w^m)$. On the other hand, since x^m decreases, by (22),

$$\begin{aligned} F_{w^{m-1}}(x^{m-1}; r^{m-1}) &= F_{w^{m-1}}(x^m; r^m) \\ &= F_{w^m}(x^m; r^m) + (w^{m-1} - w^m)\log(1 + x^m) \end{aligned}$$

decreases as well. Since $x^{(m-1)*} = x^{(m-1)t}$, we have $r^{(m-1)*} > r^{(m-1)t}$. Again by (22), $F_{w^{m-2}}(x^{m-2}; r^{m-2}) = F_{w^{m-2}}(x^{m-1}; r^{m-1})$, which decreases since r^{m-1} increases, and thus $r^{(m-2)*} > r^{(m-2)t}$; and so on. Finally, we have $r^{i*} > r^{it}$ for all i's and thus $\{(x^{it}, r^{it})\}_{i=1}^m$ cannot be optimal, which is a contradiction.

Next, for some k such that $2 \leq k \leq m$, given $x^{it} \leq \alpha(w^i)$ for $i \in \{k, \ldots, m\}$, we must have $x^{(k-1)t} \leq \alpha(w^{k-1})$. If this does not hold, i.e., $x^{(k-1)t} > \alpha(w^{k-1}) > \alpha(w^i)$ for $i \in \{k, \ldots, m\}$, then we may decrease x^{k-1} from $x^{(k-1)t}$ to $\alpha(w^{k-1})$ and keep all the other x^i's unchanged. Following a similar argument as above, we can show that the r^i's improve for $i \in \{1, \ldots, k-1\}$ (while the r^i's remain unchanged for $i \in \{k, \ldots, m\}$), which leads to a contradiction.

In fact, given fixed $x^{it} \leq \alpha(w^i)$ and r^{it} for $i \in \{2, \ldots, m\}$, the constraints in Theorem 3.1 reduce to $x^1 \geq \alpha(w^2)$ and $F_{w^1}(x^1; r^1) = F_{w^1}(x^2; r^2)$. Thus, in order to maximize r^1, we must have $x^{1t} = \alpha(w^1)$. □

3.3 Team Solution

Now based on Theorem 3.1, the team solution can be computed by induction on m. For that purpose, we equivalently write (21) and (22) as

$$r^m = Q(x^m; w^m), \tag{29}$$

$$r^i = r^{i+1} + Q(x^i; w^i) - Q(x^{i+1}; w^i), \quad i \in \{1, \ldots, m-1\}. \tag{30}$$

Then the expected profit $E[r]$ in (14) can be expressed as a function of x^i's. Therefore, the problem defined by (14) through (16) is equivalent to

$$\{x^{it}\}_{i=1}^m = \arg \max_{1 > x^1 \geq \cdots \geq x^m \geq 0} E[r], \tag{31}$$

followed by

$$r^{mt} = Q(x^{mt}; w^m), \tag{32}$$

$$r^{it} = r^{(i+1)t} + Q(x^{it}; w^i) - Q(x^{(i+1)t}; w^i), \quad i \in \{1, \ldots, m-1\}. \tag{33}$$

3.3.1 Incomplete Information with a Two-Type User

For $w^1 > w^2 > 1$ and $q_1 + q_2 = 1$, $r^2 = Q(x^2; w^2)$ by (29) and $r^1 = r^2 + Q(x^1; w^1) - Q(x^2; w^1)$ by (30). Thus, the expected profit is

$$\begin{aligned} E[r] = \sum_{j=1}^{2} q_j r^j &= q_1 Q(x^1; w^1) - q_1 Q(x^2; w^1) + Q(x^2; w^2) \\ &= q_1 Q(x^1; w^1) + q_2 Q(x^2; v^{2/2}), \end{aligned}$$

where

$$v^{2/2} := \frac{w^2 - q_1 w^1}{q_2} < w^2.$$

Remember that $Q(x; w)$ achieves a unique global maximum in $[0, 1)$ at $x = \alpha(w)$, while α is 0 in $(-\infty, 1]$ [2] and is strictly increasing in $[1, \infty)$. So, it is clear that the solution to (31) is $x^{1t} = \alpha(w^1) > x^{2t} = \alpha(v^{2/2})$. Then the optimal expected profit is $E[r^t] = q_1 Q(x^{1t}; \alpha^{-1}(x^{1t})) + q_2 Q(x^{2t}; \alpha^{-1}(x^{2t}))$.

[2] Considering the possible case that $v^{2/2} \leq 0$, here we need to extend the definition of α such that $\alpha(w) := 0$ for $w \leq 0$. However, we can still verify that $Q(x; w)$ is strictly increasing in $x \in [0, \alpha(w)]$ and strictly decreasing in $x \in [\alpha(w), 1)$, and it attains a unique global maximum in $[0, 1)$ at $x = \alpha(w)$, which will be useful for our following analysis.

3.3.2 Incomplete Information with a Three-Type User

For $w^1 > w^2 > w^3 > 1$ and $\sum_{j=1}^{3} q_j = 1$, by (29) and (30),

$$\begin{aligned}E[r] = \sum_{j=1}^{3} q_j r^j &= q_1 Q(x^1; w^1) - q_1 Q(x^2; w^1) + (q_1 + q_2) Q(x^2; w^2) \\ &\quad - (q_1 + q_2) Q(x^3; w^2) + Q(x^3; w^3) \\ &= q_1 Q(x^1; w^1) + q_2 Q(x^2; v^{2/3}) + q_3 Q(x^3; v^{3/3}),\end{aligned}$$

where

$$v^{2/3} := \frac{(q_1 + q_2)w^2 - q_1 w^1}{q_2} < w^2 \quad \text{and} \quad v^{3/3} := \frac{w^3 - (q_1 + q_2)w^2}{q_3} < w^3.$$

To solve (31), we always have $x^{1t} = \alpha(w^1)$, while the values of x^{2t} and x^{3t} depend on how $v^{2/3}$ is compared with $v^{3/3}$. The simple case is if $v^{2/3} \geq v^{3/3}$, then $x^{2t} = \alpha(v^{2/3}) \geq x^{3t} = \alpha(v^{3/3})$. On the other hand, if $v^{2/3} < v^{3/3}$, then we have the following proposition.

Proposition 3.1. *For the incomplete information case with a three-type user, if $v^{2/3} < v^{3/3}$, then $\alpha(v^{2/3}) \leq x^{2t} = x^{3t} \leq \alpha(v^{3/3})$.*

Proof. $v^{2/3} < v^{3/3}$ implies $\alpha(v^{2/3}) \leq \alpha(v^{3/3})$. Thus, $Q(x; v^{2/3})$ and $Q(x; v^{3/3})$ are both strictly increasing in $x \in [0, \alpha(v^{2/3})]$ and strictly decreasing in $x \in [\alpha(v^{3/3}), 1)$. As a result, x^{2t} and x^{3t} must be in the region $[\alpha(v^{2/3}), \alpha(v^{3/3})]$ subject to the constraint that $x^{2t} \geq x^{3t}$. In this region, $Q(x^2; v^{2/3})$ is strictly decreasing, while $Q(x^3; v^{3/3})$ is strictly increasing. Hence, to maximize $E[r]$, x^3 must take the largest possible value, which leads to $x^{2t} = x^{3t}$. □

Now for the case $v^{2/3} < v^{3/3}$, knowing that $x^{2t} = x^{3t}$, we can substitute x^2 for x^3 and further simplify the expression of $E[r]$ as

$$E[r] = q_1 Q(x^1; w^1) + (q^2 + q^3) Q(x^2; v^{2,3/3}),$$

where

$$v^{2/3} < v^{2,3/3} := \frac{w^3 - q_1 w^1}{q_2 + q_3} = \frac{q_2 v^{2/3} + q_3 v^{3/3}}{q_2 + q_3} < v^{3/3}.$$

Immediately, we have $x^{2t} = x^{3t} = \alpha(v^{2,3/3})$.

Finally, the optimal expected profit is $E[r^t] = \sum_{j=1}^{3} q_j Q(x^{jt}; \alpha^{-1}(x^{jt}))$.

3.3.3 Incomplete Information with a Multiple-Type User

For $m \geq 4$, $w^1 > \cdots > w^m > 1$, and $\sum_{j=1}^{m} q_j = 1$, by (29) and (30),

$$E[r] = \sum_{j=1}^{m} q_j r^j = \sum_{k=1}^{m-1} \sum_{l=1}^{k} q_l [Q(x^k; w^k) - Q(x^{k+1}; w^k)] + Q(x^m; w^m)$$
$$= q_1 Q(x^1; w^1) + \sum_{k=2}^{m} q_k Q(x^k; v^{k/m}),$$

where

$$v^{k/m} := \frac{\sum_{l=1}^{k} q_l w^k - \sum_{l=1}^{k-1} q_l w^{k-1}}{q_k} < w^k, \quad k \in \{2, \ldots, m\}.$$

Then we have the following corollary as an extension of Corollary 3.2. The proof is similar, and hence is omitted here.

Corollary 3.3. *For $w^1 > \cdots > w^m > 1$, $x^{it} \leq \max_{k=i}^{m} \alpha(v^{k/m})$ for $i \in \{2, \ldots, m\}$.*

Now to solve (31), we need to discuss several possible cases based on $v^{k/m}$'s.

(i) If $v^{m-1/m} < v^{m/m}$, then following a similar argument as for Proposition 3.1, we can show that $x^{(m-1)t} = x^{mt} \leq \alpha(v^{m/m})$. Note that we do not necessarily have $\alpha(v^{m-1/m}) \leq x^{(m-1)t} = x^{mt}$ here, since $x^{(m-1)t}$ and x^{mt} must be bounded above by $x^{(m-2)t}$, which could be smaller than $\alpha(v^{m-1/m})$. Knowing that $x^{(m-1)t} = x^{mt}$, we substitute x^{m-1} for x^m and further simplify the expression of $E[r]$ as

$$E[r] = q_1 Q(x^1; w^1) + \sum_{k=2}^{m-2} q_k Q(x^k; v^{k/m}) + (q_{m-1} + q_m) Q(x^{m-1}; v^{m-1,m/m}),$$

where

$$v^{m-1/m} < v^{m-1,m/m} := \frac{w^m - \sum_{j=1}^{m-2} q_j w^{m-2}}{q_{m-1} + q_m} < v^{m/m}.$$

Therefore, we can continue the discussion as for the $(m-1)$-type case and the solution can be obtained inductively.

(ii) If $1 \geq v^{m-1/m} \geq v^{m/m}$, then obviously $x^{(m-1)t} = x^{mt} = 0$. So, the expected profit can be written as $E[r] = q_1 Q(x^1; w^1) + \sum_{k=2}^{m-2} q_k Q(x^k; v^{k/m})$, and we can proceed as for the $(m-2)$-type case.

(iii) If $v^{m-1/m} > 1 \geq v^{m/m}$, then $x^{mt} = 0$ and $E[r] = q_1 Q(x^1; w^1) + \sum_{k=2}^{m-1} q_k Q(x^k; v^{k/m})$, which can be dealt with as for the $(m-1)$-type case.

(iv) If $v^{m-1/m} \geq v^{m/m} > 1$, then we need to further look at $v^{m-2/m}$.

(iv a) If $v^{m-2/m} < v^{m-1/m}$, then following a similar argument as for (i), $x^{(m-2)t} = x^{(m-1)t} \leq \alpha(v^{m-1/m})$ and

$$E[r] = q_1 Q(x^1; w^1) + \sum_{k=2}^{m-3} q_k Q(x^k; v^{k/m}) \\ + (q_{m-2} + q_{m-1}) Q(x^{m-2}; v^{m-2,m-1/m}) + q_m Q(x^m; v^m/m),$$

where

$$v^{m-2/m} < v^{m-2,m-1/m} := \frac{\sum_{j=1}^{m-1} q_j w^{m-1} - \sum_{j=1}^{m-3} q_j w^{m-3}}{q_{m-2} + q_{m-1}} < v^{m-1/m},$$

which can be dealt with as for the $(m-1)$-type case.

(iv b) If $v^{m-2/m} \geq v^{m-1/m}$, then we further look at $v^{m-3/m}$, $v^{m-4/m}$ and so on when necessary: if at some point we have $v^{i-1/m} < v^{i/m}$ for some i, then $x^{(i-1)t} = x^{it} \leq \alpha(v^{i/m})$ and the m-type case can be reduced to the $(m-1)$-type case similarly as discussed in (iv a); otherwise, $w^1 > v^{2/m} \geq \cdots \geq v^{m/m} > 1$, and as a result, $x^{1t} = \alpha(w^1)$ and $x^{it} = \alpha(v^{i/m})$ for $i \in \{2, \ldots, m\}$.

In conclusion, the solution to (31), $\{x^{it}\}_{i=1}^m$, can be obtained inductively as above and the optimal expected profit is $E[r^t] = \sum_{j=1}^m q_j Q(x^{jt}; \alpha^{-1}(x^{jt}))$. Then the complete team solution, specifically $\{r^{it}\}_{i=1}^m$, is finally computed by (32) and (33).

3.4 Solution of the Incentive-Design Problem

3.4.1 ϵ-Incentive Controllability

Now that the team solution $\{x^{it}, r^{it})\}_{i=1}^m$ has been obtained, we study incentive controllability. If the problem is incentive controllable, then there must exist a γ^t-function solving (17) through (19). Then for this γ^t, however, from (21) and (22) in Theorem 3.1, we know that a user with the type w^m must be indifferent between (x^{mt}, r^{mt}) and $(0, 0)$, while a user with the type w^i, $i \in \{1, \ldots, m-1\}$, must be indifferent between (x^{it}, r^{it}) and $(x^{(i+1)t}, r^{(i+1)t})$. Thus, for any $i \in \{1, \ldots, m\}$, x^{it} cannot be the unique solution to (17). Therefore, the problem is not incentive controllable.

Next, we will show that the problem is ϵ-incentive controllable. Similarly as for the complete information case, here for incomplete information we first find a solution to the problem defined by (17) through (19), except that the solution to (17) may not uniquely be x^{it} for $i \in \{1, \ldots, m\}$. Then we revise this solution, still denoted as γ^t, and obtain an ϵ-team optimal incentive policy $\gamma^{t\epsilon}$, which will be explained in detail in the following section.

3.4.2 Graphical Illustration by Numerical Examples

Before obtaining an inductive method to compute the solution of the incentive-design problem defined by (17) through (19), we first illustrate the idea by numerical examples. Suppose that there are $m = 5$ types with equal probabilities for

Table 1: Team solution for Example 1.

i	1	2	3	4	5
w^i	60	40	30	24	20
$\alpha^{-1}(x^{it})$	60	20	10	6	4
x^{it}	0.8256	0.7078	0.6	0.5	0.4069
r^{it}	12.1782	10.4873	8.8017	7.3655	6.1421
$F_{w^i}(x^{it}; r^{it})$	18.2024	7.4985	2.7984	0.3656	-1
$E[r^t]$	8.9949				

all types. Then from the analysis in 3.3.3, we can write the expected profit as $E[r] = \sum_{j=1}^{5} Q(x^j; v^{j/5})/5$, where $v^{1/5} := w^1$ and $v^{i/5} = i * w^i - (i-1) * w^{i-1}$ for $i \in \{2, 3, 4, 5\}$.

Example 1. Take w^1 to w^5 to be 60, 40, 30, 24, and 20, respectively. Then,

$$v^{1/5} = 60 > v^{2/5} = 20 > v^{3/5} = 10 > v^{4/5} = 6 > v^{5/5} = 4 > 1.$$

From the discussion in (iv b) in 3.3.3, we know immediately that $x^{it} = \alpha(v^{i/5})$ for $i \in \{1, 2, 3, 4, 5\}$. The numerical results are shown in Table 1.

The team solution is also depicted in Figure 3. Actually, for the five dashed curves (parts of them are solid, which will be explained shortly), the least steep one, which goes through (x^{5t}, r^{5t}) and $(0, 0)$, is $F_{w^5}(x; r) = -1$, or equivalently, $r = Q(x; w^5)$. Obviously, a user of the type w^5 gets a net utility less than -1 above this curve, while below this curve, he gets a net utility more than -1. Thus, for (17) to be satisfied, γ^t must be above or on this curve. Similarly, the curve going through (x^{4t}, r^{4t}) and (x^{5t}, r^{5t}) is $F_{w^4}(x; r) = F_{w^4}(x^{5t}; r^{5t})$, or equivalently, $r = r^{5t} + Q(x; w^4) - Q(x^{5t}; w^4)$, and γ^t must be above or on this curve; and so on. In conclusion, γ^t must be above or on the solid curve shown in Figure 3 such that (17) is satisfied for all the five types of users. Of course, it also needs to go through $(0, 0)$ and the (x^{it}, r^{it})'s to meet the requirements of (18) and (19).

Next, to solve the problem that γ^t does not make x^{it} the unique solution to (17) for $i \in \{1, \ldots, m\}$, we revise γ^t as follows to obtain an ϵ-team optimal incentive policy $\gamma^{t\epsilon}$. First, let $\gamma^{t\epsilon}(x^{5t}) = \gamma^t(x^{5t}) - \epsilon^5 = r^{5t} - \epsilon^5$, where $\epsilon^5 > 0$. Then for a user of type w^5, $(x^{5t}, \gamma^{t\epsilon}(x^{5t}))$ becomes the unique optimal choice. Next, let $\gamma^{t\epsilon}(x^{4t}) = \gamma^t(x^{4t}) - \epsilon^4 = r^{4t} - \epsilon^4$, where $\epsilon^4 > \epsilon^5$, such that $(x^{4t}, \gamma^{t\epsilon}(x^{4t}))$ becomes the unique optimal choice for a user of type w^4. On the other hand, to make sure that a user of type w^5 still sticks to $(x^{5t}, \gamma^{t\epsilon}(x^{5t}))$, we need $F_{w^5}(x^{4t}; \gamma^{t\epsilon}(x^{4t})) < F_{w^5}(x^{5t}; \gamma^{t\epsilon}(x^{5t}))$, or equivalently,

$$\begin{aligned} &F_{w^4}(x^{4t}; r^{4t}) - (w^4 - w^5)\log(1 + x^{4t}) + \epsilon^4 \\ &\quad < F_{w^4}(x^{5t}; r^{5t}) - (w^4 - w^5)\log(1 + x^{5t}) + \epsilon^5. \end{aligned}$$

Since $F_{w^4}(x^{4t}; r^{4t}) = F_{w^4}(x^{5t}; r^{5t})$ by (22), the above inequality becomes $\epsilon^4 < \epsilon^5 + (w^4 - w^5)[\log(1 + x^{4t}) - \log(1 + x^{5t})]$. Similarly, for $i \in \{3, 2, 1\}$, let $\gamma^{t\epsilon}(x^{it}) = \gamma^t(x^{it}) - \epsilon^i = r^{it} - \epsilon^i$, where $\epsilon^{i+1} < \epsilon^i < \epsilon^{i+1} + (w^i - w^{i+1})[\log(1 + x^{it}) - \log(1 + x^{(i+1)t})]$. Then by choosing ϵ^i's arbitrarily small subject to the above

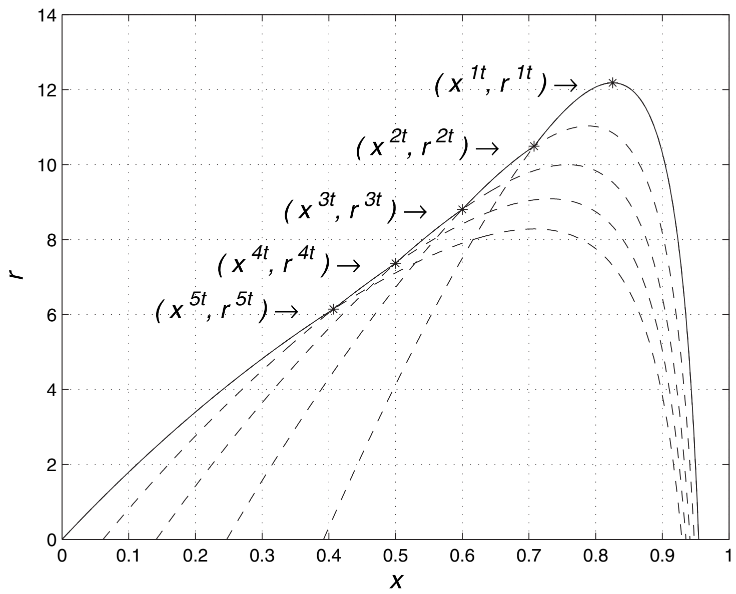


Figure 3: Graphical illustration of incentives for the incomplete information case: Example 1.

constraints, we finally arrive at an ϵ-team optimal incentive policy, $\gamma^{t\epsilon}$, and the problem is ϵ-incentive controllable.

Example 2. Now take w^1 to w^5 to be 70, 40, 33, 25, and 20, respectively. Then, $v^{1/5}$ to $v^{5/5}$ become 70, 10, 19, 1, and 0, respectively. First, $1 \geq v^{4/5} > v^{5/5}$. Then, from the discussion in (ii) in 3.3.3, $x^{4t} = x^{5t} = 0$, and the problem is reduced to the three-type case. Next, observe that $v^{2/5} < v^{3/5}$. Then, as discussed in 3.3.2, $x^{2t} = x^{3t} = \alpha(v^{2,3/5})$, where $v^{2,3/5} = (3w^3 - w^1)/2 = 14.5$, and $x^{1t} = \alpha(w^1)$. The results are listed in Table 2.

Figure 4 shows the team solution for this example. Similarly as for Example 1, the five dashed curves, from the least steep one to the most steep one, represent

Table 2: Team solution for Example 2.

i	1	2	3	4	5
w^i	70	40	33	25	20
$\alpha^{-1}(x^{it})$	70	14.5	14.5	1	0
x^{it}	0.8380	0.6615	0.6615	0	0
r^{it}	18.6491	14.8005	14.8005	0	0
$F_{w^i}(x^{it}; r^{it})$	17.7856	2.5540	-1	-1	-1
$E[r^t]$	9.6500				

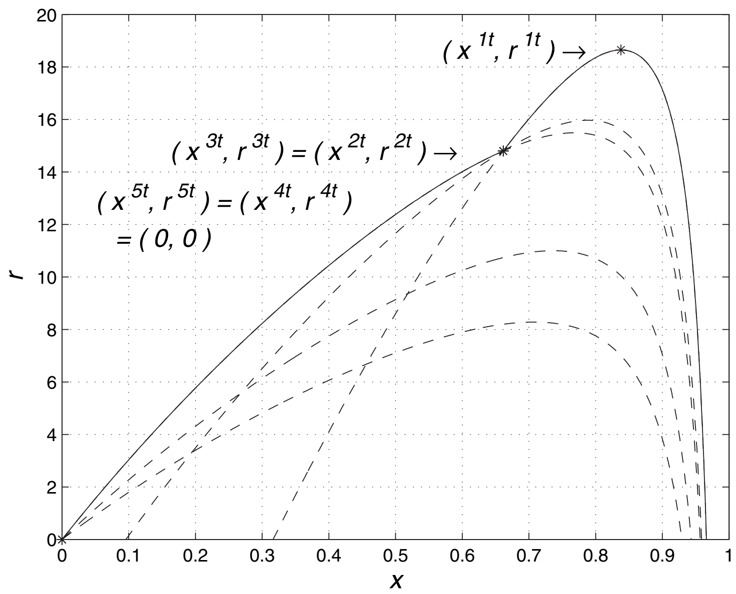


Figure 4: Graphical illustration of incentives for the incomplete information case: Example 2.

$F_{w^5}(x; r) = -1$, and $F_{w^i}(x; r) = F_{w^i}(x^{(i+1)t}; r^{(i+1)t})$ for $i \in \{4, 3, 2, 1\}$, respectively. To satisfy (17), γ^t must be above or on all these curves. Finally, any γ^t that falls above or on the solid curve and goes through (0, 0) and (x^{it}, r^{it})'s solves the incentive-design problem defined by (17) through (19), except that x^{it} is not the unique solution to (17) for $i \in \{1, 2, 3, 4, 5\}$.

The process to obtain the ϵ-team optimal incentive policy, $\gamma^{t\epsilon}$, is slightly different from that for Example 1. First, since $(x^{5t}, r^{5t}) = (x^{4t}, r^{4t}) = (0, 0)$, $\epsilon^5 = \epsilon^4 = 0$ such that $\gamma^{t\epsilon}(x^{5t}) = \gamma^t(x^{5t}) = 0$ and $\gamma^{t\epsilon}(x^{4t}) = \gamma^t(x^{4t}) = 0$. Next, since $(x^{3t}, r^{3t}) = (x^{2t}, r^{2t})$, we have $\gamma^{t\epsilon}(x^{3t}) = \gamma^t(x^{3t}) - \epsilon^3 = r^{3t} - \epsilon^3$ and $\gamma^{t\epsilon}(x^{2t}) = \gamma^t(x^{2t}) - \epsilon^2 = r^{2t} - \epsilon^2$, where $\epsilon^3 = \epsilon^2 > 0$, such that $(x^{3t}, \gamma^{t\epsilon}(x^{3t}))$, or equivalently, $(x^{2t}, \gamma^{t\epsilon}(x^{2t}))$, becomes the unique optimal choice for a user of type w^3 or w^2. On the other hand, to make sure that a user of type w^5 or w^4 still sticks to (0, 0), we need $\epsilon^3 < (w^3 - w^4)\log(1 + x^{3t})$. Finally, $\gamma^{t\epsilon}(x^{1t}) = \gamma^t(x^{1t}) - \epsilon^1 = r^{1t} - \epsilon^1$, where $\epsilon^2 < \epsilon^1 < \epsilon^2 + (w^1 - w^2)[\log(1 + x^{1t}) - \log(1 + x^{2t})]$.

3.4.3 ϵ-Team Optimal Incentive Policies

Now we can conclude with an inductive method to determine γ^t and $\gamma^{t\epsilon}$ given the team solution, $\{(x^{it}, r^{it})\}_{i=1}^m$.

Step 1: Start with $j = m$. If $x^{jt} = 0$, decrease j by 1 repeatedly until $x^{jt} > 0$. Then we know that for $i \in \{m, \dots, j+1\}$, $x^{it} = 0$ and thus $\epsilon^i = 0$ such that $\gamma^{t\epsilon}(0) = \gamma^t(0) = 0$.

Step 2: Next, let $k = j$ and repeatedly decrease j by 1 until $x^{jt} > x^{kt}$. Then for $i \in \{k, \dots, j+1\}$, the (x^{it}, r^{it})'s are all the same with $x^{it} > 0$ and $r^{it} > 0$. To make x^{it} the unique optimal choice for a user of type w^i, we need to choose $\gamma^{t\epsilon}(x) = \gamma^t(x) \geq Q(x; w^k)$ for $0 < x < x^{kt}$, and $\gamma^{t\epsilon}(x^{it}) = \gamma(x^{it}) - \epsilon^i = r^{it} - \epsilon^i$, where $\epsilon^k = \cdots = \epsilon^{j+1} > 0$. On the other hand, to make sure that a user of a lower type still sticks to (0, 0), we need $\epsilon^k < (\epsilon^k - \epsilon^{k+1}) \log(1 + x^{kt})$.

Step 3: Again, let $k = j$ and repeatedly decrease j by 1 until $x^{jt} > x^{kt}$ (for the sake of the iteration here, define $x^{0t} := 1 > x^{1t}$). For the same reason as in Step 2, for $i \in \{k, \dots, j+1\}$, we need to choose $\gamma^{t\epsilon}(x) = \gamma^t(x) \geq r^{(k+1)t} + Q(x; w^k) - Q(x^{(k+1)t}; w^k)$ for $x^{(k+1)t} < x < x^{kt}$, and $\gamma^{t\epsilon}(x^{it}) = \gamma(x^{it}) - \epsilon^i = r^{it} - \epsilon^i$, where $\epsilon^{k+1} < \epsilon^k = \cdots = \epsilon^{j+1} < \epsilon^{k+1} + (w^k - w^{k+1})[\log(1 + x^{kt}) - \log(1 + x^{(k+1)t})]$.

Step 4: Repeat Step 3 until $j = 0$. Finally, for $x^{1t} < x < 1$, let $\gamma^{t\epsilon}(x) = \gamma^t(x) \geq r^{1t} + Q(x; w^1) - Q(x^{1t}; w^1)$.

In conclusion, the resulting incentive policy, $\gamma^{t\epsilon}$, is ϵ-team optimal, which makes $(x^{it}, \gamma^{t\epsilon}(x^{it}))$ the unique optimal choice for a user with the type $w^i, i \in \{1, \dots, m\}$, and generates for the ISP an expected profit of $E[r^t] - \sum_{j=1}^m \epsilon^j$, where $\sum_{j=1}^m \epsilon^j$ can be made arbitrarily small. Therefore, the problem defined by (17) through (19) is ϵ-incentive controllable.

4 Conclusions and Extensions

This chapter has studied Internet pricing from the perspective of the ISPs by introducing nonlinear pricing policies. We have verified ϵ-incentive controllability for the single ISP-single user case, and obtained ϵ-team optimal incentive policies, under complete information and incomplete information, respectively.

Currently, we are working on an extension to the multiple user case. As we have mentioned, a Stackelberg game model was formulated in [4] and generalized in [19], where the ISP was restricted to linear pricing, and the Stackelberg game solution was also obtained, Now we introduce nonlinear pricing policies for the ISP such that his charge for one user is a function of the user's flow. Then the incentive-design problem is to find incentive functions, one for each user, that induce them to choose the team solution, where the ISP can achieve the optimal profit. Similarly as for the single user case, we consider two games: the complete information game, for which the true types of the users are commonly known, as well as the incomplete information game, for which each user's type is private information for himself. For these two games, respectively, we study incentive controllability of the incentive-design problem and, if it is incentive controllable, we obtain the team optimal incentive policies. The results will be reported in future publications.

In the future, we also plan to extend this work to the multiple ISPs case. For this case, besides playing a reverse Stackelberg game with the users, the ISPs must also play a noncooperative game with each other such that the resulting incentives compose a Nash equilibrium. The results will provide useful guidelines for the ISPs on the deployment of their pricing policies.

Acknowledgments

This research was supported by NSF grant ANI03-12976ITR. This work is an expanded and revised version of the paper with the same title, which was presented at the Eleventh International Symposium on Dynamic Games and Applications, 2004.

The authors are grateful to the anonymous reviewers for useful comments, and to Asuman E. Ozdaglar for helpful discussions and suggestions during her visit to the Coordinated Science Laboratory, University of Illinois.

Appendix: Proof of Proposition 2.1

We want to show that $F_w(x; \gamma(x))$, where $\gamma(x) = a_1 x + a_2 x^2$, and a_1 and a_2 are as given in (13), achieves a unique global maximum in (0, 1) at $x = x^t$. We first check the local optimality at this point, and then verify that it attains a unique global maximum in (0, 1).

Local Maximum

From (7) and (10), $w = (1 + x^t)/(1 - x^t)^2$. Also, from (13), $a_2 = -r^t/(x^t)^2$, where $r^t = Q(x^t; w)$. Thus,

$$\begin{aligned}
&\frac{d^2}{dx^2} F_w(x; \gamma(x))\bigg|_{x=x^t} \\
&\quad = \frac{\partial^2 Q(x; w)}{\partial x^2}\bigg|_{x=x^t} - \frac{d^2\gamma(x)}{dx^2}\bigg|_{x=x^t} \\
&\quad = -\frac{w}{(1+x^t)^2} - \frac{2}{(1-x^t)^3} - 2a_2 \\
&\quad = -\frac{2x^t + (x^t)^2 - (x^t)^3 + 2(x^t)^4 - 2(1+x^t)^2(1-x^t)\log(1+x^t)}{(x^t)^2(1+x^t)(1-x^t)^3}.
\end{aligned}$$

Showing that the above quantity is negative for $0 < x^t < 1$ is equivalent to proving

$$2x^t + (x^t)^2 - (x^t)^3 + 2(x^t)^4 > 2(1+x^t)^2(1-x^t)\log(1+x^t).$$

This is true because (i) if $x^t = 0$, both sides attain 0 and have zero first-order derivatives, and (ii) for $0 < x^t < 1$, the first-order derivative of the left-hand side is larger than that of the right-hand side, or equivalently,

$$2x^t - (x^t)^2 + 8(x^t)^3 > [2 - 4x^t - 6(x^t)^2]\log(1 + x^t),$$

which holds since the left-hand side is always positive, while for $\frac{1}{3} \le x^t < 1$, the right-hand side is nonpositive, and for $0 < x^t < \frac{1}{3}$,

$$\begin{aligned} 2x^t - (x^t)^2 + 8(x^t)^3 &> [2 - 4x^t - 6(x^t)^2]x^t \\ &> [2 - 4x^t - 6(x^t)^2]\log(1 + x^t). \end{aligned}$$

Therefore, we have proved

$$\frac{d^2}{dx^2}F_w(x; \gamma(x))\bigg|_{x=x^t} < 0, \tag{A.1}$$

which, combined with (12), guarantees that $F_w(x; \gamma(x))$ achieves a local maximum at $x = x^t$.

Unique Global Maximum

Since $F_w(x; r) = Q(x; w) - 1 - r$ and $F_w(x^t; r^t) = -1$, to show that $F_w(x; \gamma(x))$ achieves a unique global maximum in (0, 1) at $x = x^t$, we just need to prove that for $0 < x \neq x^t < 1$, $Q(x; w) < \gamma(x)$, or equivalently,

$$\eta(x; w) := \frac{Q(x; w)}{x} < \xi(x) := \frac{\gamma(x)}{x} = a_1 + a_2 x.$$

We prove this in several steps.

First, since $Q(x; w)$ and $\gamma(x)$ coincide at (x^t, r^t), and so do their first-order derivatives by (12), we have

$$\begin{aligned} \frac{\partial \eta(x; w)}{\partial x}\bigg|_{x=x^t} &= \left[\frac{1}{x}\frac{\partial Q(x; w)}{\partial x} - \frac{1}{x^2}Q(x; w)\right]\bigg|_{x=x^t} \\ &= \left[\frac{1}{x}\frac{d\gamma(x)}{dx} - \frac{1}{x^2}\gamma(x)\right]\bigg|_{x=x^t} \\ &= \frac{d\xi(x)}{dx}\bigg|_{x=x^t}. \end{aligned} \tag{A.2}$$

Hence, $\eta(x; w)$ and $\xi(x)$ are tangent at $(x^t, r^t/x^t)$.

Furthermore, since $Q(x;w)$ has a smaller second-order derivative than $\gamma(x)$ does at (x^t, r^t) by (A.1),

$$\begin{aligned}\left.\frac{\partial^2\eta(x;w)}{\partial x^2}\right|_{x=x^t} &= \left.\left[\frac{1}{x}\frac{\partial^2 Q(x;w)}{\partial x^2} - \frac{2}{x^2}\frac{\partial Q(x;w)}{\partial x} + \frac{2}{x^3}Q(x;w)\right]\right|_{x=x^t} \\ &< \left.\left[\frac{1}{x}\frac{d^2\gamma(x)}{dx^2} - \frac{2}{x^2}\frac{d\gamma(x)}{dx} + \frac{2}{x^3}\gamma(x)\right]\right|_{x=x^t} \\ &= \left.\frac{d^2\xi(x)}{dx^2}\right|_{x=x^t} = 0, \end{aligned} \tag{A.3}$$

which means that $\eta(x;w)$ is strictly concave at $(x^t, r^t/x^t)$.

We can also prove that $\frac{\partial^3\eta(x;w)}{\partial x^3} < 0$ for $0 < x < 1$. To show this, write out the expression for the second-order derivative of $\eta(x;w)$:

$$\frac{\partial^2\eta(x;w)}{\partial x^2} = w\left[-\frac{1}{x^2} + \frac{1}{(1+x)^2} - \frac{1}{x^2(1+x)} + \frac{2\log(1+x)}{x^3}\right] - \frac{2}{(1-x)^3}.$$

Note that $-2/(1-x)^3$ is strictly decreasing. Thus, it is sufficient to show that the quantity in the brackets has a negative first-order derivative, i.e.,

$$\frac{2x(1+x)^3 + 4x(1+x)^2 + x^2(1+x) - 2x^4 - 6(1+x)^3\log(1+x)}{x^4(1+x)^3} < 0.$$

Now we compare $2x(1+x)^3+4x(1+x)^2+x^2(1+x)$ with $2x^4+6(1+x)^3\log(1+x)$. It can be easily verified that the third-order derivative of the former function is always less than that of the latter one. Hence, the second-order derivatives, which assume the same value at $x = 0$, follow the same order. So do the first-order derivatives. Finally, we have $2x(1+x)^3 + 4x(1+x)^2 + x^2(1+x) < 2x^4 + 6(1+x)^3\log(1+x)$, and it is concluded that

$$\frac{\partial^3\eta(x;w)}{\partial x^3} < 0, \quad 0 < x < 1. \tag{A.4}$$

From (A.3) and (A.4), we know that as x increases from 0 to 1, $\eta(x;w)$ is either concave, or first convex and then concave. Note that $\xi(x)$ is a linear function, tangent with $\eta(x;w)$ at $(x^t, r^t/x^t)$ by (A.2). Thus, if $\eta(x;w)$ is concave for $0 < x < 1$, we always have $\eta(x;w) < \xi(x)$ except at $x = x^t$; if $\eta(x;w)$ is first convex and then concave, the same conclusion holds if and only if $\lim_{x\to 0^+}\eta(x;w) = w - 1 < \xi(0) = a_1$. Recall that from (7), (10), and (13),

$$w - 1 = \frac{1+x^t}{(1-x^t)^2} - 1 \quad \text{and} \quad a_1 = \frac{2r^t}{x^t} = \frac{2}{x^t}Q\Big(x^t; \frac{1+x^t}{(1-x^t)^2}\Big).$$

Consider them as functions of x^t. Then it can be easily verified that they assume the same value at 0^+, and the first-order derivative of $w-1$ is less than that of a_1, which implies $w - 1 < a_1$. In conclusion, $\eta(x;w) < \xi(x)$ for $0 < x \neq x^t < 1$, and hence $F_w(x;\gamma(x))$ has a unique global maximum in $(0,1)$ at $x = x^t$.

References

[1] D. Acemoglu and A. Ozdaglar, Competition and Efficiency in Congested Markets, Mathematics of Operations Research, 2006 (to appear).

[2] E. Anderson, F. Kelly and R. Steinberg, A Contract and Balancing Mechanism for Sharing Capacity in a Communication Network, Management Science, 52(1), pp. 39–53, January 2006.

[3] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory.* SIAM Series in Classics in Applied Mathematics, Philadelphia, PA, 1999.

[4] T. Başar and R. Srikant, Revenue-Maximizing Pricing and Capacity Expansion in a Many-Users Regime, *Proc. IEEE INFOCOM 2002,* pp. 1556–1563, New York, June 23–27, 2002.

[5] T. Başar and R. Srikant, A Stackelberg Network Game with a Large Number of Followers, *J. Optimization Theory and Applications,* 115(3), pp. 479–490, Dec. 2002.

[6] C. Courcoubetis and R. Weber, *Pricing Communication Networks: Economics, Technology and Modelling.* John Wiley & Sons Ltd., Queensland, Australia 2003.

[7] D. Fudenberg and J. Tirole, *Game Theory.* The MIT Press, Cambridge, MA, 1991.

[8] R. Gibbens, R. Mason and R. Steinberg, Internet Service Classes Under Competition, *IEEE J. on Selected Areas in Communications,* Vol. 18, No. 12, pp. 2490–2498, Dec. 2000.

[9] L. He and J. Walrand, Pricing Differentiated Internet Services, *Proc. IEEE INFOCOM 2005,* 1, pp. 195–204, 2005.

[10] L. He and J. Walrand, Pricing and Revenue Sharing Strategies for Internet Service Providers, *Proc. IEEE INFOCOM 2005,* 1, pp. 205–217, 2005.

[11] M. Honig and K. Steiglitz, Usage-Based Pricing of Packet Data Generated by a Heterogeneous User Population, *Proc. IEEE INFOCOM 1995,* 2, pp. 867–874, 1995.

[12] R. Jain and P. Varaiya, An Efficient Incentive-Compatible Combinatorial Market Mechanism, *Proc. Allerton Conference,* Sept. 2004.

[13] R. Johari and J. Tsitsiklis, Efficiency Loss in a Network Resource Allocation Game, *Mathematics of Operations Research*, 29(3), pp. 407–435, 2004.

[14] F. Kelly, A. Maulloo and D. Tan, Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability, *J. of the Operational Research Society*, 49, pp. 237–252, 1998.

[15] R. Maheswaran and T. Başar, Social Welfare of Selfish Agents: Motivating Efficiency for Divisible Resources, *Proc. IEEE CDC 2004,* pp. 1550–1555, The Bahamas, Dec. 14–17, 2004.

[16] P. Marbach, Analysis of a Static Pricing Scheme for Priority Services, *IEEE/ACM Trans. on Networking,* Vol. 12, No. 2, pp. 312–325, Apr. 2004.

[17] L. W. McKnight and J. P. Bailey, Eds, *Internet Economics.* The MIT Press, Cambridge, MA, 1997.

[18] S. Sanghavi and B. Hajek, Optimal Allocation of a Divisible Good to Stragegic Buyers, *Proc. IEEE CDC 2004,* pp. 2748–2753, The Bahamas, Dec. 14–17, 2004.

[19] H.-X. Shen and T. Başar, Differentiated Internet Pricing Using a Hierarchical Network Game Model, *Proc. IEEE ACC 2004,* pp. 2322–2327, Boston, MA, June 30–July 2, 2004.

[20] H.-X. Shen and T. Başar, Network Game with a Probabilistic Description of User Types, *Proc. IEEE CDC 2004,* pp. 4225–4230, The Bahamas, Dec. 14–17, 2004.

[21] J. Shu and P. Varaiya, Pricing Network Services, *Proc. IEEE INFOCOM 2003,* 2, pp. 1221–1230, Apr. 2003.

[22] A. Sundararajan, Nonlinear Pricing of Information Goods, *Management Science,* 50(12), pp. 1660–1673, Dec. 2004.

[23] P. Thomas, D. Teneketzis and J. Mackie-Mason, A Market-Based Approach to Optimal Resource Allocation in Integrated-Services Connection-Oriented Networks, *Operations Research,* 50(4), pp. 603–616, July 2002.

[24] S. Yang and B. Hajek, Revenue and Stability of a Mechanism for Efficient Allocation of a Divisible Good, Preprint, Mar. 2005.

Incentive Stackelberg Strategies for a Dynamic Game on Terrorism

Doris A. Behrens
Institute for Mathematical Methods in Economics, ORDYS
Vienna University of Technology
1040 Vienna, Austria
and
Department of Economics
Alps Adriatic University of Klagenfurt
9020 Klagenfurt, Austria
doris.behrens@aon.at

Jonathan P. Caulkins
H. John Heinz III School of Public Policy and Management
Carnegie Mellon University
Pittsburgh, PA 15213, USA
and
Qatar Campus, P.O. Box 24866
Doha, Qatar
caulkins@andrew.cmn.edu

Gustav Feichtinger
Institute for Mathematical Methods in Economics, ORDYS
Vienna University of Technology
1040 Vienna, Austria
or@eos.tuwien.ac.at

Gernot Tragler
Institute for Mathematical Methods in Economics, ORDYS
Vienna University of Technology
1040 Vienna, Austria
tragler@eos.tuwien.ac.at

Abstract

This paper presents a dynamic game model of international terrorism. The time horizon is finite, about the size of one presidency, or infinite. Quantitative and qualitative analyses of incentive Stackelberg strategies for both decision-makers of the game ("the West" and "International Terror Organization") allow statements about the possibilities and limitations of terror control interventions. Recurrent behavior is excluded with monotonic variation in the frequency of terror attacks whose direction depends on when the terror organization launches its terror war. Even optimal pacing of terror control operations does not greatly alter the equilibrium of the infinite horizon game, but outcomes from the West's perspective can be greatly improved if the game is only "played" for brief periods of time and if certain parameters can be influenced, notably those pertaining to the terror organization's ability to recruit replacements.

1 Introduction

International terrorism is by any measure a principal concern of policy makers and a complex issue. This paper complements and extends the work of Keohane and Zeckhauser [16], Kaplan *et al.* [15], and Caulkins *et al.* [4] among others by developing *incentive Stackelberg strategies for a dynamic "terror game" model*. This model aims to shed light on the question of how best to prosecute the "war on terror," specifically terror control operations intended to preempt or deter terror attacks.

The premise of the model presented here is that terror attacks (such as suicide bombings) are produced by a "stock of terrorists" [16] that represents the strength of a specific adversary (such as Al-Qaeda). Their/its "power" reflect/s some combination of human resources, financial resources, technical sophistication (particularly of weapons and weapons delivery), and sympathy/public support, particularly among the nonterrorist population within which terrorists are embedded.

The modeling is motivated by the September 11th attacks in the US, the March 11th attacks in Madrid, the July 7th attacks in London, and associated events. We will refer to the decision-maker or *player* on the terrorist side as ITO, simply as an abbreviation of some *I*nternational *T*error *O*rganization. ITO's opponents are assumed to include the West and certain Arab regimes variously labeled as "moderate," "secular," "pro-Western," or "authoritarian." For ease of exposition, we will refer to ITO's opponent simply as "the West."

The structure of this chapter is as follows. Section 2 is devoted to a crucial ingredient for the model discussed here, namely the "two-edged effect" of (potentially inflammatory) terror control activities. The "game on terror" is then presented in Section 3. Section 4 includes the derivation of the asymmetric noncooperative strategies for the West and ITO for a finite time horizon. Parameterizing the model allows us to illustrate our results in Section 5, including a thorough sensitivity analysis yielding important insights about possibilities and limitations of terror control policies. Section 6 concludes the paper with an outlook to further research.

2 The Two-Edged Effect of Terror Control Interventions

Each decision-maker in the "game-theoretic" interaction in the field of international terrorism has access to (at least) one instrument for altering the status quo. ITO can choose the rate at which it commits terror attacks. The West can choose the intensity with which it attacks ITO (to prevent terror attacks).[1]

[1]Our analysis focuses on these two interventions. Defensive aspects of homeland security operations are not modeled explicitly. When the West exercises its control, that should be

At times of "war," attacks by both sides reduce the stock of terrorists (as opposed to "peacetime" when neither side is making attacks). Attacks by the West kill, capture, or incapacitate terrorists directly; that is their purpose. ITO suicide attacks also obviously kill ITO members. Even nonsuicide attacks will tend to reduce the size of ITO because every attack exposes the organization to law enforcement risk and its operatives to an increased likelihood of imprisonment and/or death. The more aggressively the West is prosecuting its terror control campaign, the higher is the number of ITO operatives lost per successful attack. Moreover, the higher the intensity of terror control interventions, the more resources that can be devoted to investigating the aftermath of an attack by creating better opportunities for interrogations. These might in turn provide information about other ITO targets and operatives, so that subsequent suicide attackers can be stopped beforehand. So, the higher the West's intensity of terror control measures, the more attacks that ITO has to launch to produce a single successful operation. Hence, it is more costly for ITO (in terms of personnel losses) to attack when the West is in a counter-terror posture than when it is passive.

New terrorists are recruited by existing terrorists. Accordingly, when ITO is small, the growth rate in the current number of terrorists ought to be increasing. Yet, growth is not exponential and unbounded. At some point, growth per current member should slow down because of limits on the number of potential recruits, limits on the capacity of ITO to train and absorb new recruits, etc.

There is, however, one further dynamic that may be important, namely the possibility that aggressive attacks by the West generate resentment among the populations from which ITO seeks recruits. Heymann [14] mentions this possibility and notes (p. 62) that "recruitment to the Irish Republican Army (IRA) increased sharply during some periods of overly vigorous British action against suspects." That is, terror control attacks by the West may have the direct benefit of eliminating current terrorists but the undesirable indirect effect of increasing ITO's recruitment rate [15]. All these flows are depicted in Figure 1.

As far as the authors know, there is little published in the scientific literature concerning how best to model ITO's recruiting practices and abilities.[2] We imagine here that the undesirable indirect effects of what is perceived as excessive and inappropriate terror control attacks increases more than proportionally with the intensity of those attacks. That is, we imagine that in the eyes of a potential terrorist recruit, some degree of counter-attacking is to be expected, but excessive levels stimulate sympathies with ITO. This overproportional increase in sympathy could arise not only from the sheer volume of attacks, but also because of heterogeneity

---

understood to be actions such as invading Afghanistan, using Predator drones with Hellfire missiles to assassinate Al-Qaeda operatives as in Yemen, or freezing assets of organizations with links to Al-Qaeda.

[2]Kaplan *et al.* [15] estimate recruitment, attack, death, and growth rates for a related discrete-time descriptive growth model for the evolution of the strength of terrorist organizations in the West Bank, which may or may not be similar to ITO in this regard.

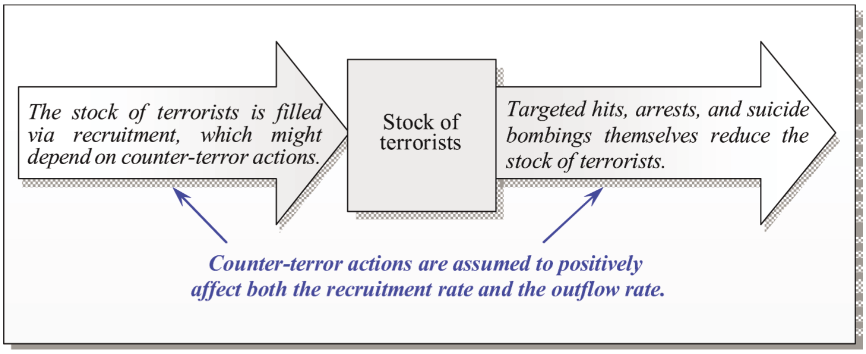


Figure 1: Flow chart of the "dynamics of terror."

in the "specificity" of attack opportunities. The less specific the attack, the more collateral damage that is caused to innocent people. This will cause the rate of sympathy among the population to increase. Because the West prefers high-specificity attacks, these attacks are undertaken whenever possible.

A "high-specificity attack" as the term is used here would refer to an attack on people known to be ITO members who are guilty of specific past or planned attacks. Therefore, a high-specificity attack would not involve significant collateral damage to innocent civilians, mosques, or other holy sites and would minimize violation of an Arab nation's sovereignty. (For example, an ITO operative is arrested in a Western country in the course of planning an attack and with unambiguous physical artifacts proving that a terror attack was being prepared.) A "low-specificity attack" would be bombing a site occupied by both ITO personnel and innocent noncombatants. A historical example would be the Clinton administration's launching cruise missiles against targets in Sudan and Afghanistan. A hypothetical example (that crosses a line not yet crossed) would be a cruise missile attack or other bombing of a mosque or shrine during Ramadan that killed many more bystanders than it did ITO operatives.

As the West expands its terror control efforts, it might be forced to resort to low-specificity tactics once it has exhausted the opportunities for high-specificity attacks. That could contribute to the adverse effect of ITO's recruitment rate increasing more than proportionally with the intensity of those attacks.

3 The Model

3.1 State Dynamics

Translating the considerations presented in Section 2 into equations, the number of terrorists at time t, $x(t)$, evolves according to the equation

$$\dot{x}(t) = f(x(t)) - \mu w(t) - g(v(t))w(t) - \phi v(t) + h(v(t)), x(0) > 0, \qquad (1)$$

$x(t) \geq 0$	number of terrorists at time t, relative to the ITO steady state population in "peacetime" (= state variable),
$v(t) \geq 0$	intensity of the West's terror control activities at time t, relative to its maximum sustainable level (= control variable of "the West"),
$w(t) \geq 0$	number of ITO attacks at time t (= control variable of ITO),
$f(x(t))$	endogenous growth of ITO at time t,
$\mu \geq 0$	average number of terrorists killed or arrested per attack,
$g(v(t))$	number of terrorists lost per terror attack due to terror control efforts $v(t)$,
$\phi \geq 0$	rate at which terror control operations would deplete ITO if the West is on full counter-offensive,
$h(v(t))$	growth of ITO at time t due to hatred caused by collateral damage induced by (low-specificity) terror control activities of the West.

We suppose that during "peacetime" (when neither player is making attacks), ITO would grow to some maximum, "natural" size which we normalize to one without loss of generality. Accordingly, a value of, e.g., $x(t) = 0.5$ corresponds to a situation where the current size of ITO is half its "peacetime" level.

We have no data or experimental opportunities with which we might test and validate alternative functional forms for the model given by equation (1). We proceed with simple linear and quadratic forms for the sake of transparency and tractability, and in recognition that the very structure of the model (e.g., reifying terrorist and terror control forces as single actors and tactics as single controls) implies that it will inevitably be highly stylized relative to the complexity of the true phenomena.

We assume that ITO (= number of terrorists $x(t)$) grows logistically in the absence of controls, i.e.,

$$f(x(t)) = \gamma(1 - x(t))x(t), \tag{2}$$

where γ denotes the endogenous growth rate.[3] ITO's size declines proportionally to the number of attacks at time t as embodied in the first outflow term $\mu w(t)$. The second outflow term originates from the fact that the West's control ($v(t)$) contributes positively to the probability of finding (and arresting) ITO operatives who knew somebody involved in an attack, i.e.,

$$g(v(t)) = \beta v(t). \tag{3}$$

The parameter β in equation (3) denotes the ratio of ITO operatives lost per successful attack when the West is on full counter-offensive ($v(t) = 1$) vs. none at all.

Moreover, the terror control activities ($v(t)$) reduce the number of terrorists directly, represented by the third outflow term $\phi v(t)$. That is, the West can attack terrorists, not only counter-attack when the terrorists strike. It is precisely that sort

[3] Note that the "carrying capacity" corresponds to the *steady state population in "peacetime"* which is normalized to one.

of unprovoked preemptive strike that might be most likely to fuel anger against the West. Consequently, it is plausible to assume that "inappropriate" terror control (seen from a non-West point of view) leads to an increase in the recruitment of new terrorists represented by the inflow term

$$h(v(t)) = \alpha v(t)^2. \tag{4}$$

The parameter α denotes the growth rate of ITO due to hatred caused by collateral damage induced by the West's "low-specificity" terror control attacks. The terms $-\phi v(t)$ and $-\beta v(t)w(t)$ and the term defined by equation (4) account for the two-edged effect of terror control activities discussed in Section 2. This means that there may be a "narrow corridor" of suitable control measures—too little or too much may be suboptimal.

The overall direction of terror control policy is stable and transparent to all. If a country is attacked, it will counter-attack. That was the bedrock of "Mutual Assured Destruction" (MAD) doctrine during the Cold War, and in some sense it is fundamental to the doctrine of essentially any country. Consequently, it is doubtful that Al-Qaeda expected that it could conduct something like the September 11th attacks without eliciting a substantial response. So in some sense the West's behavior is predictable to ITO.[4] Once it has been attacked, the West cannot unilaterally decide to stop fighting. For the sake of credibility (and deterrence) it has to stick to its "counter-strike strategy." This suggests using an open-loop information pattern for the modeled "game on terror," even though ITO has a higher degree of freedom than the West in varying the frequency and aggressiveness of its attack strategy.

Since it is not realistic to suppose that the West could never "learn" and modify its policy rule even in the long run, we examine a finite time horizon problem. Inasmuch as the US is a key driver of the West's strategy vis à vis offensive terror control operations, one might think of this time horizon as akin to the US presidential election cycle (or even shorter periods of time). In 2004 the US re-elected George Bush as president, and he can reasonably be expected to continue the Terror War more or less along the lines he had for the previous three years. The US could instead have elected John Kerry, which might have led to some change in strategy for the US and for the West more broadly. So national elections can be thought of as opportunities for altering strategy, as the Spanish elections did after the Madrid bombing, but between elections one might expect the West's strategy to be fairly stable in the sense that ITO can take it as given when choosing its strategy.

3.2 Objectives

We presume that the West may have up to three policy goals and thus up to three components of its objective function. The West certainly wants to minimize the number of terrorist attacks $w(t)$, but it may also independently seek to minimize the number of terrorists (by minimizing $x(t)$). The principal reason the West might

[4]The Terror War is assumed to be initiated by a September 11th-like attack.

dislike there being a lot of terrorists is their potential to commit attacks, which would be captured in a $w(t)$ term in the objective. However, merely living under the threat of attack can be costly, even if attacks do not occur, because of the cost of maintaining homeland security capacity, psychological stress, and activities (e.g., business investments) avoided because of the uncertainty created by an imminent threat.

The West also has an interest in minimizing the number of terror control strikes (by minimizing $v(t)$). One obvious reason is that terror control efforts cost money, just as does the exercise of almost any control policy. Decision-makers in optimal control formulations (with or without a game-theoretic component) are routinely described as trying to minimize a weighted sum of the magnitude of a problem ($x(t)$ and $w(t)$ in this case) and the costs of efforts intended to reduce that problem ($v(t)$ in this case).

With respect to the objective function we proceed with simple functional forms for the sake of transparency and tractability. In particular, we assume that the objective function components are linear in their respective arguments. Hence, for a finite time horizon $[0, T]$ and a positive discount rate r the West's objective can be written as

$$\max_{v(t)} \left\{ -\int_0^T e^{-rt}\{c_1 x(t) + c_2 w(t) + c_3 v(t)\}dt - e^{-rT} s x(T) \right\}, \tag{5}$$

where $sx(T)$ denotes the salvage value which is proportional to the stock of terrorists at the end of the time horizon.

Relative to the costs imposed by terror attacks, direct budgetary outlays on terror control attacks (as distinct from homeland security and the invasion of Iraq) are fairly modest. So if the West thought the only cost of employing $v(t)$ was budgetary or if the cost of intervention "didn't matter" because the Terror War has to be fought regardless of the efforts it takes, $v(t)$ would be a negligible part of the West's objective.

However, political stability in the Gulf region is of first-order importance to the strategic interests of all oil-importing countries. If overzealous terror control operations stirred up hostility to the West that led to an interruption of Persian Gulf oil shipments akin to the OPEC oil embargo of 1972, there could be ramifications for Western economies far larger than the direct budgetary costs of those terror control operations. Hence, from this larger politically oriented perspective, one could justify making c_3 an important component of the West's overall objective function. Hence, we explore the special case in which $c_3 = 0$ and the more general case in which it is positive.

ITO, on the other side, is interested in becoming strong and powerful (by increasing $x(t)$). Furthermore, it may value deaths of Westerners that are a direct consequence of its performing successful attacks (by increasing $w(t)$). Additionally (or alternatively), ITO might have political objectives aimed at inducing "excessive" counter-attacks by the West (eliciting high values of $v(t)$) as an indirect way of stirring up anti-Western sentiments in the Middle East. For example, these anti-Western moods could smooth the way for changes in local regimes. These considerations lead to the following objective function of ITO for a finite time horizon $[0, T]$ and a positive discount rate r (where we assume that the objective function components are linear in their respective arguments):

$$\max_{w(t)} \left\{ \int_0^T e^{-rt}\{d_1 x(t) + d_2 w(t) + d_3 v(t)\}dt + e^{-rT}\underline{s}x(T) \right\}, \tag{6}$$

where $\underline{s}x(T)$ denotes the salvage value.

Consequently, we investigate an asymmetric noncooperative open-loop differential game for a finite time horizon $[0, T]$ and a positive discount rate r with the West optimally predetermining its strategy at time $t = 0$. The West and ITO aim at satisfying equations (5) and (6), respectively, subject to

$$\begin{aligned} \dot{x}(t) &= \gamma(1 - x(t))x(t) - \mu w(t) - \beta v(t)w(t) - (\phi - \alpha v(t))v(t), \\ & x(0) > 0, \end{aligned} \tag{7}$$

$$x(t),\ v(t),\ w(t) \geq 0. \tag{8}$$

3.3 Solution Concept

The final component of the model specification is the choice of the solution concept (given an open-loop information pattern), which governs how the "game" will be played. We adapt some familiar concepts slightly to fit the present context.

We have already alluded to the first key component of the solution concept. The West announces its policy ahead of time and locks into that policy for the duration of the (finite horizon) game. Note, however, *what* is announced. The West does not commit to a particular intensity of attacks ($v(t)$) irrespective of ITO's actions, $w(t)$. Rather, the West announces a set of contingencies. It says, in effect, "If you—ITO—do not attack us, we will not attack you.[5] However, if you attack us, we will strike back. In particular, if you attack our interests abroad (e.g., bombing embassies) you can expect a limited military response. But if you kill our civilians on our soil, we are willing to commit ground troops." So what the West announces is $v(t)$ as a function of $w(t)$. ITO takes that policy as given, and chooses its own actions, represented by $w(t)$, accordingly.

Hence, our solution concept is related to *incentive strategies* that assume that each player has knowledge of the other's actions, and employs strategies making use of this information. Essentially, in *incentive problems* it is assumed that the strategy of one of the decision-makers is a function of the decisions of the other players. There is, however, a notable difference between the War on Terror and standard contexts to which incentive equilibrium concepts have been applied in

[5]Any time delay is neglected since the ITO attacks are (assumed to be) observed in real time, and the West tries to react to these attacks as soon as possible.

the past,[6] namely that the West has at best a very imperfect understanding of ITO's objectives. One hears claims varying from standard political agendas (get US troops out of the Middle East or overthrow secular pro-Western Arab regimes) to objectives that are "rational" given certain (extreme) religious beliefs (e.g., the heavenly rewards accorded to martyrs) to statements that Al-Qaeda's actions are driven by irrational hatred and are not a means to any particular end.

Ignorance of the opponent's objective has a very practical consequence in the context of dynamic games. If the West does not understand ITO's objective, it cannot know how valuable the stock of terrorists is to ITO. That is, the West can observe ITO's actions $(w(t))$, but it cannot observe ITO's motives and, hence, does not know the shadow price ITO places on the state variable, $x(t)$.

We presume that the West tries to be as rational as it can be in the face of this information deficit in the sense of optimizing over all factors that it knows about, including its own shadow price on an additional terrorist. That is, we do not relegate the West to behaving myopically. Rather, we proceed as if the West were computing a standard incentive equilibrium, but simply omit ITO's shadow price from the West's calculations.

We also omit the West's shadow price from ITO's optimization, but for an entirely different reason. Presumably ITO can observe the objectives and values of the West; democratic societies are rather open in this regard. So ITO may have fine information about the West's shadow prices. However, ITO cannot usefully manipulate those shadow prices.

Before launching a terror campaign, any given international terrorist organization is essentially a nonentity in the strategic planning and political processes of a Western country. In the US for example, defense policy before September 11th was much more influenced by the Cold War, the Vietnam War, and the Gulf War than it was by anything Al-Qaeda might do or say. It was predicated on a wide range of threat scenarios, not just large-scale terrorist attacks on the US homeland. (Indeed, that particular threat scenario received relatively little attention before September 11th and so did not figure prominently in defense policy planning.) This is very different than strategic modeling of interaction among peers, such as duopolies in business or the West and Soviet bloc in the Cold War. Whatever the West was doing before the game begins did not in any meaningful way revolve around ITO, and after the game begins, the West plays out the strategy to which it has pre-committed. Hence, ITO simply receives the West's policy as given, and cannot influence the shape of that policy by its threats or actions.

Hence, a suitable game concept is an *incentive Stackelberg game* with the West acting as the leader and ITO as the follower.[7] In such a setting the leader can

---

[6]Incentive problems are usually treated in the context of cooperative symmetric games. For an intuitive explanation, the mathematical formulation, and interesting applications of incentive equilibria, consult, e.g., the work of Ehtamo and Hämäläinen [6, 7, 8].

[7]The "incentive Stackelberg equilibrium solution concept" used in this chapter mainly differs from the well-known concept of an open-loop, finite time horizon Stackelberg game (see,

"force" the follower to play his global optimum by using a strategy that depends on the follower's action (provided the follower plays open-loop).

4 Analyzing the Terror War

As outlined above, the West determines its strategy $v(t)$ for any potential frequency with which terrorist attacks could be executed (embodied in the variable $w(t)$) yielding $v(w(t))$. It is plausible to assume that the West's terror control policy is performed "in the best possible way," i.e., by maximizing the objective function (5), yielding $v^\star(w(t))$.[8]

Therefore, the West determines the optimal terror control activities $v^\star(w(t))$ by solving an optimal control problem with respect to $v(t)$. Omitting the explicit notion of time from now on, the mathematical analysis of the Terror War starts at the current value Hamiltonian $\mathcal{H}_1$ for the West,

$$\mathcal{H}_1 = -c_1 x - c_2 w - c_3 v + \pi_1(\gamma(1-x)x - \mu w - \beta v w - (\phi - \alpha v)v). \quad (9)$$

The West's shadow price π_1 is always negative along the optimal path since $\frac{\partial \mathcal{H}_1}{\partial x} < 0$ (see Léonard [17]), i.e., ITO adding one more operative is always damaging to the West. Pontryagin's maximum principle (for the general theory see, e.g., Başar and Olsder [1], Dockner *et al.* [5], Feichtinger and Hartl [9], Léonard and Long [19]) determines the evolution of the West's shadow price of an additional terrorist:

$$\dot{\pi}_1 = (r - \gamma + 2\gamma x)\pi_1 + c_1, \;\; \pi_1(T) = -s. \quad (10)$$

Consequently, the damage implied by having one more terrorist (seen from the West's point of view) increases due to a decreasing per capita cost associated with the mere existence of ITO (c_1). Then the maximizing condition for the intensity of the West's terror control interventions (v) is given by

$$\frac{\partial \mathcal{H}_1}{\partial v} = 0 \;\Leftrightarrow\; -c_3 + \pi_1(-\beta w - \phi + 2\alpha v) = 0. \quad (11)$$

We assume that terrorist organizations know what the West will do if it is attacked (according to the policy rule implicitly determined by equation (11)). Therefore,

e.g., Dockner *et al.* [5], or Başar and Olsder [1]) with respect to the presumed behavior of the West. Here the West does not "announce" the way in which it will perform its terror control activities (to ITO), it predetermines its strategy responding to any (possible) number of terror attacks (as opposed to the number of attacks presumably being most favorable for ITO as a rational decision-maker). Hence, the deviation from the "classical" Stackelberg approach is caused by the West's ignorance of the principles or presence of the "rationality" of the ITO player and by ITO's limited ability to influence the West's policy before the beginning of the game.

[8] Considering the state and control constraints ($x, v, w \geq 0$) we perform the "direct adjoining approach" [9, p. 164] which is not explicitly described in this section. The results derived here contain, however, the analysis at the border of the admissible regions as well.

the terrorists plug the West's strategy $v^\star := v^\star(w)$ into their objective function (6), and solve an optimal control problem with respect to w. Thus, the Hamiltonian $\mathcal{H}_2$ for ITO can be determined by

$$\mathcal{H}_2 = d_1x + d_2w + d_3v^\star + \pi_2(\gamma(1-x)x - \mu w - \beta v^\star w - (\phi - \alpha v^\star)v^\star), \tag{12}$$

where π_2 is always positive along the optimal path since $\frac{\partial \mathcal{H}_2}{\partial x} > 0$ (see Léonard [17]), i.e., one more ITO operative is always a benefit for the terrorists. Since the West locks in its policy before the dynamic part of the game commences, and ITO cannot influence the West's policies before launching its terror war, at this point the analysis becomes unilateral dynamic optimization without factoring in the opposing player's shadow price. The terrorists' imputed value of an additional terrorist evolves according to

$$\dot{\pi}_2 = (r - \gamma + 2\gamma x)\pi_2 - d_1, \ \pi_2(T) = \underline{s}. \tag{13}$$

The maximizing condition for the number of ITO attacks (w) is determined by

$$\frac{\partial \mathcal{H}_2}{\partial w} = 0 \ \Leftrightarrow d_2 + d_3\frac{\partial v^\star}{\partial w} - \pi_2\left(\mu + \beta w\frac{\partial v^\star}{\partial w} + \beta v^\star\right) = 0. \tag{14}$$

The following proposition provides essential information about the interior solutions. Note that even though we normalize $v^\star = 1$ to stand for the maximum *long-run sustainable* intensity of terror control activities, *brief* periods of "extra intensity" ($v^\star > 1$) are possible due to the nature and speed of modern armed conflict.

Proposition 4.1. *If $|\pi_1| > c_3/(\beta w + \phi)$ and $\pi_2 < (2\alpha d_2 + \beta d_3)/(2\alpha\mu + \beta\phi)$ hold, the following pair of strategies constitutes an interior noncooperative incentive equilibrium solution at ($v^\star(w), w^\star$):*

$$v^\star(w) = \frac{1}{2\alpha}\left(\frac{c_3}{\pi_1} + \beta w + \phi\right) > 0, \tag{15}$$

$$w^\star = \frac{1}{\beta^2}\left(\frac{2\alpha d_2 + \beta d_3}{\pi_2} - 2\alpha\mu - \beta\phi\right) > 0. \tag{16}$$

The intensity of the West's (interior) terror control interventions evolves nearly proportional to the number of terror attacks, $w^\star$, and almost indirectly proportional to the West's shadow price of an additional terrorist, $\pi_1 < 0$. Therefore, the West's policy also depends on the current stock of terrorists. The West's incentive Stackelberg strategy does not, however, depend on the initial state of the game. Hence, it will be optimal for the Stackelberg leader to use this strategy from any initial state and time.

Moreover, equation (15) tells us that the intensity of terror control interventions is low when the cost of intervention, c_3, is high (since π_1 is negative). On the contrary, if the cost of intervention "doesn't matter" because the Terror War has to be fought regardless of the efforts it takes ($c_3 = 0$), the intensity of terror control interventions is a linear (affine) function of the number of terrorist attacks. In this case we can guarantee that $v^\star(w)$ is always positive along the optimal path, since $w \geq 0$. (See equation (8).) Proposition 4.1 states that $v^\star > 0$, as long as $c_3/|\pi_1| < (\beta w + \phi)$. This condition is computed by adding a Lagrangian-type expression to the West's Hamiltonian. If $c_3/|\pi_1| \geq (\beta w + \phi)$, the necessary conditions for interior solutions are not satisfied and we observe solutions at the border of the admissible region, because the respective Lagrangian multiplier is positive, forcing the intensity of anti-terror interventions to be zero, i.e., $v^\star(w) = 0$.

For ITO the "optimal" number of attacks is determined by equation (16). Attack rates are implicitly driven by the stock of terrorists (via its "per capita value" π_2). This is important, since the number of attacks ($w^\star$) exceeds zero only if the imputed value (π_2) of an additional terrorist (seen from the terrorists' point of view) stays below a certain threshold level determined by $\pi_2 < (2\alpha d_2 + \beta d_3)/(2\alpha\mu + \beta\phi)$.[9] (See Proposition 4.1.) Above this threshold level, a terrorist would be regarded as "too valuable" for the terrorist organization to risk losing during an attack. Thus according to equation (16), one way the West could reduce the number of ITO activities ($w^\star$) is by bringing ITO's valuation of its stock of terrorists up to a high level (especially toward the end of the game). This policy would correspond to altering the parameter $\underline{s}$ (if possible) as outlined in the following proposition.

Proposition 4.2. *If ITO's salvage value coefficient satisfies the condition $\underline{s} \geq (2\alpha d_2 + \beta d_3)/(2\alpha\mu + \beta\phi)$, the West can wipe out successively the number of terror attacks by making them less and less attractive for ITO, until the "war is over" at time $T (\rightarrow w^\star(T) = 0)$. If additionally $s \leq c_2/\phi$ holds for the West's salvage value coefficient, then the West ceases fire as well and its intensity of terror control activities is equal to zero ($\rightarrow v^\star(T) = 0$). That is, the Terror War truly comes to an end at the end of the (finite) time horizon.*

We are most interested in the finite time horizon case for reasons discussed above, but it is not entirely clear what that time horizon should be. An infinite time horizon is in some respects the limiting case as the finite time horizon gets large, so the infinite time horizon case is of interest to the extent that it may in some sense "bound" the behavior of the finite time horizon case.

For $T = \infty$ the Terror War can approach the uniquely determined "long-run equilibrium state" $\hat{x}_1$ *or* $\hat{x}_2$ given by equation (17) *or* by equation (18), respectively.[10] Denoting the two players' costates in equilibrium $\hat{x}_i$ $(i = 1, 2)$ by

[9]The condition $\pi_2 < (2\alpha d_2 + \beta d_3)/(2\alpha\mu + \beta\phi)$ is computed by adding a Lagrangian-type expression to ITO's Hamiltonian. If $\pi_2 \geq (2\alpha d_2 + \beta d_3)/(2\alpha\mu + \beta\phi)$ the Lagrangian multiplier is positive, forcing $w^\star$ to be zero.

[10]Note that $\hat{x}_1$ and $\hat{x}_2$ (as implicitly determined by equations (17) and (18)) never simultaneously satisfy the state constraint $x(t) \geq 0$.

$\hat{\pi}_{1i}$ and $\hat{\pi}_{2i}$, respectively, we get

$$\hat{X} = \begin{pmatrix} \hat{x}_1 \\ \hat{\pi}_{11} \\ \hat{\pi}_{21} \end{pmatrix} = \begin{pmatrix} \frac{\gamma(r\beta^2-\Phi)+(r-\gamma)\Xi}{2\gamma(r\beta^2-\Xi)} \\ -\frac{c_1(r\beta^2-\Xi)}{r^2\beta^2-\gamma\Phi} \\ \frac{d_1(r\beta^2-\Xi)}{r^2\beta^2-\gamma\Phi} \end{pmatrix} \tag{17}$$

and

$$\hat{Y} = \begin{pmatrix} \hat{x}_2 \\ \hat{\pi}_{12} \\ \hat{\pi}_{22} \end{pmatrix} = \begin{pmatrix} \frac{\gamma(r\beta^2-\Phi)-(r-\gamma)\Xi}{2\gamma(r\beta^2+\Xi)} \\ -\frac{c_1(r\beta^2+\Xi)}{r^2\beta^2-\gamma\Phi} \\ \frac{d_1(r\beta^2+\Xi)}{r^2\beta^2-\gamma\Phi} \end{pmatrix}, \tag{18}$$

where

$$\Phi := \beta^2\gamma + 4\mu(\alpha\mu + \beta\phi) > 0 \tag{19}$$

$$\Psi := \alpha\beta^2 d_1^2 + \gamma(2\alpha d_2 + \beta d_3)^2 > 0 \tag{20}$$

$$\Xi := \frac{1}{\alpha c_1 d_1^2}\sqrt{\alpha d_1^2(r^2\alpha\beta^4 c_1^2 d_1^2 - (r^2\beta^2 - \gamma\Phi)(c_1^2\Psi - \beta^2\gamma c_3^2 d_1^2))}. \tag{21}$$

5 Illustrating the Optimal Deterrence Policy

It is difficult to parameterize this model because of lack of data and because of its high level of abstraction. The Appendix provides rationales that we hope yield parameter values that are of the right order of magnitudes (given by Table 1),

Table 1: Base case parameter values.

	Interpretation of the parameter	**Value**
γ	Endogenous growth rate of terrorist organization	1.5
μ	Average number of terrorists killed or arrested	1
α	Growth rate of terrorist organization due to hatred caused by collateral damage induced by the West's terror control attacks	6
β	Ratio of ITO operatives lost per successful attack when the West is on full counter-offensive vs. none at all	10
ϕ	Rate at which terror control would deplete terrorists for $v = 1$	0.5
r	Discount rate	0.05
c_1	Relative cost to the West of a stock of terrorists	1
c_2	Relative cost to the West of terrorist attacks	20
c_3	Relative cost to the West of conducting terror control operations	0
s	Relative cost to the West of a stock of terrorists at time T	10
d_1	Relative value to ITO of maintaining a stock of terrorists	1
d_2	Relative value to ITO of attacking Western targets	20
d_3	Relative value to ITO of eliciting overzealous control operations	0
s	Relative value to ITO of maintaining x at the time the war ends	6
T	Length of the time horizon	0.5

and we vary certain parameters to explore the sensitivity of the conclusions with respect to those values.

We discuss the incentive Stackelberg solutions for a finite time horizon in Section 5.1. In particular, we consider the Terror War in the absence of political objectives of ITO ($d_3 = 0$) and in the absence of an explicit "interest" of the West in the cost associated with terror control interventions ($c_3 = 0$). In Section 5.2 we describe the results of the game on terror for the infinite time horizon case and list the outcomes of static comparative (sensitivity) analysis. Time horizons on the order of 2–4 years probably make the most sense, but we do not know the proper value of T with any specificity. Hence, we perform the finite time horizon analysis with a very conservative (small) value of $T = 0.5$ which, together with the infinite time horizon, brackets any plausible value of T that might be of interest. Finally, Section 5.3 briefly considers the implications of including parameters d_3 and c_3 in the analysis, which reflects the political ambitions of ITO.

5.1 Finite Horizon Incentive Stackelberg Strategies for $c_3 = 0$

When the cost of terror control operations is negligible compared to the cost of terror and hence is excluded from the West's objective function ($c_3 = 0$), the West's shadow price of an additional terrorist does not enter the analysis. All that counts is the terrorist organization's shadow price of an additional terrorist (π_2) and π_2's value at the end of the Terror War ($\underline{s}$). (See equation (15).) In particular, if $c_3 = 0$, the intensity of terror control operations is governed by

$$v^\star(w) = \frac{1}{2\alpha}(\beta w + \phi). \tag{22}$$

This can be interpreted as saying that the size of ITO becomes irrelevant for determining the terror control activity and the West responds to the number of attacks only. In particular, we state the following.

Proposition 5.1. *If the cost of terror control activities is irrelevant for the West, the intensity of its interventions is proportional to how effective terror control operations are relative to their adverse effect on terrorist recruitment (α), where the effectiveness includes both general effects (ϕ) and the losses stimulated by responding to specific terror attacks (βw). Moreover, the intensity of terror control interventions $v^\star(w)$ is a linear (affine) function of the number of attacks w.*

Accordingly, $w^\star$ and $v^\star$ evolve "hand in hand" and not "delayed" or antagonistically (where highly intensive terror control actions successfully suppress terror attacks). This "hand in hand"-type behavior is caused by the fact that the West's policy rule immediately responds to changes in the number of terrorist attacks, leaving aside the current size of ITO as a future source of attacks. Proposition 4.2 stated that the Terror War ends if ITO's value of having an additional terrorist at the end of the time horizon exceeds a certain threshold level. For the base case parameter set this threshold level is given by $\underline{s} \geq (2\alpha d_2 + \beta d_3)/(2\alpha\mu + \beta\phi) = 14.1176$.

Figure 2 illustrates for the base case parameter values the influence of increasing how much ITO values having terrorists alive at the end of the time horizon (T) by varying parameter $\underline{s}$ around π_2's equilibrium level. Figure 3 does the same when ITO values more highly maintaining a stock of terrorists throughout the entire period of interaction (specifically, when d_1 increases from 1 to 20). In both cases we start with a stock of terrorists well above its equilibrium value.

What is a reasonable value for parameter $\underline{s}$? Since the salvage value is defined by $\underline{s}x(T) = \pi_2(T)x(T)$, we might use as a "benchmark" a value near the long-run equilibrium value of ITO's relative value of an additional terrorist, i.e., $\underline{s} \cong 7.04477$. Since by definition a finite time horizon implies some present orientation, we round that value down a bit, specifically to $\underline{s} = 6$. As foils we also explore values half as large ($\underline{s} = 3$) and slightly above the threshold ($\underline{s} = 15$). As depicted by Figure 2, for $\underline{s} = 6$ ITO initially exploits its "excess" stock of terrorists by launching the terror war quite aggressively, and the West responds in kind with aggressive counter-terror operations. This high intensity of conflict attrites ITO's stock of terrorists down to slightly below its peacetime levels by $T = 0.5$, by which time terror and counter-terror operations are pursued at about half their original intensity. (That stock is still roughly twice as large as the long-run equilibrium level obtained if the Terror War is pursued for an infinite time horizon.)

If the imputed value of an additional ITO operative at the end of the war is larger than the long-run equilibrium level ($\underline{s} = 15$, dashed lines in Figure 2), ITO (and hence the West too) is much less aggressive, and the initial excess in the stock of terrorists is preserved at least partially through the end of the time horizon. Indeed, as per Proposition 4.2, terrorist attacks cease towards the end of the time horizon, i.e., $w^\star(T = 0.5) = 0$.

Decreasing $\underline{s}$ has the opposite effect. Figure 2 shows that the stock of terrorists declines drastically towards the end of the time horizon for $\underline{s} = 3$ (and the base case parameter values). If ITO does not value having many terrorists around, either during the time horizon (smallish d_1) or at its end (small $\underline{s}$), then it will commit them aggressively to suicide operations (large w). By equation (22) that in turn stimulates aggressive counter-terror efforts (large v).

So as $\underline{s}$ increases, ITO "trades" off numbers of current attacks for increases in its terminal size. One might expect similar effects from increasing parameter d_1, which denotes the relative value to ITO of maintaining a stock of terrorists. Furthermore, increasing d_1, *ceteris paribus*, is likely to reduce the intensity of the "substitution process" described above. These effects are indicated by comparing Figure 2 ($d_1 = 1$) and Figure 3 ($d_1 = 20$). As Figure 3 shows, the absolute magnitude of attacks declines as d_1 increases, *ceteris paribus*.

To summarize these insights, first, the ratio d_1/d_2 determines the structure of the trade-off between the present and future killing capacities. Second, the absolute sizes of d_1 and d_2 determine the extents of these capacities. Third, the smaller/larger d_1 is, the more/less the size of $\underline{s}$ matters.

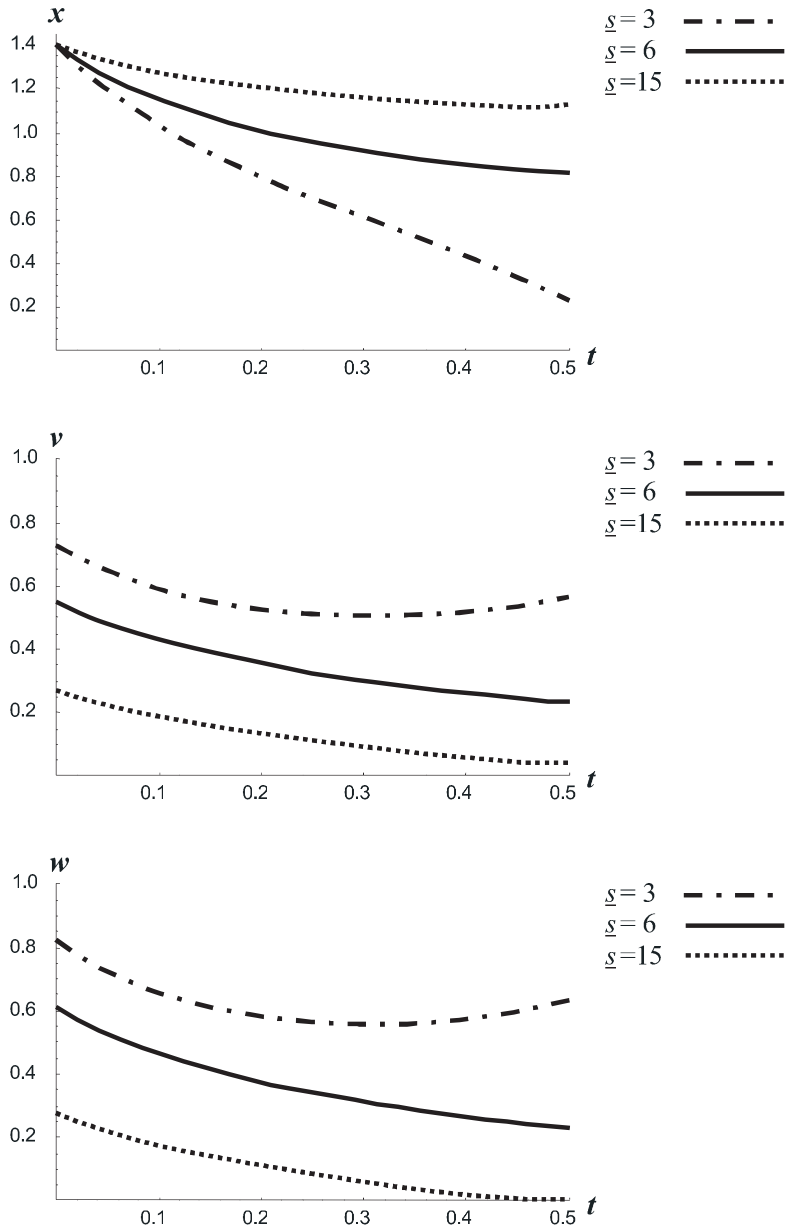


Figure 2: Noncooperative incentive equilibrium paths for the size of ITO (x), the intensity of terror control measures (v), and the number of attacks (w) for the base case parameter set (see Table 1) and different values of $\underline{s}$.

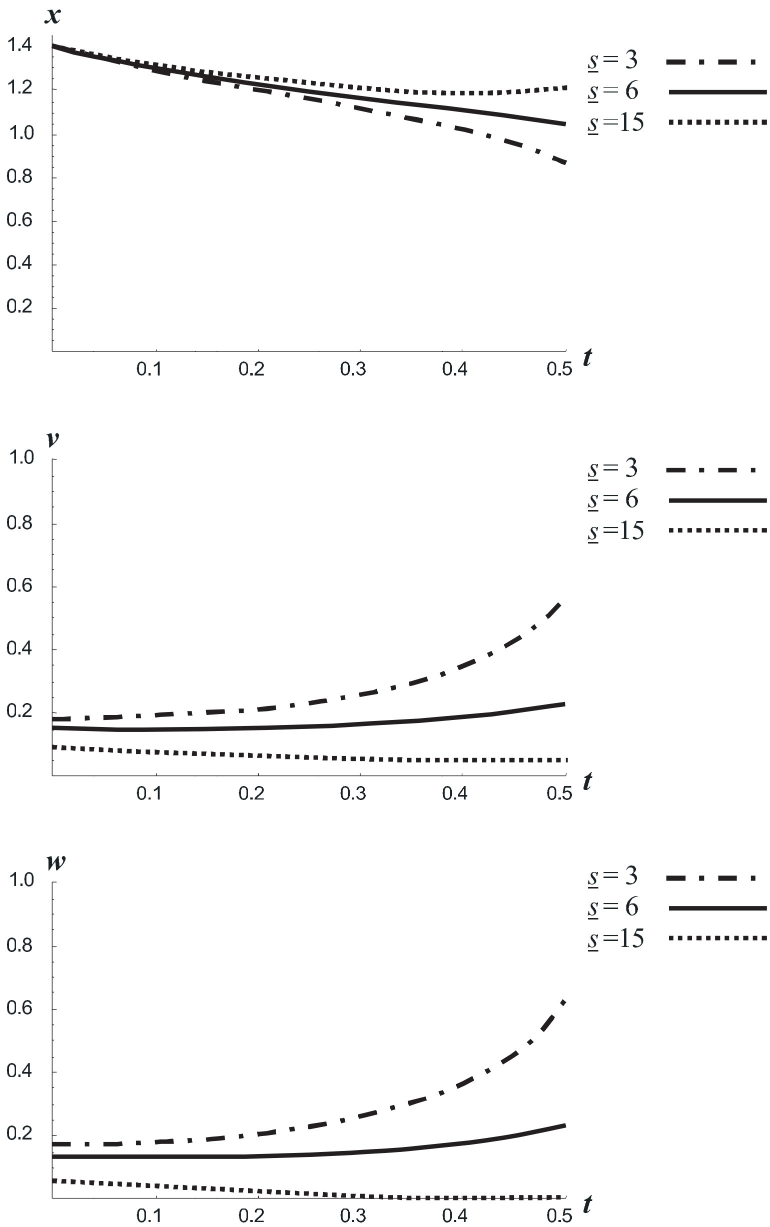
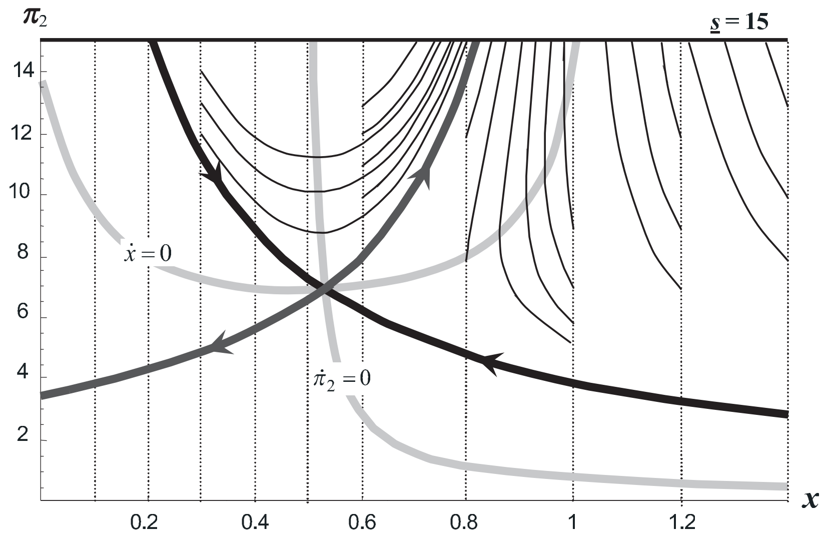


Figure 3: Noncooperative incentive equilibrium paths for the size of ITO (x), the intensity of terror control measures (v), and the number of attacks (w) for the base case parameter set (see Table 1), $d_1 = 20$, and different values of $\underline{s}$.



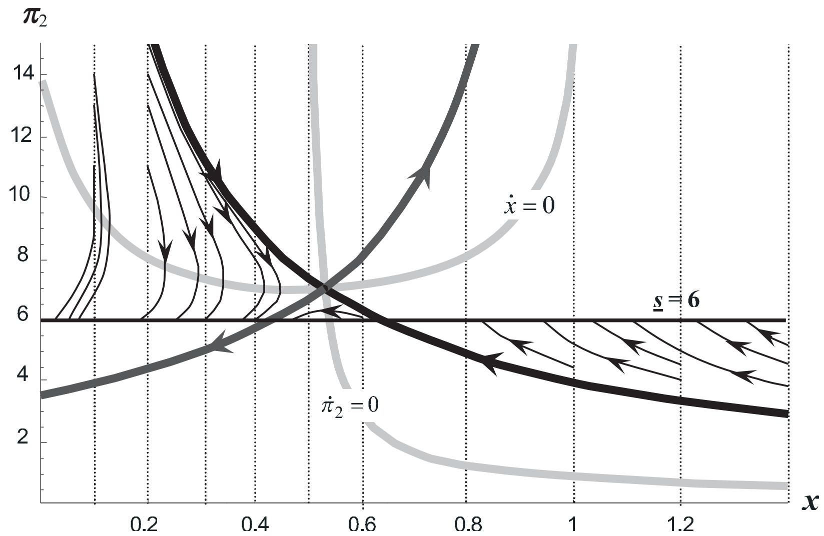


Figure 4: Phase plot in the state-costate space for the base case parameter values (see Table 1). Upper graphic for $\underline{S}(x(T)) = 15x(T)$ and lower graphic for $\underline{S}(x(T)) = 6x(T)$.

So far the analysis has presumed a particular initial size of ITO. We next consider altering the initial values x_0. For this purpose Figure 4 depicts the dynamics of the canonical system in the state-costate space (determined by equations (7) and (13)) for two terminal values $\underline{s}$ chosen above and below the long-run equilibrium level (i.e., for $\underline{s} = 6$ and $\underline{s} = 15$).

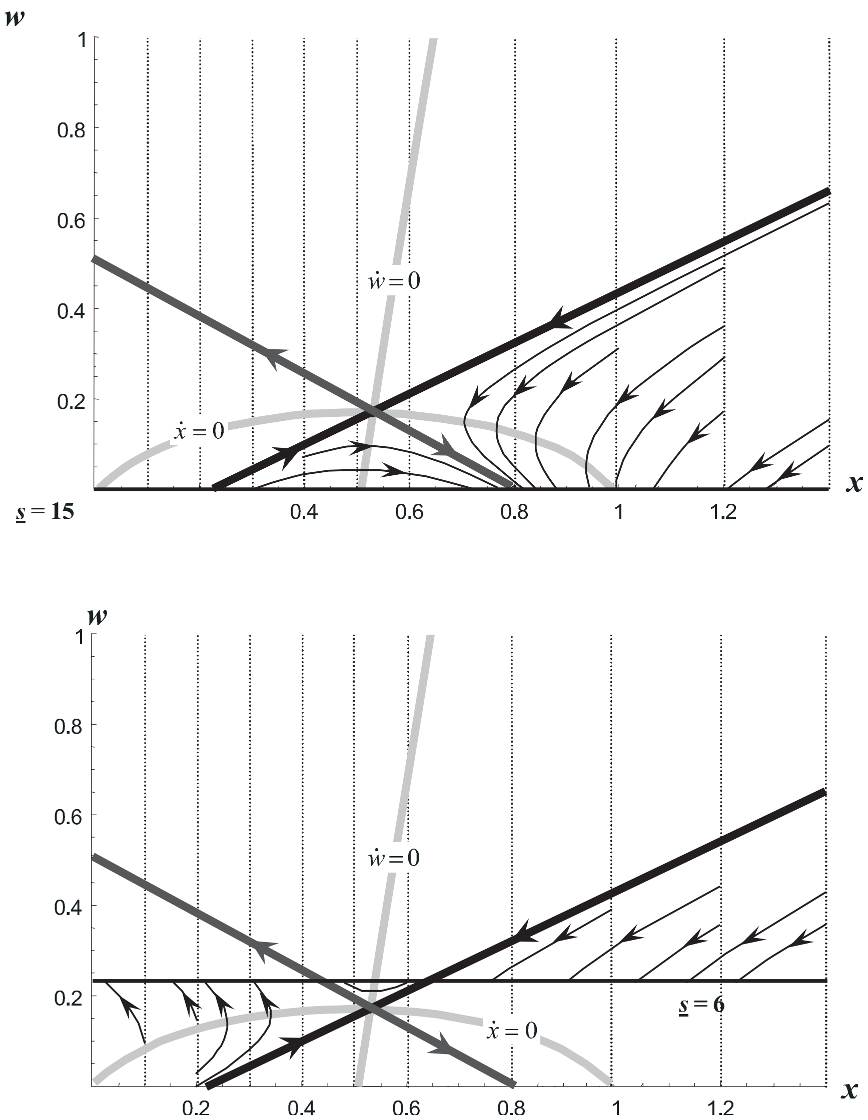


Figure 5: Phase plot in the state-control space for the base case parameter values (see Table 1). Upper graphic for $\underline{S}(x(T)) = 15x(T)$ and lower graphic for $\underline{S}(x(T)) = 6x(T)$.

For base case parameter values and different values of x_0, Figures 4, 5, and 6 depict multiple trajectories where each one is optimal for a different length of the time horizon T.

As shown by Figure 4, if the size of the terrorist organization is initially larger than in peacetime ($x_0 > 1$), the shadow price of an additional terrorist (π_2)

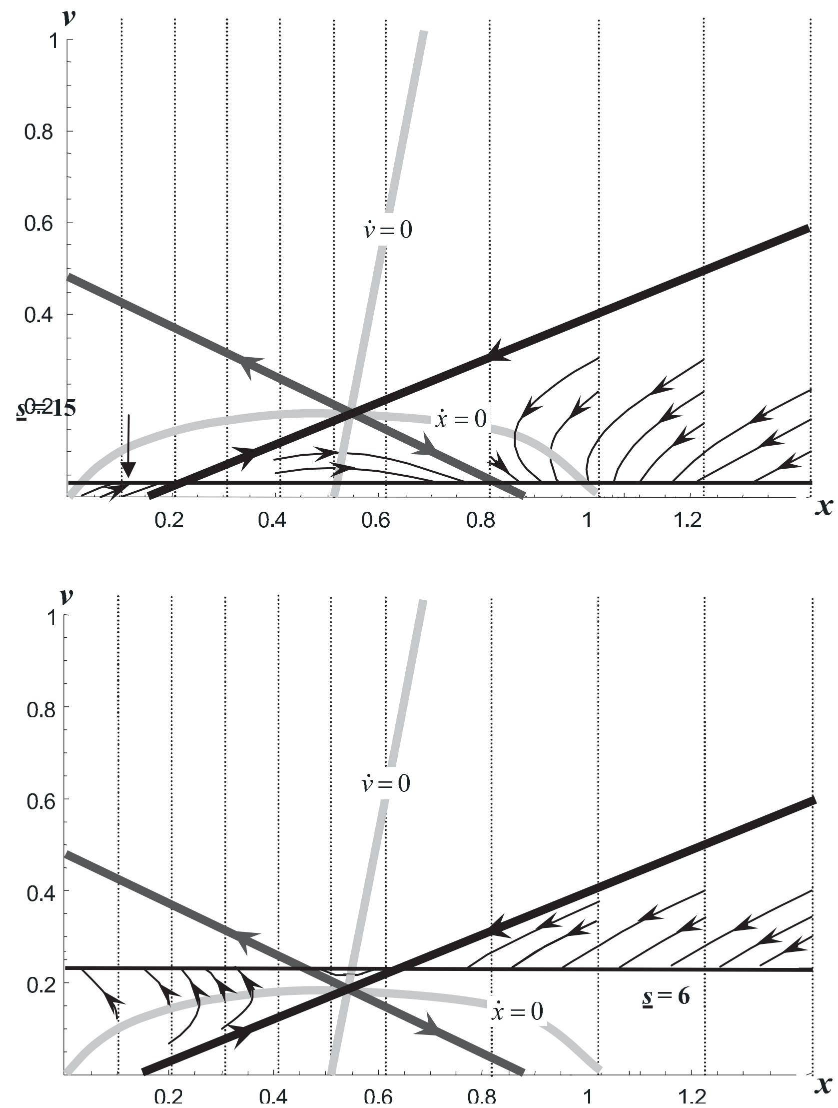


Figure 6: Phase plot in the state-control space for the base case parameter values (see Table 1). Upper graphic for $\underline{S}(x(T)) = 15x(T)$ and lower graphic for $\underline{S}(x(T)) = 6x(T)$).

monotonously increases as the number of terrorists (x) declines. This corresponds to a monotonous decline in the number of attacks (depicted by Figure 5) and, consequently, to a reduction of terror control interventions (shown by Figure 6).

If the size of the terrorist organization is initially smaller than in peacetime ($x_0 < 1$) the size of ITO does not always evolve monotonically over time (see

Figures 4–6). The number of terrorists can first decrease then increase (large $\underline{s}$ and $x_0 > \hat{x}_1$) or first increase and then decrease (small $\underline{s}$ and $x_0 < \hat{x}_1$). As the time horizon tends to infinity, the system dynamics follow the saddle path (represented by the bold arrow) approaching the long-run equilibrium stock of terrorists ($\hat{x}_1$). For shorter time horizons, the optimal path takes the "faster" path farther away from the saddle path.

From Figure 5 we derive that if $\underline{s}$ is close to $\hat{\pi}_2$, then the number of terror attacks w is nearly a linear function of the size of the terrorist organization x, because the stable two-dimensional saddle-point manifold is (almost) linear. This "linear response" (equation (22)) to the frequency of terror attacks seems to quickly squeeze the growth and size of ITO x (as depicted by Figure 6) as well as the number of attacks. However, in fact it is the reduction in v per se that reduces the growth of x (and consequently the level of x and w).

Static comparative analysis providing a brief discussion of the system behavior along the saddle-point manifold is discussed in the next section (to contrast with the results of the "short-run analysis" for the finite time horizon game).

5.2 Infinite Horizon System Behavior for $c_3 = 0$

Qualitatively, the basic behavior of the system is rather straightforward when $T = \infty$. In the long run, the stock of terrorists will converge to a positive number smaller than the peacetime steady state size. (See equations (19), (20).) Wartime operations of both ITO and the West erode the stock of terrorists relative to peacetime levels, but since ITO controls the tempo of conflict, it does not pursue operations in a manner that leads to its own eradication.

One can think of this as vaguely akin to the Vietcong's strategy in Vietnam. There was little possibility of truly destroying the United States. Indeed, there were essentially no attacks on the US mainland during the Vietnam War. So the best the Vietcong could do was to stay in the field fighting at the maximum intensity they could sustain, while using the ability of irregular forces to disappear into the jungles to prevent their forces from ever truly being eradicated.

This insight implies that in the long run the West has no way to bring the war to a close. Its optimal strategy is essentially reactive, with ITO controlling the tempo. However, the West may be able to manipulate certain parameter values as well as its current terror control operations (v), so we explore some parametric sensitivity analysis.

The first set of results of the static comparative analysis is not surprising. The ratio of ITO operatives lost per successful attack when the West is on full counter-offensive vs. none at all (β) and the rate at which terror control operations would deplete ITO operatives for $v = 1$ (ϕ) both calibrate the efficiency of terror control interventions. The more efficient terror control activities are (i.e., the larger parameters β and ϕ are) the fewer terror attacks there will be (though there will be marginally more terror control operations).

Let us next have a look at the sensitivity results for some other parameters. Consider first parameter α, which describes the extent to which terror control operations stimulate recruitment of terrorists. Changes in α have minimal effect on the stock of terrorists for the infinite time horizon case. However, a larger α makes it "more attractive" for terrorists to perform attacks (w high) and less attractive for the West to suppress terror over time. Conversely, if the West could somehow reduce α, that would free it to pursue more aggressive terror control operations (raise v). ITO would respond in ways that protected its terror stock, but that means sharply curtailing the number of terror attacks, which is precisely one of the West's principal objectives.

Similar comments pertain to γ, which governs ITO's rate of recruiting in peacetime and, more generally, recruiting that is not in reaction to terror control attacks by the West. While the number of terrorists x along the saddle-point path hardly changes for different growth rates, γ, the control variables, w and v, do. The root of this behavior is the fact that the additional terrorists caused by an increase in γ quickly flow through state x—being immediately "translated" into a higher number of attacks w. Hence, anything the West can do to make it more difficult for ITO to recruit terrorists translates directly into fewer terror attacks on the West.

To extend the logic above, the single most important factor for determining the size of the terrorist organization, the number of attacks, and the intensity of terror control interventions is the relative value ITO places on attacking the West relative to maintaining its stock of terrorists. If the ratio d_2/d_1 decreases, being strong and powerful ($= x$ large) becomes relatively more important to ITO than actually attacking the West ($= w$ large). If x and w are valued equally (d_2/d_1), the long-run equilibrium stock of terrorists is almost identical to the carrying capacity of terrorists ($x = 1$) and w is extremely low (see Figure 7).

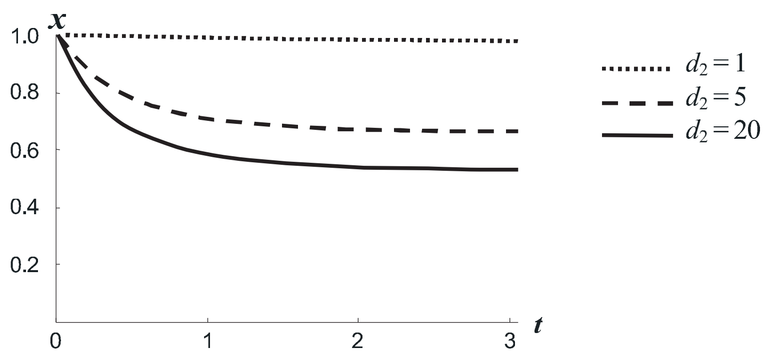


Figure 7: Incentive equilibrium paths for the strength of ITO (x) for the base case parameter values and for different values of the relative importance of bringing threat to the West d_2; $T = \infty$ and $\underline{S}(x(T)) = \underline{s}x(T) = \pi_2^\star(T)x(T)$ as a function of d_2. Solid line indicates the base case scenario.

5.3 Implications of Varying c_3 for the Infinite Horizon Game

So far we have assumed that the cause of ITO was not motivated by any political objective (i.e., $d_3 = 0$). Suppose, however, that ITO derives value by eliciting an over-reaction from the West that turns the "Arab street" against the West (or a pro-Western Arab regime). In this case $d_3 > 0$. If the West is aware of this change in the policy goals, but does not adapt its strategy by increasing c_3 above zero, then the solution is quite insensitive to changes in the value of d_3. In the long run the stock of terrorists x is only slightly smaller for $d_3 > 0$ than for $d_3 = 0$. (For example, increasing d_3 from 0 to 100 increases $\hat{x}$ by only 7.2 %.) Moreover, d_3 doesn't drive up w beyond $d_3/\beta\pi_2$, which could cause additional terror control measures.

This implies an important insight. Even if d_3 changes a lot, the terrorists' actions (w) do not change very much. If we assume the West had no clear independent knowledge about ITO's real objectives, the West could not effectively discern what ITO's objectives are merely by observing its actions. The pattern of attacks that would be optimal for ITO if it were trying to elicit an over-reaction from the West would be very similar to the pattern of attacks it would pursue even if its only goals were more simply to kill Westerners and increase its own size.

6 Conclusions and Suggestions for Further Research

According to this model, whether the intensity of the terror war grows or ebbs in the long run (i.e., for an infinite time horizon) just depends on how much ITO has built up its forces before initiating the conflict. If ITO waited until its initial stock of terrorists exceeded the long-run equilibrium, then the opening rounds of warfare would be the most intense, with the intensity ebbing toward the long-run stable equilibrium (in the case of an infinite time horizon). Conversely, if ITO launched the war while still small, its forces and attack rate would grow toward that equilibrium.

In the short run the situation is different. The intensity of terror attacks and counter-terror efforts might ebb then subsequently increase or vice versa, depending on the relative magnitudes of the initial number of terrorists, the long-run equilibrium number, the length of the time horizon, and the shadow price of a terrorist at the end of the time horizon.

The West has the ability to significantly squeeze size, strength, and attack rate of ITO for brief periods of time—supported or exacerbated by the shadow price of an additional terrorist at the end of the war. Nevertheless, the West has limited capacity to bring about a cessation of hostilities through its dynamic counter-terror efforts.

If one views the West as being additionally able to manipulate certain parameter values (e.g., by increasing the efficiency of terror control activities) as well as its current terror control strategy, the analysis is less depressing for the West. Inasmuch as the war on terror is more realistically represented by a chain of short-term (= finite horizon) games instead of a single finite or infinite time horizon game, parameter changes are possible.

The level of terrorism seems sensitive to two sets of parameters in particular. The first governs recruitment in peacetime and its response to counter-terror operations. In brief, the harder it is for ITO to recruit, the fewer terror attacks it can commit.

More subtly, ITO will also moderate its terror attacks the more it values maintaining a stock of terrorists in the present or at the end of the time horizon, relative to how much value it derives from committing terrorist attacks.

If ITO's objectives primarily revolve around killing Westerners, that might suggest "target hardening" investments in homeland defense. However, ITO's objectives may not be so straightforward. (Indeed, the very ambiguity of Al-Qaeda's objectives figured into the solution concept we applied to this strategic interaction.)

An interpretation of Al-Qaeda's strategic goals is that it has more conventional political objectives [3, 14] so that not only is "war an extension of diplomacy by other means," but also fatwa-endorsed international terror can be an extension of diplomacy by other means. It is not hard to list political objectives whose accomplishment might bring satisfaction to Al-Qaeda's leadership: evacuation of Western military units from Arab lands, notably Saudi Arabia; the destruction of Israel; the closing of international trade, particularly of popular fashions that are seen as destructive to traditional Muslim values; replacing "moderate" or "pro-Western" regimes with theocracies that adhere more strictly to Muslim holy law (Sharia); and deposing the House of Saud.

The view that Al-Qaeda is driven by political objectives is vociferously rejected by some [13] and might seem contradicted by the fact that at least until the eve of the 2004 presidential election Al-Qaeda had not issued demands. The absence of a quid pro quo framing does not, however, necessarily imply that Al-Qaeda's attacks on the West are not a means to some other end. It is at least a theoretical possibility that Al-Qaeda is trying through terror to induce the West into taking actions that it would not or could not agree to explicitly in a negotiation. So the absence of political demands from Al-Qaeda does not imply the absence of political objectives. Elaborating on the actors' objectives and knowledge and understanding of their opponents' objectives is a sensible area for further research. It is not inconsistent with Frey and Luechinger's [10] suggestion that making terrorist attacks less attractive (for the terrorists) can be superior to deterrence.[11] This certainly opens a wide field of further research.

[11] Frey and Luechinger [10] contrast the potential benefits of "benevolence" versus "deterrence" strategies to dissuade terrorists from violent activities. A deterrence strategy, as they use the term, raises the opportunity cost of terrorist activities by defending potential targets, hitting terrorist training centers, infiltrating terrorist groups, and the like. Such deterrence is fundamentally confrontational. A benevolence strategy also raises the opportunity cost of terrorist violence, but it does so by reducing the cost of nonviolent activity, or what Frey and Luechinger call "ordinary activity." Unlike a deterrence strategy, however, a benevolence strategy can improve the well-being of terrorists (if they have more ordinary goods) and the public (if less terrorism occurs). In this way, a benevolence strategy has the potential to achieve a positive-sum outcome. A crucial assumption in this analysis is that a terrorist goal is to have more goods. If the terrorists' objectives were different, e.g., if their goal were simply to harm the public, than Frey and Luechinger's conclusions might not pertain.

Other extensions of the analysis presented here are more methodological. They include issues like an explicit treatment of time delays in the controls and/or recruiting process. Moreover, it is plausible to assume that the model outcomes might change with the solution concept chosen for the analysis. For example, a classic Stackelberg game might lead to different policy recommendations for the West. More generally, one would ideally like to investigate the dynamic game on terror in all the game variants (Nash vs. Stackelberg, open loop vs. closed loop, etc.). Analyzing dynamic games is, however, quite difficult. Therefore, future research might include numerical approximation of the Terror War by using available software (e.g., DYNGAME [19], OPTGAME [2], and Pakes and McGuire's computer algorithm [20]).

Acknowledgments

This research was partly financed by the Austrian National Bank (OeNB) under contract No. 10107 and contract No. 11216. Special thanks to Fouad ElOuardighi, Harri Ehtamo, Dieter Grass, Steffen Jørgensen, Peter M. Kort, Gerhard Sorger, Franz Wirl, Georges Zaccour, and an anonymous referee for their helpful comments.

Appendix: Parameterization

Here we provide rationales that we hope yield parameter values of the right order of magnitudes (given by Table 1). The parameter that is the easiest to interpret is the social discount rate r. It is common in policy analysis to assume a real annual discount rate of 3–5% [13]. We take the high end of this range ($r = 0.05$) inasmuch as war has an urgency that domestic policy does not. One could argue that perhaps ITO's discount rate should be higher than that of the West inasmuch as ITO's objectives are driven more by a single person whereas national security objectives are, ultimately, defending a constitutional system of government whose lifespan exceeds that of any mortal human being. Nevertheless, such changes are unlikely to make a major difference.

The parameter μ is also easy to define. It is the number of terrorists lost per terrorist sent on a successful attack mission. In the absence of terror control measures that figure is basically 1, inasmuch as ITO frequently uses suicide attack tactics. (For these purposes the September 11th attacks could be viewed as 19 attacks on one day since 19 Al-Qaeda operatives died in that one combined attack.)

The larger v is, the more people ITO loses per successful attack. Parameter β governs by how much more, i.e., β denotes the ratio of ITO operatives lost per successful attack when the West is on full counter-offensive status vs. peacetime status. We have no empirical information on this parameter, but we can imagine that ITO might lose an order of magnitude more operatives per successful attack if the West were exerting maximum sustainable terror control effort, i.e., $v = 1$), so in the base case we set $\beta = 10$, but also consider values of 5 and 20.

Parameter γ governs how quickly ITO would grow toward its natural size during peacetime (i.e., when $w = v = 0$). The maximum rate of growth occurs when $x = 0.5$ and that rate is $\gamma/4$. Likewise the time necessary to grow over a certain range (say, from 20% of its final size to 80% of that size) can be readily computed. Values in the range $1 < \gamma < 3$ generate reasonable rates of growth in this regard. They imply growth from 20% to 80% of final, peacetime size in a bit less than 1–3 years. Values of $\gamma = 1.5$ and 2 might be the most sensible point estimates.

Given γ we can estimate α if we view $v = 1$ as the maximum sustainable terror control effort and have some judgment about how such efforts would increase ITO's popularity among its potential recruits. In particular, consider the following hypothetical question: How large would ITO get given the indirect recruitment benefits generated by the maximum sustainable counter-strikes if those counter-strikes effort were not simultaneously eliminating terrorists? That is, imagine a world in which potential recruits were as angry with the US as they would be if the West were attacking Al-Qaeda with all means available, but those attacks were not in fact depleting the population of terrorists.

The state equation for that hypothetical thought experiment would be $\dot{x} = \gamma(1 - x)x + \alpha v^2$. The equilibrium values of x in that case would be $\hat{x}_{1,2} = 0.5(1 \pm \sqrt{1 + 4\hat{v}^2\alpha/\gamma})$. So if one thought the level of anti-West hostility associated with such attacks would double ITO's "peacetime" natural size, then the term $\sqrt{1 + 4\alpha/\gamma} = 3$, so $\alpha/\gamma = 2$. More generally, if one thought it would increase ITO's peacetime natural size by a factor of k, then $\alpha = k(k - 1)\gamma$. There is abundant room for disagreement about how much animosity terror control operations generates and what role such animosity plays in ITO's recruiting success. We are not experts in such matters, but we explore the implications of there being a doubling or tripling of ITO's size, i.e., $\alpha = 2\gamma$ and $\alpha = 6\gamma$.

The last remaining state equation parameter is ϕ, the rate at which terror control operations would deplete ITO operatives if those operations were pursued with maximum sustainable intensity (i.e., if $v = 1$). We take as our base case that $\phi = 0.5$, implying that if ITO started at its steady state peacetime size and never received any new recruits, it would take the West two years of maximum intensity terror control operations to effectively destroy ITO. Once again we are not aware of empirical data with which to estimate this parameter, and we do not have deep domain knowledge, so this value is little more than a guess. We also consider values of $\phi = 0.25$ and $\phi = 1.0$.

Turning to the objective function values, we normalize relative to the value of d_1, the benefit per year to ITO of having its full (peacetime) complement of operatives available to it as a resource. That is, without loss of generality, we set $d_1 = 1$.

The conventional interpretation of ITO's objective is to kill Westerners. For such purposes, the mere existence of a pool of ITO operatives is not of much direct value. They represent potential, but the actual "value" is created by carrying out successful terrorist attacks (w), so in our base case 1 we assume d_2 is much larger than d_1, in particular $d_2 = 20$. One way to think about that value is to consider that an individual ITO operative generates as much value for the organization by blowing himself or herself up in a terrorist attack as he or she would if he/she remained in the organization, in reserve, as a perpetual threat never exercised (since $1/r = 20$ when $r = 0.05$).

In the absence of political motives the ITO player does not value terror control operations per se, so $d_3 = 0$. In the alternative view, those terror control operations are valued, so we also explore the case of $d_3 > 0$. A parameter value of $d_3 = 20$ means that one year of the West exerting maximum sustainable terror control effort is worth as much to ITO's objectives as having its full peacetime complement of operatives in perpetuity. To complete the contrast between the two interpretations of ITO's objectives, when $d_3 = 20$ we set $d_2 = 0$ (i.e., killing Westerners is purely a means to an end, not an end in itself). We use the same relative magnitudes of objective function coefficients for the West.

Without any restriction to generality we assume that the War on Terror is limited to a time period of half a year, $T = 0.5$.

References

[1] Başar, T., Olsder, G.J. (1999) *Dynamic Noncooperative Game Theory*. SIAM Series in Classics in Applied Mathematics, Philadelphia, PA.

[2] Behrens, D.A., Hager, M., Neck, R. (2003) OPTGAME 1.0: A numerical algorithm to determine solutions for two-person difference games. In: R. Neck (ed.), *Modelling and Control of Economic Systems 2002*, Elsevier, Oxford, 47–58.

[3] Bergman, P.L. (2001) *Holy War Inc.: Inside the Secret World of Osama bin Laden*. Free Press, New York.

[4] Caulkins, J.P., Grass, D., Feichtinger, G., Tragler, G. (2005) Optimizing Counter-Terror Operations: Should One Fight Fire with "Fire" or "Water"? *Computers and Operations Research*, to appear .

[5] Dockner, E., Jorgensen, S. Long, N.V., Sorger, G. (2000) *Differential Games in Economics and Management Science*. Cambridge University Press, Cambridge.

[6] Ehtamo, H., Hämäläinen, R.P. (1986) On Affine Incentives for Dynamic Decision Problems. In: Başar, T. (ed.), *Dynamic Games and Applications in Economics*. Springer, Berlin, 46–63.

[7] Ehtamo, H., Hämäläinen, R.P. (1989) Incentive Strategies and Equilibria for Dynamic Games with Delayed Information. *Journal of Optimization Theory and Applications* **63**, 355–370.

[8] Ehtamo, H., Hämäläinen, R.P. (1993) A Cooperative Incentive Equilibrium for a Resource Management Problem. *Journal of Economic Dynamics and Control* **17**, 659–678.

[9] Feichtinger, G., Hartl, R.F. (1986) *Optimale Kontrolle Ökonomischer Prozesse—Anwendungen des Maximumprinzips in den Wirtschaftswissenschaften*. DeGruyter, Berlin.

[10] Frey, B.S., Luechinger, S. (2003) How to Fight Terrorism: Alternatives to Deterrence. *Defence and Peace Economics* **14**(4), 237–249.

[11] Friedman, T.L. (2002) *Longitudes and Attitudes: Exploring the World after September 11*. Farrar, Straus, and Giroux, New York.

[12] Glain, S. (2004) *Mullahs, Merchants, and Militants*. St. Martin's Press, New York.

[13] Gold, M.R., Siegel, J.E., Russell, L.B., Weinstein, M.C. (1996) *Cost-Effectiveness in Health and Medicine*. Oxford University Press, London.

[14] Heymann, P.B. (2003) Dealing with Terrorism after September 11, 2001: An Overview. In: Howitt, A., Pangi, R. (eds.) *Preparing for Domestic Terrorism*. MIT Press, Cambridge, MA, 57–72.

[15] Kaplan, E.H., Mintz, A., Mishal, S., Samban, C. (2005) What Happened to Suicide Bombings in Israel? Insights from a Terror Stock Model. *Studies in Conflict & Terrorism* **28**, 225–235.

[16] Keohane, N.O, Zeckhauser, R. (2003) The Ecology of Terror Defense. *Journal of Risk and Uncertainty* **26**, 201–229.

[17] Léonard, D. (1981) The Signs of the Costate Variables and Sufficiency Conditions in a Class of Optimal Control Problems. *Economic Letters* **8**, 321–325.

[18] Léonard, D., Long, N.V. (1992) *Optimal Control Theory and Static Optimization in Economics*. Cambridge University Press, Cambridge.

[19] McKibbin, W., Sachs, J.D. (1991) *Global Linkages. Macroeconomic Interdependences and Cooperation in the World Economy*. The Brookings Institution, Washington, DC.

[20] Pakes, A., Gowrisankaran, G., McGuire, P. (1993) Implementing the Pakes-McGuire Algorithm for Computing Markov-Perfect Equilibria in GAUSS. *Working paper of the Department of Economics at the Yale University*.

Capital Accumulation, Mergers, and the Ramsey Golden Rule

Roberto Cellini
Department of Economics
University of Catania
Corso Italia 55, 95129 Catania, Italy
cellini@unict.it

Luca Lambertini
Department of Economics
University of Bologna
Strada Maggiore 45, 40125 Bologna, Italy
and
ENCORE, University of Amsterdam
Roetersstraat 11
WB1018 Amsterdam, The Netherlands
lamberti@spbo.unibo.it

Abstract

We take a differential game approach to model the Ramsey growth model from the standpoint of the representative firm. We identify parametric conditions such that the economy cannot reach the Ramsey golden rule, due to the presence of a stable demand-driven equilibrium. This may happen under Cournot and Bertrand behaviour, as well as social planning. We show that a wave of horizontal mergers can indeed drive the economy towards the Ramsey golden rule.

1 Introduction

The neoclassical literature on economic growth is based upon either Ramsey ([12]) or Solow ([15]; see also Swan, [16]). Both approaches describe an economy supplied by perfectly competitive firms that need to accumulate capacity (i.e., physical capital) in order to produce. In particular, the Ramsey model describes the problem of a representative consumer that maximises the discounted flow of utility from consumption, under the constraint given by the dynamics of capital accumulation, thereby attaining a unique saddle point equilibrium which is known as the modified golden rule. In this equilibrium, the marginal productivity of capital is equal to the sum of discounting and depreciation.

Here, we (i) drop the assumption of a perfectly competitive industry, and (ii) model the optimal capacity and output decisions of oligopolistic firms in a dynamic game with capital accumulation *à la* Ramsey, in order to verify the effects of market power on the long-run equilibrium performance of the economic system. We allow for nonlinear market demand functions, borrowing the demand structure from Anderson and Engers ([1], [2]). In this respect, the present model can be seen as an extension of the analysis contained in Cellini and Lambertini ([4]) which is limited to the case of a linear market demand.

First of all, we prove that the Cournot oligopoly produces multiple steady state equilibria. In particular, the economically meaningful solution is dictated either by demand conditions, or by the Ramsey rule of capital accumulation, depending upon (i) the number of firms; (ii) market size; and (iii) the curvature of market demand. Therefore, the economy may not, in general, converge to the modified golden rule in the long run. In particular, while the Ramsey golden rule is independent of such parameters, the market-drive solution is not, and the capital (or capacity) endowment associated with the Cournot solution becomes progressively smaller than the Ramsey one as the number of firms increases.

We also illustrate the relationship holding, in steady state, between demand curvature and capital commitment. This can be summarised in the following terms. Suppose that demand turns from linear into convex. If so, each firm's capital decreases. The opposite happens if we turn a linear demand into a concave one, and it is increasingly so as concavity increases. The intuition is that in the first case market size and firms' optimal outputs decrease, while in the second the opposite is true.

Then, we evaluate the behaviour of a benevolent social planner aiming at welfare maximisation, to find that the socially efficient output is independent of the curvature of market demand. This, in connection with the aforementioned results, entails that the welfare loss associated to the Cournot equilibrium decreases as market size increases.

The last step illustrates a feature of the present model that is inherently linked to its dynamic nature, and sheds some new light on a long-standing issue that has been largely debated using, with very few exceptions,[1] static games. This step consists in investigating the feasibility of horizontal mergers, to find that they have a role in leading the economy towards the modified golden rule. This would ultimately eliminate the aforementioned distortion affecting capital accumulation when firms' behaviour is driven by demand conditions rather than by discounting and the marginal productivity of capital. Since the Cournot capacity comes closer to the Ramsey golden rule as the number of firms decreases, we find that a wave of horizontal mergers can drive the economy towards the modified golden rule. This of course requires that there be a profit incentive for firms to merge, and we identify

[1] See Gowrisankaran ([8]) and Dockner and Gaunersdorfer ([6]), who investigate horizontal mergers in a dynamic oligopoly game with sticky prices introduced in the literature by Simaan and Takayama ([14]).

the conditions ensuring it, in the vein of previous work carried out on this topic in the theory of industrial organization (see Salant *et al.*, [13]; Perry and Porter, [11], *inter alia*). Accordingly, there emerges that the present setup not only offers a new way of treating the issue of horizontal mergers, but also relates this issue to the the efficiency of the capital accumulation process characterising an industry.

The remainder of the chapter is structured as follows. The setup is laid out in Section 2. Section 3 describes the Cournot oligopoly. Section 4 contains the analysis of the social optimum. The analysis of the incentives towards horizontal mergers is performed in Section 5. The concluding remarks are given in Section 6.

2 The Model

We examine the Ramsey model from the standpoint of firms rather than the representative consumer. To do so, we assume that the industry producing the final good does not necessarily operate under perfectly competitive conditions, and we will examine perfect competition as a special case. Accordingly, instead of illustrating the problem of constrained utility maximisation, we rely upon a demand structure which is specified as in Anderson and Engers ([1], [2]) to characterise the dynamic profit-maximising behaviour of a population of symmetric firms. The market is served by n firms selling a homogeneous product over time $t \in [0, \infty)$. The number of firms is constant over time,[2] and the market demand function is defined as follows:

$$Q(t) = A - (p(t))^{\alpha}, \ \alpha > 0. \tag{1}$$

This function is always downward sloping, and can be either convex ($\alpha \in (0, 1)$) or concave ($\alpha > 1$). Firms are quantity-setters, the inverse demand function being

$$p(t) = (A - Q(t))^{\frac{1}{\alpha}}, \tag{2}$$

where $Q(t) = \sum_{i=1}^{n} q_i(t)$, and $q_i(t)$ is the individual output of firm i at time t. Production requires physical capital k,[3] accumulating over time to create capacity. At any t, the output level is $y_i(t) = f(k_i(t))$, with $f' \equiv \frac{\partial f(k_i(t))}{\partial k_i(t)} > 0$ and $f'' \equiv \frac{\partial^2 f(k_i(t))}{\partial k_i(t)^2} < 0$.

A reasonable assumption is that $q_i(t) \leq y_i(t)$, that is, the level of sales is at most equal to the quantity produced. As in Ramsey ([12]), we assume that unsold output $\widetilde{q}_i(t) \equiv f(k_i(t)) - q_i(t)$ can be costlessly transformed into physical capital

[2]One can assume that each consumer in the economy owns one firm. If so, then the output level chosen to maximise quantity corresponds to the level of consumption that each individual would choose to maximise utility, given the market price of the consumption good.
[3]As usual, we assume that technology is homogeneous of degree one in labour and capital, and investigate the problem in per capita terms.

and reintroduced into the production process, yielding accumulation of capacity according to the following process:

$$\frac{dk_i(t)}{dt} = f(k_i(t)) - q_i(t) - \delta k_i(t), \tag{3}$$

where $\delta \in [0, 1]$ denotes the rate of depreciation of capital. Therefore, the cost of capital is represented only by the opportunity cost of intertemporal relocation of unsold output.

In order to further simplify the analysis, suppose that unit variable cost is constant and equal to zero. Firm i's instantaneous profits i are

$$\pi_i(t) = p(t)q_i(t). \tag{4}$$

Firm i maximizes the discounted flow of its profits:

$$J_i = \int_0^\infty e^{-\rho t}\pi_i(t)dt \tag{5}$$

under the constraint (3) imposed by the dynamics of the state variable $k_i(t)$, and initial conditions $k_i(0) \equiv k_{i0}$ for all i. Note that the state variable does not enter directly the objective function.

For future reference, we first outline the features of the demand function (2) in terms of the elasticity of demand w.r.t. price, $\varepsilon_{Q,p}$. The price elasticity of demand can be written as follows:

$$|\varepsilon_{Q,p}| = -\frac{\partial Q(\alpha)}{\partial p(Q(\alpha))} \cdot \frac{p(Q(\alpha))}{Q(\alpha)} = \frac{\alpha p^\alpha}{A - p^\alpha}. \tag{6}$$

3 The Cournot Equilibrium

In solving the quantity-setting game between profit-seeking agents, we shall focus upon a single firm. The relevant objective function of firm i is

$$J_i = \int_0^\infty e^{-\rho t}\, q_i(t) \cdot [A - q_i(t) - Q_{-i}(t)]^{\frac{1}{\alpha}} dt, \tag{7}$$

where $Q_{-i}(t) = \sum_{j \neq i} q_j(t)$ is the total output of the $n-1$ rivals of firm i at time t. The function (7) must be maximised w.r.t. $q_i(t)$, under (3). The corresponding Hamiltonian function is

$$\begin{aligned}\mathcal{H}(t) = e^{-\rho t} \cdot \Big\{ & q_i(t) \cdot [A - q_i(t) - Q_{-i}(t)]^{\frac{1}{\alpha}} \\ & + \lambda_{ii}(t)[f(k_i(t)) - q_i(t) - \delta k_i(t)] \\ & +_{j\neq i}\lambda_{ij}(t)[f(k_j(t)) - q_j(t) - \delta k_j(t)]\Big\},\end{aligned} \tag{8}$$

where $\lambda_{ij}(t) = \mu_{ij}(t)e^{\rho t}$, and $\mu_{ij}(t)$ is the co-state variable associated to $k_j(t)$.

From an informational point of view, the characterisation of feedback strategies seems more satisfying. However, given the structure of the game at hand, this task is impossible when one insists on analytically derived equilibria. Therefore, we characterise the open-loop solution of the game.

The necessary conditions for the optimum are[4]

$$\frac{\partial \mathcal{H}_i(t)}{\partial q_i(t)} = \left[A - q_i(t) - \sum_{j \neq i} q_j(t) \right]^{\frac{1}{\alpha}} - \left\{ q_i(t) \cdot \frac{[A - q_i(t) - \sum_{j \neq i} q_j(t)]^{\frac{1-\alpha}{\alpha}}}{\alpha} + \lambda_{ii}(t) \right\} = 0; \tag{9}$$

$$-\frac{\partial \mathcal{H}_i(t)}{\partial k_i(t)} = \frac{\partial \mu_{ii}(t)}{\partial t} \Leftrightarrow \frac{\partial \lambda_{ii}(t)}{\partial t} = \left[\rho + \delta - f'(k_i(t))\right] \lambda_{ii}(t); \tag{10}$$

$$-\frac{\partial \mathcal{H}_i(t)}{\partial k_j(t)} = \frac{\partial \mu_{ij}(t)}{\partial t} \Leftrightarrow \frac{\partial \lambda_{ij}(t)}{\partial t} = [\rho + \delta - f'(k_j(t))]\lambda_{ij}(t), \tag{11}$$

along with the vector of initial conditions $\mathbf{k}(0) = \mathbf{k}_0$ and the transversality conditions:

$$\lim_{t \to \infty} \mu_i(t) \cdot k_i(t) = \lim_{t \to \infty} e^{-\rho t} \cdot \lambda_i(t) \cdot k_i(t) = 0. \tag{12}$$

Note that $\lambda_{ij}(t)$, $j \neq i$, does not appear in (9); moreover, equation (11) is a separable differential equation admitting the solution $\lambda_{ij}(t) = 0$ at all t. Accordingly, we can set $\lambda_{ij}(t) = 0$ for all $i \neq j$ at any t, and $\lambda_{ii}(t) = \lambda_i(t)$, and confine our attention to a single co-state equation for each firm. The necessary conditions for a path to be optimal become

$$\frac{\partial \mathcal{H}_i(t)}{\partial q_i(t)} = \left[A - q_i(t) - \sum_{j \neq i} q_j(t) \right]^{\frac{1}{\alpha}} - \left\{ q_i(t) \cdot \frac{\left[A - q_i(t) - \sum_{j \neq i} q_j(t)\right]^{\frac{1-\alpha}{\alpha}}}{\alpha} + \lambda_i(t) \right\} = 0; \tag{13}$$

$$-\frac{\partial \mathcal{H}_i(t)}{\partial k_i(t)} = \frac{\partial \mu_i(t)}{\partial t} \Rightarrow \frac{\partial \lambda_i(t)}{\partial t} = [\rho + \delta - f'(k_i(t))]\lambda_i(t). \tag{14}$$

[4]Second-order conditions for concavity are not shown explicitly. However, it can be easily shown that the Hamiltonian satisfies Arrow's ([3]) sufficiency condition. In writing the first-order conditions, the indication of exponential discounting is omitted for brevity.

Condition (13) implicitly defines the reaction function of firm i to the rivals' output decisions. We rewrite (13) as follows:

$$\left[A - q_i(t) - \sum_{j\neq i} q_j(t)\right]^{\frac{1}{\alpha}} - \left\{\frac{q_i(t)}{\alpha} \cdot \frac{\left[A - q_i(t) - \sum_{j\neq i} q_j(t)\right]^{\frac{1}{\alpha}}}{A - q_i(t) - \sum_{j\neq i} q_j(t)} + \lambda_i(t)\right\} = 0, \tag{15}$$

that is,

$$\left[A - q_i(t) - \sum_{j\neq i} q_j(t)\right]^{\frac{1}{\alpha}} \cdot \left[1 - \frac{q_i(t)}{\alpha\left(A - q_i(t) - \sum_{j\neq i} q_j(t)\right)}\right] - \lambda_i(t) = 0. \tag{16}$$

In order to simplify calculations and to obtain an analytical solution, we adopt the following assumption, based on firms' *ex ante* symmetry:

$$\sum_{j\neq i} q_j(t) = (n-1)q_i(t). \tag{17}$$

Due to symmetry, in the remainder we drop the indication of the identity of the firm and rewrite the first-order condition as follows:

$$[A - nq_i(t)]^{\frac{1}{\alpha}} \cdot \left\{\frac{\alpha[A - nq_i(t)] - q_i(t)}{\alpha[A - nq_i(t)]}\right\} - \lambda_i(t) = 0, \tag{18}$$

where we can write $[A - nq_i(t)]^{\frac{1}{\alpha}} = p(t)$, with $q \leq A/n$. Therefore,

$$p(t) \cdot \left\{\frac{\alpha[A - nq_i(t)] - q_i(t)}{\alpha[A - nq_i(t)]}\right\} - \lambda_i(t) = 0, \tag{19}$$

that is,

$$\frac{p(t)\alpha[A - nq_i(t)] - p(t)q_i(t) - \alpha[A - nq_i(t)]\lambda_i(t)}{\alpha[A - nq_i(t)]} = 0, \tag{20}$$

from which we get

$$A[p(t) - \lambda_i(t)]\alpha - q_i(t)\{p(t) + n\alpha[p(t) - \lambda_i(t)]\} = 0. \tag{21}$$

Then, we obtain the symmetric per-firm output:

$$q^*(t) = \frac{A[p(t) - \lambda(t)]\alpha}{(1+n)p(t) - n\lambda(t)\alpha}, \tag{22}$$

which can be rewritten in several equivalent ways, e.g.,

$$\lambda(t) = p(t) - \frac{q^*(t)(p(t))^{1-\alpha}}{\alpha}. \tag{23}$$

The preceding discussion, in particular (22)–(23), produces the following result, which needs no further proof.

Lemma 3.1. *In equilibrium, the following necessarily holds:*

$$\frac{\partial q^*(t)}{\partial t} = 0 \Rightarrow \frac{\partial p(t)}{\partial t} = 0 \Rightarrow \frac{\partial \lambda(t)}{\partial t} = 0$$

and

$$\frac{\partial p(t)}{\partial t} = 0 \Rightarrow \frac{\partial q^*(t)}{\partial t} = 0 \Rightarrow \frac{\partial \lambda(t)}{\partial t} = 0.$$

Equation (23) can be differentiated w.r.t. time to obtain

$$\frac{\partial \lambda(t)}{\partial t} = 0 \Rightarrow \frac{\partial p(t)}{\partial t} - \frac{(p(t))^{1-\alpha}}{\alpha} \cdot \frac{\partial q^*(t)}{\partial t} - \frac{(1-\alpha)q^*(t)(p(t))^{-\alpha}}{\alpha} \cdot \frac{\partial p(t)}{\partial t} = 0. \tag{24}$$

Using the symmetry assumption (17) and differentiating the direct demand function (1) w.r.t. time, we get

$$\frac{\partial q^*(t)}{\partial t} = -\frac{\alpha(p(t))^{\alpha-1}}{n} \cdot \frac{\partial p(t)}{\partial t}, \tag{25}$$

which can be plugged into (24) to yield

$$\frac{\partial \lambda(t)}{\partial t} = 0 \Rightarrow \frac{\partial p(t)}{\partial t}\left(1 + \frac{1}{n}\right) - \frac{(1-\alpha)q^*(t)(p(t))^{-\alpha}}{\alpha} \cdot \frac{\partial p(t)}{\partial t} = 0, \tag{26}$$

from which we derive the following.

Lemma 3.2. *The condition* $\frac{\partial \lambda_i(t)}{\partial t} = 0$ *is satisfied when*

$$\textit{either } \frac{\partial p(t)}{\partial t} = 0 \quad \textit{or} \quad q^*(t) = \frac{A\alpha(n+1)}{n(\alpha n+1)}$$

or

$$\left\{ \frac{\partial p(t)}{\partial t} = 0 \quad \textit{and} \quad q^*(t) = \frac{A\alpha(n+1)}{n(\alpha n+1)} \right\},$$

provided $A \neq nq^*(t)$. *The solution* $q^*(t) = \frac{A\alpha(n+1)}{n(\alpha n+1)}$ *is relevant only for* $\alpha \in (0, 1)$.

However, the condition $q^*(t) = \frac{A\alpha(n+1)}{n(\alpha n+1)}$ is relevant only for $\alpha \in (0, 1)$; as $q^*(t) = \frac{A\alpha(n+1)}{n(\alpha n+1)} > \frac{A}{n}$ for $\alpha > 1$, the condition would imply $p(t) < 0$. See also below.

Now we rewrite (13) as follows:

$$q^*(t) = \alpha[p(t) - \lambda(t)](p(t))^{\alpha-1}. \tag{27}$$

This expression can be differentiated w.r.t. time:

$$\frac{dq^*(t)}{dt} = \alpha \left\{ \left[\frac{dp(t)}{dt} - \frac{d\lambda(t)}{dt} \right] (p(t))^{\alpha-1} + (\alpha - 1)(p(t))^{\alpha-2}[p(t) - \lambda(t)] \frac{dp(t)}{dt} \right\} \tag{28}$$

or

$$\frac{1}{\alpha} \cdot \frac{dq^*(t)}{dt} = (p(t))^{\alpha-1} \left\{ \frac{p(t) + (\alpha - 1)[p(t) - \lambda(t)]}{p(t)} \right\} \frac{dp(t)}{dt} - \frac{d\lambda(t)}{dt} (p(t))^{\alpha-1}. \tag{29}$$

Since $p(t) - \lambda(t) = \dfrac{q^*(t)(p(t))^{1-\alpha}}{\alpha}$, (29) rewrites as

$$\frac{1}{\alpha} \cdot \frac{dq^*(t)}{dt} = (p(t))^{\alpha-1} \left\{ \left[1 + \frac{\alpha - 1}{\alpha} \cdot q^*(t)(p(t))^{-\alpha} \right] \frac{dp(t)}{dt} - \frac{d\lambda(t)}{dt} \right\}. \tag{30}$$

Now, using the following information:

$$\begin{cases} \dfrac{dp(t)}{dt} = -\dfrac{n(p(t))^{1-\alpha}}{\alpha} \cdot \dfrac{\partial q^*(t)}{\partial t} \\ \dfrac{\partial \lambda(t)}{\partial t} = [\rho + \delta - f'(k(t))]\lambda(t) \ , \\ \lambda(t) = p(t) - \dfrac{q^*(t)(p(t))^{1-\alpha}}{\alpha} \end{cases} \tag{31}$$

we obtain

$$\frac{1}{\alpha} \cdot \frac{dq^*(t)}{dt} = -\frac{n}{\alpha} \left[1 + \frac{\alpha - 1}{\alpha} \cdot q^*(t)(p(t))^{-\alpha} \right] \frac{dq(t)}{dt} - (p(t))^{\alpha-1}[\rho + \delta - f'(k(t))] \left[p(t) - \frac{q^*(t)(p(t))^{1-\alpha}}{\alpha} \right], \tag{32}$$

that is,

$$\frac{dq^*(t)}{dt} \cdot \left\{ \frac{1}{\alpha} + \frac{n}{\alpha} \left[1 + \frac{\alpha - 1}{\alpha} \cdot q^*(t)(p(t))^{-\alpha} \right] \right\} = -(p(t))^{\alpha-1}[\rho + \delta - f'(k(t))] \left[p(t) - \frac{q^*(t)(p(t))^{1-\alpha}}{\alpha} \right]. \tag{33}$$

Define

$$\frac{1}{\alpha} + \frac{n}{\alpha} \left[1 + \frac{\alpha - 1}{\alpha} \cdot q^*(t)(p(t))^{-\alpha} \right] \equiv \beta \tag{34}$$

with

$$\beta \begin{matrix} > 0 \\ = 0 \\ < 0 \end{matrix} \quad \text{if } q^*(t) \begin{matrix} < \dfrac{A\alpha(n+1)}{n(\alpha n+1)} \\ = \dfrac{A\alpha(n+1)}{n(\alpha n+1)} \\ > \dfrac{A\alpha(n+1)}{n(\alpha n+1)} \end{matrix} \tag{35}$$

so that (33) simplifies as follows:

$$\frac{dq^*(t)}{dt} = -\frac{1}{\beta}(p(t))^{\alpha-1}[\rho + \delta - f'(k(t))]\left[p(t) - \frac{q^*(t)(p(t))^{1-\alpha}}{\alpha}\right]. \tag{36}$$

Then, note that for all $\alpha > 1$, we have $\beta > 0$ if $n > 1/\alpha$. Otherwise, the sign of β is ambiguous.

We are now in a position to state the following.

Theorem 3.1. *The steady state requirement $dq^*(t)/dt = 0$ is satisfied if*

$$\left.\begin{matrix} f'(k(t)) = \rho + \delta \\ p(t) = \dfrac{q^*(t)(p(t))^{1-\alpha}}{\alpha} \Rightarrow q^*(t) = \dfrac{A\alpha}{1+\alpha n} < \dfrac{A}{n} \end{matrix}\right\} \ \forall\, \alpha > 0$$

and

$$p(t) = 0 \Rightarrow q^*(t) = \frac{A}{n}, \qquad \forall\, \alpha \geq 1.$$

If $\alpha \in (0, 1)$, $q^*(t) = A/n$ does not represent a solution to $dq/dt = 0$.

Proof. See the Appendix. □

We can draw a phase diagram in the space $\{k, q\}$, in order to characterise the steady state equilibrium. For simplicity, consider first the case $\alpha \in (0, 1)$. The locus $\dot{q} \equiv dq/dt = 0$ is given by $q = \frac{A\alpha}{1+\alpha n}$ and $f'(k) = \rho + \delta$ in Figure 1.

Note that the horizontal locus $q = \frac{A\alpha}{1+\alpha n}$ denotes the usual equilibrium solution we are well accustomed with from the existing literature dealing with static market games (see Anderson and Engers, [1], [2]). The two loci partition the space $\{k, q\}$ into four regions, where the dynamics of q is given by (36), as summarised by the vertical arrows. The locus $\dot{k} \equiv dk/dt = 0$ as well as the dynamics of k, depicted by horizontal arrows, derive from (3). Steady states, denoted by M, L along the horizontal arm, and P along the vertical one, are identified by intersections between loci. Note that the situation illustrated in Figure 1 is only one out of five possible configurations, due to the fact that the position of the vertical line $f'(k) = \rho + \delta$ is independent of the demand parameters, while the horizontal locus $q = \frac{A\alpha}{1+\alpha n}$ shifts

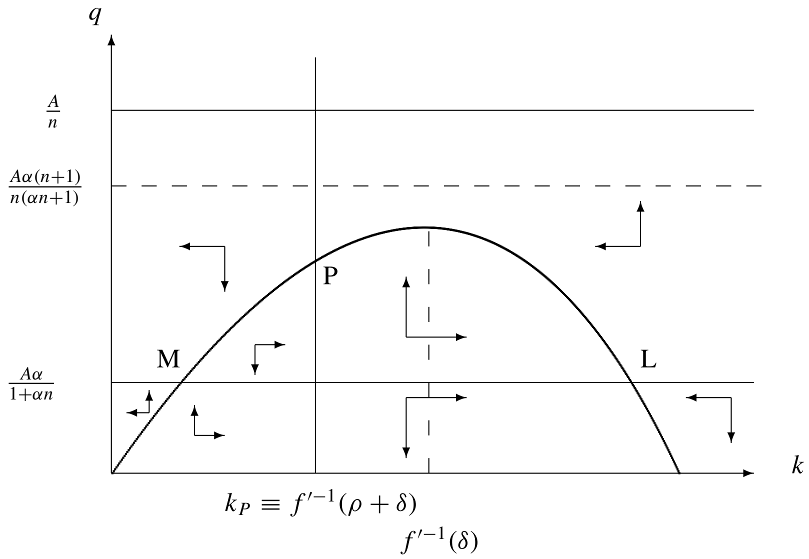


Figure 1: Cournot competition, $\alpha < 1$ (and $\beta > 0$ in ss).

upwards as A and/or α increase. Therefore, we obtain one out of five possible regimes:

(1) There exist three steady state points, with $k_M < k_P < k_L$ (this is the situation depicted in Figure 1). M is a saddle point; P is an unstable focus. L is again a saddle point, with the horizontal line as the stable arm.

(2) There exist two steady state points, with $k_M = k_P < k_L$. Here, M coincides with P, so that we have only two steady states which are both saddle points. In $M = P$, the saddle path approaches the saddle point from the left only, while in L the stable arm is again the horizontal line.

(3) There exist three steady state points, with $k_P < k_M < k_L$. Here, P is a saddle; M is an unstable focus; L is a saddle point, as in regimes (1) and (2).

(4) There exist two steady state points, with $k_P < k_M = k_L$. Here, points M and L coincide. P remains a saddle, while $M = L$ is a saddle whose converging arm proceeds from the right along the horizontal line.

(5) There exists a unique steady state point, corresponding to P. Here, there exists a unique steady state point, P, which is also a saddle point.

An intuitive explanation may be given as follows. The vertical locus $f'(k) = \rho + \delta$ identifies a constraint on optimal capital embodying firms' intertemporal preferences, i.e., their common discount rate. Accordingly, the maximum output level in steady state would be that corresponding to the capacity such that $f'(k) = \delta$. Yet, a positive discounting (that is, impatience) induces producers to install a smaller steady state capacity, which identifies the Ramsey *modified golden rule*. Define this level of capital as $\widehat{k}$. When the reservation price A is very large (or α is large, or n is low), points M and L either do not exist (regime (5)) or fall to the right of P (regimes (2), (3), and (4)). Under these circumstances, the capital constraint is operative and firms choose the capital accumulation corresponding to P.

Hence, if the steady state is unique, it is also a saddle point, and it can be reached for any vector of initial states $\mathbf{k}_0$. If instead there are several steady state points, then there can be more than one saddle point. In such a case, firms reach the one that is characterised by the smaller capacity (again, this solution is reachable from any given $\mathbf{k}_0$). The explanation lies in the fact that, since both steady state points located along the horizontal locus entail the same levels of sales, point L is surely inefficient in that it requires a higher amount of capital. Point M, as already mentioned, corresponds to the optimal quantity emerging from the static version of the game. It is hardly the case of emphasising that this solution encompasses both monopoly (when $n = 1$) and perfect competition (as, in the limit, $n \to \infty$). In M, marginal instantaneous profit is nil.

We can sum up the discussion as follows. The unique efficient and non-unstable steady state point is P if $k_P < k_M$, while it is M if the opposite inequality holds. Such a point is a saddle. The individual equilibrium output is $q^*_{ss} = \frac{A\alpha}{1+\alpha n}$ (where subscript *ss* stands for *steady state*) if the equilibrium is identified by point M, or the level corresponding to the optimal capital constraint $\widehat{k}$ if the equilibrium is identified by point P. The reason is that, if the capacity at which marginal instantaneous profit is nil is larger than the optimal capital constraint, the latter becomes binding. Otherwise, the capital constraint is irrelevant, and firms' decisions in each period are solely driven by the unconstrained maximisation of single-period profits.

When the optimal output is q^*_{ss}, the steady state price is

$$p^*_{ss} = \left(\frac{A}{1+\alpha n}\right)^{\frac{1}{\alpha}}. \tag{37}$$

The per-firm instantaneous profits in steady state are

$$\pi^*_{ss}\big|_{q^*_{ss}} = \alpha \cdot \left(\frac{A}{1+\alpha n}\right)^{\frac{1+\alpha}{\alpha}}, \tag{38}$$

while they are $\pi^*_{ss}\big|_{f(\widehat{k})} = f(\widehat{k})\left[A - nf(\widehat{k})\right]^{\frac{1+\alpha}{\alpha}}$ if the optimal output is $f(\widehat{k})$.

Now consider the case $\alpha > 1$. Here, there exists the additional horizontal arm given by $q = \frac{A}{n} > \frac{A\alpha}{1+\alpha n}$. The overall situation is depicted in Figure 2.

Points M, P, and L are characterised as in the previous case (where $\alpha \in (0, 1)$). The features of points U and V can be quickly summarised as follows. From the direction of the arrows in Figure 2, it appears that point U is completely unstable, while point V is clearly inefficient, and can be disregarded. Note finally that U portrays a situation where Cournot players would indeed behave as perfect competitors. This also appears in the first-order condition of the static game (see,

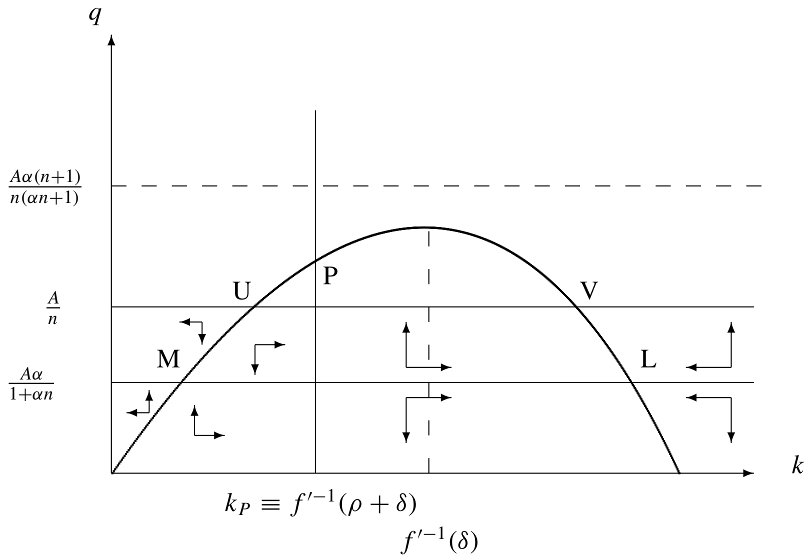


Figure 2: Cournot competition, $\alpha > 1$ (and $\beta > 0$ always).

e.g., Lambertini, [9], p. 331), where $q = A/n$ is a minimum. Steady state profits perfirm are as above.

We stress that the foregoing analysis encompasses the settings examined by Fershtman and Muller ([7]) and Cellini and Lambertini ([4]). The former paper considers a homogeneous-good duopoly where firms always sell a quantity equal to the installed capacity at time t. Therefore, in their paper the steady state at $q^*_{ss} = \frac{A\alpha}{1+\alpha n}$ does not appear. The latter paper accounts for the equilibrium dictated by demand parameters; although allowing for product differentiation, it confines itself to a linear demand setup.

Theorem 3.1 produces the following relevant corollaries.

Corollary 3.1. *In the steady state at M, the elasticity of demand w.r.t. price is constant and it is* $\left|\varepsilon^*_{Q_{ss},p_{ss}}\right| = \frac{1}{n}$ *for all* $\alpha > 0$.

Proof. See the Appendix. □

The explanation for this result is intuitive. In general, the pricing behaviour of a Cournot oligopoly is described by

$$p(Q)\left(1 + \frac{s_i}{|\varepsilon|}\right) = c_i'(q_i), \tag{39}$$

where $s_i \equiv q_i/Q$ and $c_i'(q_i)$ define, respectively, the market share and marginal cost of firm i (see Novshek, [10]). In our setting, $c_i'(q_i)$ is constant, so that the optimal per-firm output is chosen to determine the price as a constant markup over marginal cost.

Corollary 3.2. *As the number of firms tends to infinity, the steady state equilibrium at point M reproduces perfect competition, for all $\alpha > 0$.*

Proof. See the Appendix. □

Corollary 3.3. *The steady state capital is everywhere nondecreasing in α.*

Proof. See the Appendix. □

4 The Social Optimum

The solution of the planner's problem can be quickly dealt with, as its analysis is largely analogous to, but simpler than, the oligopoly equilibrium. The instantaneous maximand of the benevolent planner is social welfare, defined as the sum of producer and consumer surplus:

$$SW(t) = \Pi(t) + CS(t), \tag{40}$$

where

$$\Pi(t) = \sum_i \pi_i(t) \tag{41}$$

with $\pi_i(t) = q_i(t)(A - Q(t))^{\frac{1}{\alpha}}$, and

$$CS(t) = \int_0^{Q(t)} (A - s(t))^{\frac{1}{\alpha}} ds(t) = \frac{\alpha}{1+\alpha}\left[A^{\frac{\alpha+1}{\alpha}} - (A - Q(t))^{\frac{1+\alpha}{\alpha}}\right]. \tag{42}$$

Technicalities are largely analogous to the case of the Cournot oligopoly. Therefore, we confine ourselves to the characterisation of the planner's solution in terms of outputs and capital endowments.

The steady state solutions for the planner are $f'(k_i(t)) = \rho + \delta$ and $q_{ss}^{SP}(t) = A/n$, where the superscript SP stands for *social planning*. Note that the demand-driven solution corresponds to an overall output $Q_{ss}^{SP}(t) = A$. The latter is the perfectly competitive output for the whole market, at which $p(t) = 0$. It is then trivial to prove that this also coincides with the steady state of the Bertrand market game.[5]

In the steady state given by $nq_{ss}^{SP}(t) = A$, we have

$$SW_{ss}^{SP} = CS_{ss}^{SP} = \frac{\alpha A^{\frac{1+\alpha}{\alpha}}}{1+\alpha} \tag{43}$$

while obviously $\Pi_{ss}^{SP} = 0$.

[5]This is due to the assumption of product homogeneity. With differentiated products (as in Cellini and Lambertini, [4]), equilibrium outputs under both social planning and Bertrand competition would depend upon α.

Finally, we can assess the welfare distortion associated to the steady state Cournot equilibrium where $Q^*_{ss} = nq^*_{ss}$, compared to the above social optimum. We obtain the following.

Proposition 4.1. *The welfare distortion due to Cournot competition is decreasing both in α and in n.*

Proof. See the Appendix. □

The fact that the welfare loss associated to oligopoly is decreasing in the number of Cournot agents is not surprising at all. The intuition behind an analogous effect, associated with an increase in α, can be immediately interpreted as follows. Any increase in α entails that the area between the demand function and the axes (p and Q) becomes larger. The same holds for each individual output q^*_{ss}. Thus, increasing the size of the market translates into increasing the toughness of competition and welfare. Given the socially efficient output at $Q^{SP}_{ss}(t) = A$, the foregoing argument implies that the Cournot welfare loss must decrease as α becomes higher.

5 Horizontal Mergers and the Modified Golden Rule

The foregoing analysis highlights that, unlike the standard analysis of the Ramsey model with a representative consumer facing a competitive single-good industry, here there may exist demand-driven steady state equilibria that fall short of the modified golden rule. This is the case both in the Cournot model (point M in figures 1 and 2) and in the planning (or Bertrand) model. This happens whenever parameters $\{A, n, \alpha, \delta, \rho\}$ are such that the intersection between the horizontal arm given by the market-driven equilibrium behaviour and the concave locus $\dot{k} = 0$ locates to the left of the vertical arm where $f'(k) = \rho + \delta$.

Consider point M in the Cournot model, where $q^*_{ss} = \frac{A\alpha}{1+\alpha n}$ and suppose that point M is to the left of point P. The equilibrium output has the following properties:

$$\frac{\partial q^*_{ss}}{\partial A} > 0; \; \frac{\partial q^*_{ss}}{\partial n} < 0; \; \frac{\partial q^*_{ss}}{\partial \alpha} > 0. \tag{44}$$

This entails that the larger A and α are, the closer the economy gets to the modified golden rule. Since A and α are measures of market size, we may say that ‘the larger the market, the easier it is to reach the Ramsey golden rule.’ However, exactly the opposite holds the larger n is: if the number of firms is very high, the horizontal arm can be arbitrarily close to the horizontal axis, thus utterly preventing the economy from reaching the modified golden rule. Much the same conclusion holds under social planning (or Bertrand behaviour), where $q^{SP}_{ss} = \frac{A}{n}$.[6]

This leads to a seemingly counterintuitive result.

[6]To this regard, recall that, whenever M is to the left of p, then M is a saddle point while P is an unstable focus irrespective of the market regime under consideration.

Proposition 5.1. *Suppose the market-driven equilibrium is a saddle point. Then, the economy can be driven to the modified golden rule through a sufficiently large wave of horizontal mergers.*

This is surely feasible under social planning, given the assumption of a constant returns to scale technology (observe that SW_{ss}^{SP} in (43) is independent of n). Given product homogeneity, this also entails that firms's profits are unaffected by a reduction of n in the Bertrand case, and therefore horizontal mergers are feasible as well under price competition.[7] There remains to assess the Cournot case.

In the existing literature on mergers, the static Cournot model is usually adopted. The well-known paper by Salant *et al.* ([13]) proposed a puzzle concerning the profitability of horizontal mergers in a Cournot oligopoly where all firms use the same technology, characterised by constant returns to scale. According to their analysis, we should not worry about horizontal mergers, unless it involves *almost all* the firms in the industry. However, their model drastically underestimates the incentive towards merging, as shown by Perry and Porter ([11]) by introducing fixed costs into the picture. They assume that the aggregate amount of capital is fixed at the industry level, and the associated fixed costs are distributed across firms in proportion to their individual holdings of the capital factor. By doing so, Perry and Porter allow the model to account for the intuitive fact that a merger gives rise to a new firm that is bigger than its parts (previously independent firms), and the new productive unit may produce a larger output at any given marginal costs, as compared to the rivals (as well as the previously independent firms that participated in the merger). This perspective leads Perry and Porter to find that, contrary to the conclusions reached by Salant *et al.* ([13]), the price increase generated by the merger can often suffice to compensate for the decrease in the output of the merged firm, and therefore may make such a merger a profitable one. The presence of a fixed cost, and the related efficiency evaluation, may lead one to conclude, in some circumstances, that the gain in efficiency is sufficiently high to more than offset the decrease in output caused by the merger, and therefore the merger is also socially efficient and should not be obstacled or forbidden.

The setting proposed in this chapter can be profitably applied to the investigation of the private and social incentives towards horizontal mergers, as a tool for manoeuvring the long-run equilibrium path of the economic system.

To our aim, it suffices to investigate the case where the pre-merger equilibrium is determined by the horizontal arm, while the post-merger equilibrium is the Ramsey one.[8] Assume m firms consider the perspective of merging. The profitability of the

---

[7]This would be true *a fortiori* with differentiated products, as shown by Deneckere and Davidson ([5]).

[8]Note that the opposite cannot occur. If firms are in the Ramsey equilibrium before the merger, then they cannot move to a demand-driven equilibrium after a merger of any size, since the merger itself reduces the number of firms.

merger is ensured iff:

$$\frac{\pi_{ss}^{*}(m-n+1)\big|_{f(\widehat{k})}}{m} > \pi_{ss}^{*}\big|_{q_{ss}^{*}}$$
$$\Rightarrow \frac{f(\widehat{k})\left[A-(n-m+1)f(\widehat{k})\right]^{\frac{1+\alpha}{\alpha}}}{m} > \alpha \cdot \left(\frac{A}{1+\alpha n}\right)^{\frac{1+\alpha}{\alpha}}, \tag{45}$$

which cannot be evaluated analytically. However, with reference to the previous literature (Salant *et al.*, [13]), a natural case to explore is that where $n = 3$ and $m = 2$. Moreover, set A at an admissible value, e.g., $A = 10$, and

$$f\left(\widehat{k}\right) = \frac{q_{ss}^{*}(n) + q_{ss}^{*}(n-m+1)}{2} = \frac{\alpha}{2}\left[\frac{A}{1+\alpha n} + \frac{A}{1+\alpha(n-m+1)}\right]. \tag{46}$$

Under these assumptions, inequality (45) is satisfied for all admissible values of parameter α, with

$$\lim_{\alpha\to\infty} \frac{f(\widehat{k})\left[10-2f(\widehat{k})\right]^{\frac{1+\alpha}{\alpha}}}{2} - \alpha \cdot \left(\frac{A}{1+3\alpha}\right)^{\frac{1+\alpha}{\alpha}} = \frac{5}{18}, \tag{47}$$

where $f(\widehat{k})$ is defined as in (46). Proceeding likewise, one can check the profitability of a merger of any admissible size, such that the *ex post* equilibrium is the Ramsey golden rule. As to the social desirability of such merger operations, this can be quickly dealt with by observing that the merger involving m firms always produces a decrease in the industry output and therefore an increase in the equilibrium price. As a consequence, social welfare would always decrease after such a merger. Accordingly, an antitrust agency should prevent firms from merging. However, the regulator, instead of explicitly prohibiting a merger, could intervene indirectly by manoeuvring the interest rate in the financial market. To see this, consider briefly the following perspective. If the capital markets are efficient, the discount rate ρ used by firms is given by the long-run interest rate on such markets. Therefore, by changing the interest rate, a public authority is able to affect the position of the horizontal arm identifying the Ramsey equilibrium in the phase diagram. Then, if the pre-merger oligopoly equilibrium locates on the horizontal (Cournot) arm and prevents the economy from reaching the modified golden rule, the policy maker may appropriately reduce the interest rate, thus making possible the attainment of the Ramsey equilibrium.

6 Concluding Remarks

We have taken a differential game approach in order to revisit the Ramsey growth model from the standpoint of the representative firm rather than the consumer. We have confined our attention to the steady state alloactions of Nash equilibria under the open-loop information structure. We have shown that there are configurations of parameters where the economy cannot reach the Ramsey golden rule in steady state, due to the presence of a stable equilibrium dictated by demand conditions. This may happen irrespective of the market regime, be that Cournot, Bertrand, or social planning. The distance between the market-driven equilibrium and the modified golden rule decreases in the number of firms. Therefore, a wave of horizontal mergers can indeed drive the economy towards the Ramsey golden rule.

Acknowledgments

We thank Steffen Jørgensen, two anonymous referees, Simon Anderson, Gustav Feichtinger, Yuri Kaniovski, and George Leitmann for useful comments and discussion on earlier drafts. The usual disclaimer applies.

Appendix

Proof of Theorem 3.1. The proof largely relies on Lemma 3.1 and the above discussion. To complete it, just observe that, for all $\alpha \geq 1$, we have

$$\frac{A\alpha}{1+\alpha n} < \frac{A}{n} < \frac{A\alpha(n+1)}{n(\alpha n+1)} \tag{48}$$

so that $\beta > 0$ everywhere, while for all $\alpha \in (0, 1)$,

$$\frac{A\alpha}{1+\alpha n} < \frac{A\alpha(n+1)}{n(\alpha n+1)} < \frac{A}{n} \tag{49}$$

so that $\beta > 0$, when evaluated in the steady state, where surely $\alpha > \frac{1}{n}$. □

Proof of Corollary 3.1. To prove the claim, it suffices to plug the steady state output of the overall population of firms, $Q_{ss}^* = nq_{ss}^* = \frac{nA\alpha}{1+\alpha n}$, into (6) to obtain $\left|\varepsilon^*_{Q_{ss}p_{ss}}\right| = 1/n$. □

Proof of Corollary 3.2. Recall that $Q_{ss}^* = nq_{ss}^* = \frac{nA\alpha}{1+\alpha n}$, and check that

$$\lim_{n\to\infty} \frac{nA\alpha}{1+\alpha n} = A. \tag{50}$$

This proves the result. □

Proof of Corollary 3.3. To prove the claim, it suffices to observe that, given the assumptions about technology, then in general $\frac{\partial q(t)}{\partial k(t)} > 0$ to the left of point P, and

$$\text{sign}\left\{\frac{\partial k^*(t)}{\partial \alpha}\right\} = \text{sign}\left\{\frac{\partial q^*(t)}{\partial \alpha}\right\}$$

in the same range, where clearly $\frac{\partial q^*(t)}{\partial \alpha} > 0$. This holds for all $k^*(t)$ as determined by point M. If M coincides with P, then the optimal capital endowment is given by the Ramsey rule. This argument implies the Corollary. □

Proof of Proposition 4.1. Consider that the welfare distortion, i.e., $SW_{ss}^{SP} - SW_{ss}^*$ is proportional to the difference between the output level of the planner and the overall steady state production of the Cournot firms:

$$Q_{ss}^{SP} - Q_{ss}^* = \frac{A}{1+\alpha n}, \tag{51}$$

which is everywhere decreasing both in α and (obviously) in the number of firms operating in the Cournot setting, n. This implies the proposition. □

References

[1] Anderson S.P. and Engers M., Stackelberg vs Cournot oligopoly equilibrium, *Int. J. of Industrial Organization*, 10, 1992, 127–135.

[2] Anderson S.P. and Engers M., Strategic investment and timing of entry, *Int. Ec. Rev.*, 35, 1994, 833–853.

[3] Arrow K., Applications of control theory to economic growth, in G.B. Dantzig and A.F. Veinott, Jr (eds), *Mathematics of the Decision Sciences, Part 2*, American Mathematical Society, Providence, 1968.

[4] Cellini R. and Lambertini L., A dynamic model of differentiated oligopoly with capital accumulation, *J. Econ. Theory*, 83, 1998, 145–155.

[5] Deneckere R. and Davidson C., Incentives to form coalitions with Bertrand competition, *RAND J. Econ.*, 16, 1985, 473–486.

[6] Dockner E.J. and Gaunersdorfer A., On the profitability of horizontal mergers in industries with dynamic competition, *Japan and the World Economy*, 13, 2001, 195–216.

[7] Fershtman C. and Muller E., Capital accumulation games of infinite duration, *J. Econ. Theory*, 33, 1984, 322–339.

[8] Gowrisankaran G., A dynamic model of endogenous horizontal mergers, *RAND J. Econ.*, 30, 1999, 56–83.

[9] Lambertini L., Cartel stability and the curvature of market demand, *Bulletin Econ. Research*, 48, 1996, 329–334.

[10] Novshek W., Cournot equilibrium with free entry, *Rev. Econ. Studies*, 47, 1980, 473–486.

[11] Perry M.K. and Porter R.H., Oligopoly and the incentive for horizontal merger, *American Econ. Rev.*, 75, 1985, 219–227.

[12] Ramsey F.P., A mathematical theory of saving, *Economic Journal*, 38, 1928, 543–559.

[13] Salant S.W., Switzer S. and Reynolds R.J., Losses from horizontal merger: The effects of an exogenous change in industry structure on Cournot-Nash equilibrium, *Quarterly Journal of Economics*, 98, 1983, 185–213.

[14] Simaan M. and Takayama T., Game theory applied to dynamic duopoly problems with production constraints, *Automatica*, 14, 1978, 161–166.

[15] Solow R., A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70, 1956, 65–94.

[16] Swan T.W., Economic growth and capital accumulation, *Economic Record*, 32, 1956, 334–361.

Economic Growth and Process Spillovers with Step-by-Step Innovation

Shravan Luckraz
Faculty of Business
Bond University
Queensland, 4229
Australia
sluckraz@bond.edu.au

Abstract

This paper extends previous research on the effects of process imitation on economic growth by accounting for stochastic intra-industry spillovers. We employ a non-Schumpeterian growth model to determine the impact of such spillovers on investment in industries where firms are either neck-and-neck or unleveled. Our central finding is that, in an economy where the representative industry is a duopoly, research and development (RD) spillovers positively affect economic growth. While other non-Schumpeterian models assume that the imitation rate of laggard firms is unaffected by the RD effort of the leader firm, we consider the case where the latter's RD activity generates some positive externality on its rivals' research. In this construct, the duopolists in each industry play a two-stage game. In the first stage, they invest in RD which can reduce their costs of production only if they successfully innovate and they compete with each other by using Markovian strategies. In the second stage, they compete in the product market. At any point in time, an industry can be either in the neck-and-neck state or in an unleveled state where the leader is n steps ahead of the follower. At the steady state, the inflow of firms to an industry must be equal to the outflow. By determining the steady state investment levels of each industry, we demonstrate a positive monotonic relationship between the spillover rate and economic growth.

1 Introduction

Endogenous growth theorists have investigated extensively the impact of low appropriability on the growth rate of an economy. While Schumpeterian [14] models[1] posit a negative relationship between them, more recently developed non-Schumpeterian models have shown that the relationship is nonmonotonic and that a strict negative relationship only holds whenever the level of appropriability is very low.[2] Aghion et al. [2] made a seminal contribution to the non-Schumpeterian

[1]Also see [10], [11], and [15].

[2]In particular, they find an inverted U-shaped relationship between innovation and a parameter which promotes competition in their models.

branch of endogenous growth by showing that static monopoly is not always a necessary evil for long-run efficiency in a step-by-step innovation growth model (referred to as the AHV model). Their result is consistent with the empirical findings in [5] and [12] and the theoretical predictions in [3]. The AHV model together with the contributions noted above, however, downplay the role of externalities in strategic interactions among firms. In contrast, an important strand in the industrial organization literature argues that process spillovers play a key role in two-stage noncooperative RD games[3] because they capture the diffusion of technology between leaders and laggards. Since such externalities depend on the level of appropriability in an industry, the effect of lower or higher appropriability on growth can be observed by the impact of spillovers on the latter. The purpose of this study is to determine the effects of process RD spillovers on growth by extending the AHV framework.

The relationship between spillovers and RD incentives has two aspects. First, in many industries, firms undertake RD investments in order to develop new products or processes. One feature of RD investment that distinguishes from other forms is that the firms which do the investing are often not able to exclude others from freely benefiting from their investments. Thus, the benefits from RD investments spill over to other firms in the economy.[4] Now since the laggards can improve their own technology by free-riding on the leader's research, technologically more advanced firms might have a disincentive to undertake more research since their productivity lead might be significantly reduced in the presence of such spillovers. Hence, the first characteristic of RD spillovers is that it can potentially reduce research incentives. The second aspect of spillovers is related to the concept of "escape competition."[5] When the laggard firms benefit from process RD spillovers, they improve their own technology and thereby reduce the technology gap between the leader and themselves. As a result, there will be competitive pressure on the leader to innovate further to maintain its lead. Those two opposing forces of spillovers on RD are observed in the empirical findings in [6]. Given the importance of spillovers in the strategic interactions between firms investing in RD, we allow them to play a major role in our model.

The purpose of this chapter is to explore the interdependence of spillovers, appropriability, and growth in a framework where strategic interactions between firms are taken into account. Our work is primarily motivated by the empirical study in [16] where strong support for technological spillovers from aggregate research intensity to industry-level innovation success is found. We develop a dynamic general equilibrium model that is distinct from the AHV model in two major ways.

[3] D'Aspremont and Jacquemin [4] pioneered this framework.

[4] Griliches [9] emphasized the significance of spillovers in modeling and estimating the effects of RD investments.

[5] "Escape competition" refers to the motive of innovating in order to escape competition; that is, firms in the neck-and-neck state will innovate to obtain a productivity lead over their rivals.

First, we consider the case for homogeneous Cournot competition rather than differentiated Bertrand competition.[6] Second, we assume the hazard rate of imitation to be dependent on the spillovers induced by the leader's RD. Thus, owing to the presence of externalities, the probability of the laggard making a successful innovation is a fraction of the leader's probability of doing so. We therefore highlight the role of spillovers in an economy in which firms play a differential RD game. We consider a one-sector endogenous growth model whereby in the second stage, duopolists in the representative industry sell a homogeneous good to consumers who spend a fixed proportion of their income in each period. In the first stage, while the industry leader innovates and moves one step up the technology ladder with some probability, the follower imitates and catches up with the leader with some hazard rate. Thus, at any point in time, we can have industries in different states with the technology gap ranging from 0 to n.[7] Stationarity implies that for some state n, the inflow of industries in that state should be equal to the outflow. By computing the growth rate at the steady state, we derive two sets of results.

First, as in the AHV model, we look at the case where the innovation lead of the leader is so high that it has no incentive to increase its lead by one step. This simplification allows us to reduce the number of states to only state 0 and state 1. By computing the fraction of industries in those states and by deriving the optimal neck-and-neck and unleveled innovation rates, we can derive the steady state growth rate of the economy by using the stationarity condition that the fraction of industries in every state is constant in long-run equilibrium. We then use comparative statics to find the impact of the spillover parameter on the growth rate and consider whether lack of appropriability necessarily reduces the growth rate for the case of a large innovation lead.

Second, we consider the case where the innovation lead can be small and hence, in the case at hand, the leader has incentives to extend its lead further than one step. While the AHV model uses a method of asymptotic expansion to derive results for the "small innovation lead" case, our results are derived directly from Bellman's equations since our assumption of Cournot competition allows our profit functions to be independent of the competition parameter as opposed to the differentiated Bertrand case. Thus, we also solve for the optimum RD effort of the leader who wants to move more than one step ahead of its follower and derive the fraction of industries which might be in that state in equilibrium. We shall then have three optimal levels of RD effort as well as three steady state fractions of industries in the respective states. Results are derived in a similar fashion as in the previous case with "large innovation lead" and we can therefore analyze how the policymaker, by varying the appropriability rate of the industry, can affect the research incentives in the neck-and-neck state, as well as in the unleveled states. Some policy implications on whether larger appropriability promotes growth are then drawn.

[6] Although, in another study, Aghion et al. [1] compared the Bertrand and the Cournot cases, they did not allow the industry leader to extend its lead by more than one step in their model.

[7] Note that state 0 is also known as the neck-and-neck state.

Our results show that the growth rate is unaffected by the spillover rate for the "large innovation lead" case. Thus, in contrast to the traditional Schumpeterian argument, we find that lower appropriability does not necessarily reduce growth, when the rate of imitation depends on the RD spillovers. Moreover, we find as in the AHV model that the level of RD effort is greatest at the neck-and-neck state and that this constitutes a major component of the economy's growth rate. For the case of "small innovation lead," our findings indicate that process spillovers affect growth positively and that imitation and innovation can be strategic substitutes. We also note that the fraction of industries in the state in which the leader is more than one step ahead is positively related to the spillover rate and to the RD effort in the neck-and-neck state. Clearly, it follows from our results that the trade-off between short-run monopoly and long-run efficiency is not observable in a framework where both strategic interactions between firms and diffusion of technology are taken into account. Hence, an immediate policy implication is that greater appropriability is not always good for the growth rate of the economy.

On the normative side, one interpretation of the main result is that the more technologically advanced firm will innovate even further in the face of process spillovers in order to maintain its productivity lead. It is noteworthy that unlike the AHV result, ours is not heavily dependent on the product differentiability parameter. Thus, the model identifies the "pure" effect of process spillovers, which enhance imitation, on welfare. Specifically, imitation and unintended technological diffusion can promote growth. Consequently, we shed some light on the ongoing debate as to whether or not restricting the act of reverse engineering is justifiable on economic grounds. We believe that in an industry where reverse engineering can hasten the diffusion of technology via process spillovers, the strategic interaction between rival firms will guarantee that a competitive environment always prevails. Furthermore, the innovator does not have an incentive to lay back as a monopolist as its technological lead might fall. Hence, growth is always enhanced by more competition.

The remainder of this chapter is organized as follows. Section 2 presents the formal model, derives the steady state growth rate, and provides some comparative static results. Section 3 offers some concluding remarks.

2 The Model

2.1 Overview

We consider a model with d goods, d industries with 2 firms each, and infinitely lived identical consumers. The latter face two optimization problems: temporal and intertemporal utility maximization. Preferences across goods and time are logarithmic. In the intertemporal problem, the consumer chooses the optimal labor supply and consumption (or expenditure) for each period. The remaining income is invested in the industries' RD. To simplify the model, we make the following assumptions. First, we normalize expenditures to allow the rate on return on savings (savings of agents) to be constant and equal to the exogenously given discount rate. Second, we assume that labor supply is perfectly elastic. With the optimal amount to be spent in each period chosen, the representative agent can thus derive his demand function for each industry from his temporal optimization problem.

On the production side, the industry demand is derived from the consumer problem and taking the demand schedule as given, the duopolists in the representative industry compete in Cournot fashion to choose their respective output and research intensities (which are innovation rate for the leader and imitation rate for the laggard). If the productivity of one firm is higher than that of its competitor, then the former is the leader and the latter is the follower. Moreover, the leader moves up the technology ladder with a Poisson hazard rate by employing some units of labor, while the follower catches up with some Poisson hazard rate which consists of two components (one of which is proportional to the leader's success rate). The leader does not benefit in terms of externality from the follower, but the latter, practicing process reverse engineering, benefits from the leader.

The choices of the firm's variables which are quantity and research intensities are found sequentially. We therefore use backward induction in a two-stage non-cooperative game setting to formulate the firms' optimum behavior. In the second stage, the firms play a Cournot game to determine their respective quantities. Their respective profits as functions of their productivity levels can thus be derived. In the first stage, the leader (follower) plays a differential game to choose its optimal innovation (imitation) in each possible state taking the technology gap as given. The Markovian Nash equilibrium is then found. The steady state growth rate is determined by the optimal values of imitation and innovation and is used to derive results. The effects of changes in appropriability and growth are subsequently analyzed.

2.2 Formal Model

2.2.1 Consumers

Let d, C_t, L_t, Q_{it}, R_t, and W_t be the number of industries, the consumption of the representative agent, his labor supply, the quantity produced in the industry i for $i = 1, 2, \dots, d$, the interest rate on savings, and the wage rate, respectively, at time t. Then the intertemporal preference of the agent can be written as

$$U(\rho) \equiv \int_0^{\infty} e^{-\rho t}(u(C_t) - L_t)\, dt, \quad \rho > 0, \tag{2.1}$$

where ρ is the discount rate and where

$$u(C_t) \equiv \ln(C_t), \quad t \geq 0. \tag{2.2}$$

The intertemporal utility maximization problem results in (i) $W_t = 1$ and (ii) $R_t = \rho$ after normalization.

Remark 2.1. We observe that the rate of return on savings is not equal to the rate of return on capital in general as it usually depends on a financial sector which is not modeled here. While R_t is the rate of return on savings which savers accrue with certainty when they place their money not spent on consumption into the bank, the rate of return on capital is uncertain (probabilistic) in the model. Thus, R_t does not imply that the rate of return on capital is constant and equal to the discount rate. It just says that the rate of return on savings is equal to the discount rate. We now give the mathematical proof for (i) $W_t = 1$ and (ii) $R_t = \rho$. Let the consumer's wealth at time t be A_t, let P_t be the price of consumption, and let $P_t C_t = 1$ due to normalization. Then we have

$$H = \ln(C_t) - L_t + \lambda_t (R_t A_t + W_t L_t - P_t C_t).$$

Furthermore, since $\frac{dH}{dC_t} = 0$ implies $\frac{1}{P_t C_t} = \lambda_t$ and $P_t C_t = 1$, it follows that $\lambda_t = 1$. Moreover, since $\frac{dH}{dL_t} = 0$, $W_t \lambda_t = 1$ and hence $W_t = 1$. Also

$$\frac{d\lambda_t}{dt} = \lambda_t (\rho - R_t),$$

but since $\lambda_t = 1$, $\frac{d\lambda_t}{dt} = 0$; hence $\rho - R_t = 0$ and therefore $R_t = \rho$. The normalization assumption simply states that the aggregate income $P_t C_t = 1$. The above shows that optimality conditions from the Hamiltonian function would imply that $W_t = 1$. But then, as a corollary, we must have $L_t = 1$. Therefore, aggregate income being equal to one is consistent with wages equal to unity.

The temporal consumer preference is given by

$$u(C_t) = \sum_{i=1}^{d} \ln(Q_{it}), \quad t \geq 0. \tag{2.3}$$

Fix t. Static utility maximization of $u(C_t)$ w.r.t. the variables $Q_{1t}, \ldots, Q_{dt}$ subject to the constraint $\sum_{i=1}^{d} P_{it} Q_{it} = 1$ gives

$$Q_{it} = \frac{1}{d P_{it}}, \quad i = 1, \ldots, d.$$

Hence the industry demand curve is

$$Q_{it} = \frac{M}{dP_{it}}, \quad i = 1, \ldots, d, \tag{2.4}$$

where $M = \frac{1}{d}$.

2.2.2 Producers

Given the industry demand (2.4) each firm will choose its respective optimal production q_{ijt} such that[8] $\sum_{j=1}^{2} q_{ijt} = Q_{it}, i = 1, \ldots, d; t \geq 0$. We assume that

[8]In this subsection, we sometimes omit subscript i for simplicity.

firm 1 is the leader and firm 2 is the follower. Each firm's production function is given by

$$q_j = A_j L_j, \quad j = 1, 2. \tag{2.5}$$

It can easily be inferred from (2.5) that the per unit cost of the jth firm is given by $\frac{W^*}{A_j}$ where W^* is the economy level wage rate. Also, due to our assumption that $W_t = 1$, the per unit cost becomes

$$c_j = \frac{1}{A_j}, \quad j = 1, 2. \tag{2.6}$$

We denote the productivity lead of a leader who is n steps ahead as follows:

$$\left(\frac{c_2}{c_1}\right)^n = \left(\frac{A_1}{A_2}\right)^n, \quad n = 0, 1, 2, \ldots. \tag{2.7}$$

We also denote the size of the lead by

$$\gamma \equiv \left(\frac{A_1}{A_2}\right). \tag{2.8}$$

Thus an increase in γ and/or n will increase (decrease) the leader's (follower's) profit. We assume innovative and imitative activities to be randomly determined. Specifically, we assume (as in AHV) that the leader or a neck-and-neck firm in state n, by employing $\psi(x)$ units of labor in RD, moves one step ahead with a Poisson hazard rate of x_n for $n = 0, 1, 2$. (Formally, we let $H(n) = x_i u(n)$ be the hazard rate in state n where $u(n)$ is the hazard function. Using the exponential distribution which has been widely used in the literature, $u(n) = 1$ and hence $H(n) = x_i$ for firm i.) The follower catches up with its rival with a Poisson hazard rate of $\overline{x}_i + h_n$, where the RD cost function $\psi(x)$ is an increasing and convex function of the RD effort and h_n is the ease of imitation or RD spillover parameter.[9] We define such spillovers as follows:

$$h_n = bx_n. \tag{2.9}$$

Our definition of spillovers is similar to that used in [6] although we pursue some extensions. In particular, we define spillovers to include valuable knowledge generated in the research process of the leader, which becomes accessible to the follower if and only if the latter is reverse engineering the innovator's research process. Given that spillovers favor imitation, it becomes a better strategy for the follower to imitate by feeding off the leader's innovation at least initially. Thus, the follower is necessarily an imitator.

Also note that while AHV make use of an "ease to imitate" parameter and a "competition parameter" to proxy the absence of institutional, legal, and regulatory

[9] In any state n, x_n is the leader's success rate while $\overline{x}_n$ is the follower's catching up probability.

impediments connected with patent laws and regulations, in our model b in (2.9) includes all of these factors.

We shall consider the stationary closed-loop Nash equilibrium in Markovian strategies in which each firm's RD effort depends on its current state as well as on its current RD level and not on the industry to which the firm belongs or the time. We assume without loss of generality that the RD cost function is given by

$$\psi(x) \equiv \frac{\beta x^2}{2}, \quad \beta > 0. \tag{2.10}$$

This completes the model.

2.3 Solving the Model

We solve the model by backward induction.

2.3.1 Stage 2

Using the inverse demand function (2.4), firm $j's$ profit maximization problem is to compute the quantity

$$\max_{q_j} \left(\frac{Mq_j}{q_i + q_j} - c_j q_j \right) \tag{2.11}$$

for $i \in \{1, 2\} \backslash \{j\}$. The Cournot–Nash quantity for firm j is given by

$$q_j = \frac{Mc_i}{(c_i + c_j)^2} \tag{2.12}$$

for $i \in \{1, 2\} \backslash \{j\}$. The profit function for firm j is given by

$$\pi_j = \frac{Mc_i^2}{(c_i + c_j)^2} \tag{2.13}$$

for $i \in \{1, 2\} \backslash \{j\}$.

Proof. See the Appendix. □

Remark 2.2. (i) If $\gamma \geq 1$ then

$$\overline{\pi}_n = \frac{M}{(1 + \gamma^n)^2}$$

is a strictly decreasing function and $\pi_n = \frac{M}{(1+\gamma^{-n})^2}$ is a strictly increasing function.
(ii) If $\gamma = 1$ then

$$\frac{M}{(1 + \gamma^n)^2} = \frac{M}{(1 + \gamma^{-n})^2} = \frac{M}{4}$$

and if $\gamma > 1$ then

$$\frac{M}{(1+\gamma^{n})^{2}} + \frac{M}{(1+\gamma^{-n})^{2}} > \frac{M}{2},$$

where π_n and $\overline{\pi}_n$ are the profits of the leader and the follower, respectively, in state n.

Proof. See the Appendix. □

The first part of Remark 2.2 states that a higher (lower) relative cost, that is, the larger (lower) the technology gap in favor of the leader (follower) is always strictly advantageous to its profit. The second part of Remark 2.2 states first that when the firms are in the neck-and-neck state, they have equal profits and, second, that the sum of the firm's profit in an asymmetric duopoly is larger than the sum of profits when firms are symmetric. Thus, when there is more than the minimal degree of competition, total profits are lower if firms are neck and neck with identical costs than if one has a relative cost advantage. This fact, which is also consistent with the AHV Bertrand differentiated product case, is important for the derivation of our results.

The intuition behind Remark 2.2 can be given as follows. We first observe that in an industry with identical demand and cost specifications of the model used here a monopolist's profit is always higher than the sum of profits of two symmetric duopolists irrespective of whether the latter compete in prices (Bertrand) or in quantities (Cournot). This is because lack of competition in a monopolistic industry enables the monopolist to extract a greater share of consumer surplus (in the form of higher prices) from consumers than it would have extracted if the market were competitive. Since a duopolistic industry is more competitive than a monopolistic one, consumer surplus is always less in the latter and, hence, industry profits are higher.

Now note that an asymmetric industry is always less competitive (closer to a monopolist) than a symmetric industry with the same number of firms. This is because there is always a large firm which has some monopoly power in the asymmetric industry. Thus, by analogous arguments to those used in the preceding paragraph, the summation of profits in the asymmetric case must be higher than in the symmetric case.

2.3.2 Stage 1 (The Markovian Nash Equilibrium)[10]

Let V_0, V_n, and $\overline{V}_n$ denote the expected present value of the profits of the neck-and-neck firm, the leader, and the follower, respectively. Given that the equilibrium interest rate equals the rate of time preference, we derive V_0, V_n, and $\overline{V}_n$ heuristically from the Bellman equations as follows:

$$V_n(\overline{x}_n, x_0) = \max_{x_n} F_n(x_n, \overline{x}_n, x_0), \tag{2.14}$$

[10]The general case is formulated by Dockner et al. [7].

where

$$F_n(x_n, \overline{x}_n, x_0) \equiv \left(\pi_n - \frac{\beta x_n^2}{2}\right) dt + e^{-r\,dt}[x_n\,dt V_{n+1} + (\overline{x}_n + bx_n)\,dt V_0 + (1 - x_n\,dt - (\overline{x}_n + bx_n)\,dt) V_n].$$

Furthermore,

$$\overline{V}_n(x_n, x_0) = \max_{\overline{x}_n} \overline{F}_n(\overline{x}_n, x_n, x_0), \tag{2.15}$$

where

$$\overline{F}_n(\overline{x}_n, x_n, x_0) \equiv \left(\overline{\pi}_n - \frac{\beta \overline{x}_n^2}{2}\right) dt + e^{-r\,dt}[x_n\,dt\overline{V}_{n+1} + (\overline{x}_n + bx_n)\,dt V_0 + (1 - x_n\,dt - (\overline{x}_n + bx_n)\,dt)\overline{V}_n].$$

Finally

$$V_0(\overline{x}_1) = \max_{x_0} F_0(x_0, \overline{x}_1), \tag{2.16}$$

where

$$F_0(x_0, \overline{x}_1) \equiv \left[\left(\pi_0 - \frac{\beta x_0^2}{2}\right) dt + e^{-r\,dt}(x_0\,dt\overline{V}_1 + x_0\,dt V_1 + (1 - x_0\,dt - x_0\,dt) V_0)\right].$$

As in AHV (2.14) can be interpreted as follows: the value of currently being a leader n steps ahead at time t equals the discounted value at time $(t + dt)$, plus the current profit flow $\pi_n dt$ minus the current RD cost $\frac{\beta x^2}{2} dt$ plus the expected discounted capital gain from innovation, thereby moving one step ahead of the follower, minus the discounted expected capital "loss" from having a follower catch up. Similar interpretations can be made for (2.15) and (2.16). For dt small, $e^{-r\,dt} \approx 1 - r\,dt$ and the second-order terms in (dt) can be ignored. Then (2.14)–(2.16) can be rewritten as follows:[11]

$$rV_n = \pi_n - \frac{\beta x_n^2}{2} + x_n[V_{n+1} - V_n] + (\overline{x}_n + bx_n)[V_0 - V_n] \tag{2.17}$$

$$r\overline{V}_n = \overline{\pi}_n - \frac{\beta \overline{x}_n^2}{2} + x_n[\overline{V}_{n+1} - V_n] + (\overline{x}_n + bx_n)[V_0 - \overline{V}_n] \tag{2.18}$$

$$rV_0 = \pi_0 - \frac{\beta x_0^2}{2} + x_0[\overline{V}_1 - V_0] + x_0[V_1 - V_0]. \tag{2.19}$$

[11] Note that r is the borrowing rate.

Maximizing the right-hand side (RHS) of (2.17)–(2.19), we have

$$x_n = \frac{(V_{n+1} - V_n)}{\beta} + \frac{(V_0 - V_n)b}{\beta} \tag{2.20}$$

$$\overline{x}_n = \frac{(V_0 - \overline{V}_n)}{\beta} \tag{2.21}$$

$$\overline{x}_0 = \frac{(V_1 - V_0)}{\beta}. \tag{2.22}$$

We can now use equations (2.17)–(2.22) to solve recursively for the sequences $\{x_n, \overline{x}_{n+1}, V_n, V_{n+1}\}_{n\geq 0}$. It can be shown[12] that after some recursions, the system above reduces to the following three equations:

$$rV_1 = \pi_1 + \frac{\beta x_1^2}{2} - \beta \overline{x}_1 x_0 \tag{2.23}$$

$$rV_0 = \pi_0 + \frac{\beta x_0^2}{2} - \beta \overline{x}_1 x_0 \tag{2.24}$$

$$r\overline{V}_1 = \overline{\pi}_1 + \frac{\beta \overline{x}_1^2}{2} - \beta b \overline{x}_1 x_1. \tag{2.25}$$

In addition, the following two equations solve the above system:

$$\frac{x_0^2}{2} + rx_0 = \frac{\Gamma_0}{\beta} + \frac{x_1^2}{2}, \tag{2.26}$$

where

$$\frac{\overline{x}_1^2}{2} + (r + [x_0 + bx_1])\overline{x}_1 = \frac{\Gamma_{-1}}{r} + \frac{x_0^2}{2}, \tag{2.27}$$

where Γ_0 and Γ_{-1} are given by $\pi_1 - \pi_0$ and $\pi_0 - \overline{\pi}_1$, respectively. A corollary of Remark 2.2 is that $\Gamma_0 > \Gamma_{-1}$ and this implies there is more incentive for the leader to do research in order to escape competition when it is in the neck-and-neck state.

On a technical note, we now show that if the objective function is maximized w.r.t. x_i and q_i the solution of the game will fail to be subgame perfect. The solution to the game with x_i and q_i chosen simultaneously by a leader firm in state n can be derived from the following:

$$W_n(\overline{x}_n, q_i) = \max_{x_n, q_j} G_n(x_n, q_j, \overline{x}_n, q_i),$$

where

$$\begin{aligned} G_n(x_n, q_j) \equiv{} & \left(\pi_n(q_i, q_j) - \frac{\beta x_n^2}{2}\right) dt \\ & + e^{-r\,dt}[x_n\,dt V_{n+1}(\overline{x}_n + bx_n)\,dt V_0 \\ & + (1 - x_n\,dt - (\overline{x}_n + bx_n)\,dt)V_n]. \end{aligned}$$

[12]Proof can be provided upon request.

Suppose player i commits to q_i after x_n and $\overline{x}_n$ have been chosen. Then player j can make strictly higher profits by choosing

$$q_j^* \equiv q_j^*(x_n, \overline{x}_n) = \text{argmax}_{q_j} \frac{Mq_j}{q_i + q_j} - c_j q_j, \quad i = 1, 2.$$

Thus, $W_n(\overline{x}_n, q_i)$ fails to be subgame perfect as player j can choose to blow up the game if player i commits to q_i. Now it is straightforward to show that the solution derived by the Bellman equation (2.14) that follows is subgame perfect:

$$W_n^*(\overline{x}_n) = \max_{x_n} G_n^*(x_n, \overline{x}_n),$$

where

$$\begin{aligned} G_n^*(x_n, \overline{x}_n) \equiv &\left[\pi_n(q_i^*(x_n, \overline{x}_n), q_j^*(x_n, \overline{x}_n), x_n, \overline{x}_n) - \frac{\beta x_n^2}{2}\right] dt \\ &+ e^{-r\,dt}[x_n\, dt V_{n+1}(\overline{x}_n + bx_n)\, dt V_0 \\ &+ (1 - x_n dt - (\overline{x}_n + bx_n)\, dt) V_n]. \end{aligned}$$

Intuitively, the Bellman method in this context only guarantees subgame perfection in subgames starting at any state but it does not ensure subgame perfection within each state.

2.4 Steady State

We now characterize the steady state of the economy by finding the steady state of one industry and assuming that all other industries are operating at their respective steady state levels. Let μ_n denote the steady state fraction of industries with technological gap $n \geq 0$ so that we have

$$\sum_{n \geq 0} \mu_n = 1. \tag{2.28}$$

The steady state is defined as the stationarity condition whereby the inflow of industries in each state is equal to the outflow of industries from it. Since the steady state is a fixed point, it must be the case that the fraction of industries in each state is also fixed (constant). The steady state is indeed equilibrium. Suppose not, then there would exist some state in which the outflow of industries exceeds the inflow which would contradict the stationarity condition. Intuitively, there should be no pressure for more industries to leave the state than industries entering, as this would lead to a decrease in the relative share[13] of that state vis-à-vis other states. At the steady state, the structure of the states cannot change.

Thus, stationarity will imply that for any state n, the flow of industries into it should be equal to the flow out. For example, during time interval dt, in

[13]We refer to its share of industries.

$\mu_n(\overline{x}_n + bx_n)dt$ industries with technological gap $n \geq 1$ the follower catches up with the leader and thus the total flow of industries into state 0 is

$$\sum_{n\geq 0} \mu_n(\overline{x}_n + bx_n)\, dt. \tag{2.29}$$

Also, in $\mu_0(2x_0)dt$ neck-and-neck industries one firm secures a lead, and the total flow of industries out of state 0 is $2\mu_0(x_0)dt$. We thus have

$$2\mu_0 x_0 = \sum_{n\geq 0} \mu_n(\overline{x}_n + bx_n). \tag{2.30}$$

For state 1 and then for states $n \geq 2$, we have

$$\mu_1(x_1 + \overline{x}_1 + bx_1) = 2\mu_0 x_0 \tag{2.31}$$

and

$$\mu_n(x_n + \overline{x}_n + bx_n) = \mu_{n-1}x_{n-1}, \quad n \geq 2. \tag{2.32}$$

The asymptotic growth rate of the representative industry (if it exists) is given by

$$g = \lim_{t\to\infty} \frac{d}{dt} \ln Q_i(t) \tag{2.33}$$

hence

$$g = \lim_{t\to\infty} \frac{lnQ_i(t)}{t}.$$

As in AHV we say that an industry i is said to go through a $(p+1)$ cycle over a time interval $[0, t]$ if the technological gap n goes through the sequence $\{0, 1, \ldots, p-1, p, 0\}$. Since the value of $\ln Q_i(t)$ rises by $\ln \gamma^p = p \ln \gamma$, during an industries transition from state n to $n+1$, $\ln Q_i(t)$ can be approximated by

$$\ln Q_i(t) \approx (\ln \gamma) \sum_{p\geq 1} pN_p(t),$$

where $N_p = N_p(t)$ is the probability that the industry has gone through exactly $(p+1)$ cycles over the interval $[0, t]$. Thus we rewrite (2.33) as

$$g = \ln \gamma \lim_{t\to\infty} \frac{1}{t} \sum_{p\geq 1} pN_p(t), \tag{2.34}$$

where $\sum_{p\geq 1} pN_p(t)$ is the expected number of cycles over the interval $[0, t]$ which is approximately equal to $(g \ln \gamma)t$. Note that $N_p(t)$ is also equal to the fraction of industries in state p at time t times the probability that the follower catches up with the leader in such an industry. Thus, again as in AHV we have

$$g = \sum_{p\geq 1} \mu_p(\overline{x}_p + bx_p)(p \ln \gamma). \tag{2.35}$$

Equations (2.30)–(2.32) and (2.35) imply that

$$g = \left(2\mu_0 x_0 + \sum_{k\geq 1} \mu_k x_k\right)(p \ln \gamma). \tag{2.36}$$

Proof. See the Appendix. □

It is clear from (2.36) that the largest component of growth comes from the neck-and-neck state. Intuitively, this happens since there are two firms trying to advance the technology frontier in that state compared to only one in any other state. Hence, technology would advance twice as fast on average in a neck-and-neck state if all efforts were the same.

2.5 Very Large Innovative Lead

In this section, we consider the case where a one-step lead is so large that the leader has no incentives to increase its lead by more than one step. In other words, we consider the case where $\gamma \to \infty$. Thus, in this section the maximum permissible lead is one step. Consequently, $x_1 = 0$.

Proposition 2.3. *Assume that the conditions in the above game hold and that the productivity lead of the leader is large. Then (i) an industry with a relatively lower degree of appropriability does not necessarily grow at a slower rate; that is, an increase in the ease of spillovers or an improvement in the reverse engineering environment does not necessarily lead to a fall in the rate of innovation and (ii) the level of RD effort is higher in the neck-and-neck state than in the unleveled state.*

Proof. (i) Since the maximum permissible lead is one step, we have $x_1 = 0$. Thus, (2.26) and (2.27) can be rewritten as

$$\frac{x_0^2}{2} + rx_0 = \frac{\Gamma_0}{\beta}, \tag{2.37}$$

$$\frac{\overline{x}_1^2}{2} + (r + x_0)\overline{x}_1 = \frac{\Gamma_{-1}}{\beta} + \frac{x_0^2}{2}. \tag{2.38}$$

Using the fact that there are only 2 states in this case, we have, using (2.28),

$$\mu_1 = 1 - \mu_0. \tag{2.39}$$

Substituting (2.39) in (2.31), we have

$$\mu_0 = \frac{\overline{x}_1}{2x_0 + \overline{x}_1}. \tag{2.40}$$

Using the fact that for all in (2.36) and (2.40), we have the growth rate

$$g = \frac{2x_0\overline{x}_1}{2x_0 + \overline{x}_1}. \tag{2.41}$$

Now since x_0^* and $\overline{x}_1^*$ can be found by solving (2.37) and (2.38) simultaneously, we can derive g^* only in terms of the exogenous parameters. Visual inspection of (2.37), (2.38), and (2.41) show that g^* is independent of the spillover parameter b. (ii) Using the fact that $\Gamma_0 > \Gamma_{-1}$ and after some algebraic manipulation of (2.37) and (2.38), we can establish that $x_0 > \overline{x}_1$.

Part (i) of Proposition 2.3 states that the spillover parameter b does not affect growth whenever the lead of one step is large enough. Part (ii) of Proposition 2.3 states that when firms are in the leveled state they have more incentive to undertake innovation than in any other states. Thus, the usual Schumpeterian effect of more intense competition in the neck-and-neck state is outweighed by the increased incentive for firms to innovate in order to escape competition. Moreover, unlike the AHV model, our result does not depend on the product differentiability parameter. Hence, we find that a competitive environment can stimulate RD by increasing the incremental profit from innovating, that is, by strengthening the motive for neck-and-neck rivals to innovate so as to become the leader. Intuitively, since externalities are present only in the unleveled state and the RD level of the leader is zero in that state, the spillover rate in the model (for the case of large innovative lead) becomes zero too. Since growth is driven by the innovation rate in the neck-and-neck state and the imitation rate of the unleveled state, it is independent of the spillover rate. Therefore, this proposition reinforces the case put forward by AHV by showing that RD incentives are higher at the neck-and-neck state and that greater appropriability does not necessarily increase growth even when there are externalities to the leader's RD. □

2.6 Very Small Innovative Lead

In this section, we look at the extreme opposite case of the previous section; that is, the case in which the one-step lead is small and, hence, the leader does have an incentive to increase its lead by more than one step. Therefore, we look into the case where $x_1 \neq 0$. Thus, we consider the case where $\gamma \to 0$. For simplicity we assume only three states: state 0 which is the neck-and-neck-state, state 1 where the leader has a one-step lead, and state 2 where the leader has a two-step lead. The following result asserts similar results for more than two states.

Proposition 2.4. *Assume that the conditions in the above game hold and that the productivity lead of the leader is small. Then growth rate and process spillovers are positively related; that is, an increase in the spillover rate unambiguously leads to an increase in the growth rate.*

Proof. See the Appendix. □

Proposition 2.4 states our main result here. It gives us an important long-run relationship between imitation rate and growth at the steady state without making the assumption of a large technology gap (large innovative lead). In this Markovian game between the duopolists, with imitation and innovation as strategic variables, we observe that, at any point in time, an increase in imitation rate will always prompt the leader to increase its innovation rate in equilibrium. Thus, process imitation creates a source of competitive pressure that deters the leader from maximizing short-run monopoly profit and instead "forces" him to innovate further. The mechanism driving this result can be observed from the construction of the proof. For a small productivity lead, we show that the relationship between the RD effort of the follower and the spillover rate is negative since the laggard has less incentive to innovate when it can feed off the leader's effort. Since we also show that the RD efforts of the two rivals in the unleveled state are inversely related at the steady state, the leader's innovation must be positively related to the level of externalities.

Thus, an increase in the spillover rate reduces the effort of the follower as it can free-ride on the leader who, by receiving the signal that his technological advantage is shrinking, makes an effort to restore its lead. Proposition 2.4 implies that the policymaker can enhance the economy's growth rate by choosing a lower level of appropriability. Hence, there is always (at any point in time) a follower who will prompt the leader to innovate further in such a market configuration, and this will lead to higher growth. Also note that the above phenomenon might be due to increasing returns on the RD when the gap is large. According to Glass [8], an important factor in Japan's recent economic slowdown has been the exhaustion of all imitation possibilities as they move closer to the world's technology frontier.

3 Conclusion

We have presented an analytical model that deals with process imitation and spillovers in a non-Schumpeterian framework. Our motivation stems mainly from the fact that the existing non-Schumpeterian models depend heavily on the level of competition in showing the interrelation between process imitation and spillovers and their impact on growth. Moreover, existing Schumpeterian models lack adequate empirical evidence to explain growth using the concept of creative destruction. Indeed, most of these studies rely heavily on the price undercutting mechanism of the homogeneous Bertrand game. We demonstrate, without relaxing the assumption of product homogeneity, that competitive behavior can still prevail by using a Cournot quantity competition setting. Two main factors drive competitive behavior in the long run; first, the RD level in the neck-and-neck state and, second, spillovers occurring due to a lack of appropriability.

Moreover, this chapter can offer a basis for understanding how the dynamic strategic interactions between two firms with a technology gap can determine the economy's growth rate when there is uncertainty. In particular, imitation acts as a spur by putting pressure on the industry leader to innovate further, and this drives the economy's engine of growth. Furthermore, this research can contribute to the literature on "the law and economics of reverse engineering" (see [13]) by providing some economic grounds in favor of process reverse engineering.

In this regard, it demonstrates the existence of a non-Schumpeterian element in the innovator's best response function. One immediate policy implication of our model is that relaxing laws and regulations that hinder process imitation might not always be a good thing in an industry characterized by spillovers.

4 Appendix

4.1 Derivation of the Second Stage Quantity and Profit Functions and Proof of Remark 2.2

In this section we derive (2.12), (2.13), and Remark 2.2. From (2.11), firm j's problem is given by

$$\max_{q_j} \left(\frac{Mq_j}{q_i + q_j} - c_j q_j \right).$$

First-order condition for firm j is given by

$$(q_i + q_j)M - q_j M - c_j (q_j + q_i)^2 = 0. \tag{4.1}$$

By symmetry we have

$$(q_j + q_i)M - q_i M - c_i (q_i + q_j)^2 = 0. \tag{4.2}$$

Simplifying gives

$$q_i M - c_j (q_j + q_i)^2 = 0 \tag{4.3}$$

$$q_j M - c_i (q_i + q_j)^2 = 0. \tag{4.4}$$

Solving (4.3) and (4.4) simultaneously gives (2.12). Substituting (2.12) for both firms in (2.4) gives P_t:

$$P_t - c_j = \frac{c_i}{(c_i + c_j)^2}. \tag{4.5}$$

Thus, the profit for firm j is given by

$$(P_t - c_j) q_j = \pi_j. \tag{4.6}$$

(4.6) verifies (2.13). Using (2.6) and (2.13) for firm 1 we have

$$\pi_1 = \frac{M\left(\frac{1}{A_2}\right)^2}{\left(\frac{1}{A_2} + \frac{1}{A_1}\right)^2} \tag{4.7}$$

$$\pi_2 = \frac{M\left(\frac{1}{A_1}\right)^2}{\left(\frac{1}{A_2} + \frac{1}{A_1}\right)^2}. \tag{4.8}$$

Now using (2.7), (2.8) on (4.7) and (4.8) and differentiating the result w.r.t. γ establishes the first part of Remark 2.2. Part (ii) of Remark 2.2 is obtained by replacing $\gamma = 1$ in (2.7), (2.8) and hence in (4.7) and (4.8). Part (ii) of Remark 2.2 holds since

$$\frac{1+(\gamma^n)^2}{(1+\gamma^n)^2} > 1/2. \tag{4.9}$$

□

4.2 Derivation of the Steady State Growth Rate (2.36)

Note that (2.35) can be rewritten as

$$g = \ln \gamma \sum_{p\geq 1} \mu_p(\overline{x}_p + bx_p)p. \tag{4.10}$$

Moreover,

$$\sum_{p\geq 1} \mu_p(\overline{x}_p + bx_p)p = \sum_{p\geq 1} \mu_p(\overline{x}_p + bx_p) + \sum_{p\geq 2} \mu_p(\overline{x}_p + bx_p) + \cdots \tag{4.11}$$

The first term of the RHS of (4.11) is $2mu_0x_0$ from (2.30). Now taking the summation on both sides of (2.32) and rearranging, we have

$$\sum_{p\geq k} \mu_p(\overline{x}_p + bx_p)p = \mu_{k-1}x_{k-1}. \tag{4.12}$$

Substituting (4.12) in (4.11) and then in (4.10), we have (2.36). □

4.3 Proof of Proposition 2.4

Assume that $n = 0, 1, 2$. We first derive the fraction μ of industries in state 2. We know that

$$\mu_0 + \mu_1 + \mu_2 = 1 \text{ or } \mu_0 = 1 - \mu_1 - \mu_2. \tag{4.13}$$

Using (2.31) and (2.32) we have the stationarity condition

$$\mu_1(x_1 + \overline{x}_1 + bx_1)p = 2\mu_0x_0. \tag{4.14}$$

Furthermore,

$$\mu_1(x_2 + \overline{x}_2 + bx_2)p = 2\mu_1x_1. \tag{4.15}$$

Solving (4.13)–(4.15) simultaneously gives

$$\mu_2^* = \frac{2x_0x_1}{2x_0x_1 + [(1+b)x_1 + \overline{x}_1 + 2x_0][(1+b)x_2 + \overline{x}_2]}. \tag{4.16}$$

Thus,

$$\mu_1 = 1 - \mu_0 - \mu_2^*. \tag{4.17}$$

Now by solving for μ_0 as we did for Proposition 2.3, we derive the growth rate also in a similar fashion, and it is given by[14]

$$g = \frac{2x_0[(1+b)x_1 + \overline{x}_1][1-\mu_2^*]}{2x_0 + \overline{x}_1 + (1+b)x_1}. \tag{4.18}$$

Now it can be shown that the partial derivatives of g in (4.18) w.r.t. to b, x_0, x_1, and $\overline{x}_1$ are all positive as long as the partial derivatives of μ_2^* w.r.t. to b, x_0, x_1, and $\overline{x}_1$ are small enough. We next solve (2.26) and (2.27) to find $\overline{x}_1^*$ in terms of b, x_0, and other exogenous parameters only. We thus have

$$\overline{x}_1^* = \frac{\sqrt{\Omega^2 + 4\Lambda} - \Omega}{2}, \tag{4.19}$$

where

$$\Omega = 2\left(r + x_0 + b\sqrt{x_0^2 + 2 + x_0 - \frac{2\Gamma_0}{\beta}}\right)$$

and $\Lambda = \Gamma_{-1} + x_0^2$. It can be shown that the partial derivatives of $\overline{x}_1^*$ w.r.t. b, x_0, and x_1 are negative[15] and, thus, by the chain rule the partial derivatives of x_0^* and x_1^* w.r.t. b must be positive. We can now find the total change in the growth rate g as a result of a change in b:

$$\operatorname{sgn}\left(\frac{dg}{db}\right) = \operatorname{sgn}\left\{\frac{dg}{dx_0} \times \frac{dx_0}{db} + \frac{dg}{dx_1} \times \frac{dx_1}{db} + \frac{dg}{d\overline{x}_1} \times \frac{d\overline{x}_1}{db} + \frac{dg}{db}\right\}. \tag{4.20}$$

It can be shown that the RHS of (4.20) is always positive as long as $x_1 \geq \frac{d\overline{x}_1}{db}$, which we can reasonably impose as a restriction. □

Acknowledgments

This chapter is a substantially revised version of a chapter of the author's Ph.D. thesis. I would like to thank Julian Wright and Mark Donoghue for their encouragement and helpful suggestions. Many thanks are also due to Wayne Lawton for his help in editing the material. I am grateful to two anonymous referees of this journal for their invaluable comments and suggestions. I am also thankful to the editor Steffen Jørgensen.

[14] For the case where the number of states is greater than 2, we replace μ_2^* in (4.18) by $\sum_{n=2}^{N} \mu_n$. Thus, the proof can be extended for the case of more than one state.

[15] Note that although x_1 is not present in (4.19), we can still deduce this relationship since we know from (2.26) that x_1 and x_0 are positively related.

References

[1] P. Aghion, C. Harris, P. Howitt, J. Vickers, *Competition, Imitation and Growth with Step by Step Innovation.* European Economic Review **41** (1997), 771–782.

[2] P. Aghion, C. Harris, P. Howitt, J. Vickers, *Competition, Imitation and Growth with Step by Step Innovation.* Review of Economic Studies **68** (2001), 467–492.

[3] C. D'Aspremont, R. Ferreira, L. Varet, *Strategic RD Investment, Competition Toughness and Growth.* Unpublished (2002).

[4] C. D'Aspremont, A. Jacquemin, *Cooperative and Non cooperative RD in Duopoly with Spillovers.* American Economic Review **78** (1988), 1133–1137.

[5] R. Blundell, R. Griffith, J. Reenen, *Dynamic Count Data Models of Technological Innovation.* Economic Journal **105** (1995), 333–344.

[6] W. Cohen, D. Levinthal, *Innovation and Learning: The Two Faces of RD.* Economic Journal **99** (1989), 569–596.

[7] E. Dockner, V. N. Long, S. Gerhard, S. Jorgensen, *Differential Games in Economics and Management Science.* Cambridge University Press (2000).

[8] A. J. Glass, *Imitation as a stepping stone to Innovation.* Working Paper, Department of Economics, Ohio State University (2000).

[9] Z. Griliches, *Issues in Assessing the Contribution of Research and Development to Productivity Growth.* Bell Journal of Economics **10** (1979), 92–116.

[10] G. Grossman, H. Helpman, *Quality Ladders in the Theory of Growth.* Review of Economic Studies **58** (1991), 43–61.

[11] G. Grossman, H. Helpman, *Quality Ladders and Product Cycles.* The Quarterly Journal of Economics **106** (1991), 92–116.

[12] S. Nickell, D. Nicolitsas, N. Dyden, *Competition and Corportate Performance.* Journal of Political Economy **104** (1996), 724–46.

[13] P. Samuelson, S. Scotchmer, *The Law and Economics of Reverse Engineering.* Yale Law Journal (forthcoming).

[14] J. A. Schumpeter, *The Theory of Economic Development.* Oxford University Press (1934).

[15] S. Sergerstrom, *Innovation, Imitation and Economic Growth.* The Journal of Political Economy **99** (1991), 807–827.

[16] M. Zachariadis, *Research and Development, Innovation, and Technological Progress: A Test of the Schumpeterian Framework without Scale Effects.* Canadian Journal of Economics, **36** (2003), 566–586.

Supplier-Manufacturer Collaboration on New Product Development

Bowon Kim
KAIST Graduate School of Management
207-43 Cheongryangri
Dongdaemoon-Ku
Seoul 130-722, Korea
bowonKim@kgsm.kaist.ac.kr

Fouad El Ouardighi
ESSEC Business School
Avenue Bernard Hirsch
B.P. 105
95021 Cergy Pontoise, France
elouardighi@essec.fr

Abstract

In this chapter, we develop a dynamic model of collaboration between a manufacturer and its supplier, where the fundamental issue is, for each player, how to allocate own resources between improving an existing product and developing a new one. We study the optimal time path of effort allocation for each player in a noncooperative setting, and then look into the cooperative setting. We finally conduct a numerical analysis and observe how the optimal dynamics of key variables change as parameters of the game vary.

Key words. Quality improvement, new product development, supply chain management, differential games.

1 Introduction

Collaboration between supplier and manufacturer is prevalent and can be of diverse forms [1, 3, 6, 9]. A supplier often collaborates with its customer, e.g., a manufacturing company, on developing new products. But there is a potential trade-off in their endeavor. On one hand, there needs to be a short-term collaboration between the supplier and the manufacturer on the manufacturer's current product. For instance, in order to gain a larger market share for the existing product, it would be important for the product quality to improve and/or for its price to decrease. To do so, the manufacturer expects its supplier to support its efforts to improve its existing product value. On the other hand, the manufacturer needs its supplier's help in developing a new product, which is necessary for the manufacturing company to continue to grow in the market. Compared with the joint efforts to improve the existing product's quality, collaborating for new product development is a long-term endeavor.

Consider, for example, the case of SKT, a Korean telecommunications company. In the early 1990s, SKT had struggled to make a transition from an analog to a digital mobile phone system. While it was using the analog technology, the company tried to improve the telecommunications quality with help from its key suppliers. At the same time, since SKT knew the new digital technology would replace the old one, it started to invest in developing the digital technology and asked those suppliers who were also involved in improving the quality of the analog telecommunications system to participate in the project. Then, the relevant decision problem was how to allocate resources between the current technology (product quality) improvement and the development of new technology (product). This was the problem faced by not only SKT, but also its supplier(s).

Now the fundamental issue facing the supplier and the manufacturer respectively is how to allocate resources between improving the existing product and developing a new one. By improving the existing product, they expect to earn more *current profit*, i.e., their short-term profitability will increase. By enhancing the chance to have a successful new product development, they can expect to earn *future profit*, which is vital for their continuous growth. Since they have to allocate their *present* resources between *current* operations and *future* possibility, they are facing a decision problem of trade-off.

In this chapter, we formulate this problem as a differential game model with two players, a manufacturer and its supplier. In Section 2, we design the differential game model in a noncooperative setting, i.e., assuming a competitive relationship between the manufacturer and the supplier. In Section 3, we analyze the qualitative solutions for the noncooperative game. In Section 4, we look into the cooperative setting, where the manufacturer and the supplier make decisions as if they were a single firm, i.e., a supply chain. To gain further insights into the optimal dynamics, we conduct a numerical analysis in Section 5. We observe how the optimal dynamics of key variables change as the game parameters vary. Finally, we draw conclusions and managerial implications from our qualitative and numerical analyses.

2 Model Formulation

To date, the differential game literature on new product development has focused on racing issues. Early contributions are most notably those of Reinganum [10, 11] and Kamien and Schwartz [8]: they considered dynamic games of innovation involving two competitors, based on the Nash equilibrium in innovation investment strategies.

In this chapter, we consider a chain with two firms, a manufacturer and its supplier, both involved in two major activities, respectively the improvement of quality of an existing product and the development of a new product. We first describe the dynamics of the quality improvement activity. Let $q(t) > 0$ denote the quality level at time t of the product on the market (i.e., current product). Quality is a state variable incremented by the respective efforts of the manufacturer, $u_1(t)$, and the supplier, $u_2(t)$, that is,

$$\dot{q}(t) = g(u_1(t), u_2(t)), \quad q(0) = q_0 > 0.$$

For the dynamics of new product development (NPD), let $z(t) \geq 0$ be the cumulative level of knowledge at time t on a product in development (i.e., future product). The NPD activity is also a state variable incremented by the respective contributions of both players, $v_i(t)$, $i = (1, 2)$, that is,

$$\dot{z}(t) = h(v_1(t), v_2(t)), \quad z(0) = z_0 \geq 0.$$

Under negative exponential distribution, the breakthrough cumulative probability at t is given by $\phi(z(t)) = 1 - e^{\lambda z(t)}$, which indeed depends on the cumulative effort z. Moreover, $\frac{d\phi(z(t))/dt}{(1-\phi(z))} = \lambda \dot{z}$ is the conditional probability of completion, which is a standard setting as we are dealing with instantaneous probabilities [7, 11].

Using a finite planning horizon, i.e., $t \in [0, T]$, the objective criterion of the manufacturer is

$$\begin{aligned}\underset{p_1,u_1,v_1}{\text{Max}}\ J^1 = \int_0^T & \{[(p_1(t) - p_2) f(p_1(t), q(t)) - c_1(u_1(t)) - d_1(v_1(t))] \\ & \times (1 - \phi(z(t))) + \omega \phi_z \dot{z}(t) N\}\, dt + \theta_1 z(T).\end{aligned}$$

For simplicity, we assume a short enough planning horizon so that the discounting effect can be disregarded. At time t, provided the new product is not completed yet, the manufacturer's net profit from the existing product is

$$(p_1 - p_2) f(p_1, q) - c_1(u_1) - d_1(v_1),$$

with probability $(1 - \phi(z))$. It is composed of the following elements: gross profit margin $(p_1 - p_2)$, that is, sales price minus (constant) transfer price to the supplier, times sales of the existing product, $f(p_1, q)$, minus expense to improve the existing product, $c_1(u_1)$, minus expense to develop the new product, $d_1(v_1)$. The assumption of a constant transfer price follows from the fact that when manufacturers and suppliers decide to collaborate, the transfer price to the supplier is one of the terms that are usually fixed at the time of signing the contract, especially when the contract period is short term. We assume that the demand function $f(p_1, q)$ is influenced negatively by price, i.e., $f_{p_1} < 0$, and positively by quality, i.e., $f_q > 0$.

In addition, with probability [7]

$$d\phi(z(t)) = \phi_z \dot{z}(t) dt = \phi_z h(v_1(t), v_2(t))\, dt,$$

the new product will become available during the time interval $(t, t + dt)$, providing a future utility stream with discounted value ωN at time t. The constant ω, $0 \le \omega \le 1$ (resp., $0 \le 1 - \omega \le 1$) denotes the constant rate which should be allocated to the manufacturer (resp., the supplier). On the other hand, N represents the total future benefits from the new product, from the (unknown) time it becomes available forward. We assume that the players have an *a priori* knowledge about the size of the lump-sum benefit from the new product development success, so that N may be considered as constant. Note that the manufacturer and the retailer agree beforehand on each one's share in the revenues of the new product. This refers to the fact that partnering companies generally want to design their collaboration contract in a predictable manner as much as possible. Predetermining the profit sharing scheme is part of such an effort. For our analysis, we first define ω *ex ante* and analyze our model. Then, we study how the changes in ω affect the optimal dynamics while conducting the numerical analysis. Finally, letting $\theta_1 \ge 0, \theta_1 z(T)$ represents the salvaged value of the stock of knowledge at the end of the planning horizon.

Using similar arguments, the supplier's problem is

$$\begin{aligned}\underset{u_2,v_2}{\text{Max}} J^2 = \int_0^T &\{[p_2 f(p_1(t), q(t)) - c_2(u_2(t)) - d_2(v_2(t))] \\ &\times (1 - \phi(z(t))) + (1 - \omega)\phi_z \dot{z}(t) N\}\, dt + \theta_2 z(T).\end{aligned}$$

Then the manufacturer controls its sales price and its quality improvement and knowledge accumulation efforts, while the supplier controls its quality improvement and knowledge accumulation efforts.

For the demand function, the following multiplicative expression is introduced:

$$f(p_1(t), q(t)) = eq(t)(a - bp_1(t)),$$

$a, b, e > 0$, where demand is linearly decreasing with price and increasing with quality.

On the other hand, for the probability distribution function, we assume that the probability of successful development by a given date is known to both players and should be given by a negative exponential distribution [11], that is, $\phi(z(t)) = 1 - e^{-\lambda z(t)}$, $\lambda > 0$, and $\frac{d\phi(z(t))}{dt} = \lambda \dot{z}(t) e^{-\lambda z(t)}$. Thus, the conditional probability that the new product is completed immediately beyond time t given no successful development until time t, i.e., the hazard rate, is $\frac{\lambda \dot{z}(t) e^{-\lambda z(t)}}{1 - \phi(z(t))} = \lambda \dot{z}(t)$.

In addition, for tractability, the following linear additive specifications are respectively introduced:

$$g(u_1(t), u_2(t)) = \sum_{i=1}^{2} \alpha_i u_i(t),$$

$$h(v_1(t), v_2(t)) = \sum_{i=1}^{2} \beta_i v_i(t),$$

where $\alpha_i > 0, i = (1, 2)$, is the marginal impact of player i's effort to the quality improvement process, and $\beta_i > 0, i = (1, 2)$, is the marginal influence of player i's contribution to the development process.

Finally, we assume convex increasing cost functions for both players, with explicit functions respectively for quality improvement effort,

$$c_i(u_i(t)) = \frac{c_i(u_i(t))^2}{2},$$

$c_i > 0, i = (1, 2)$, and for new product development effort,

$$d_i(v_i(t)) = \frac{d_i(v_i(t))^2}{2},$$

$d_i > 0, i = (1, 2)$.

Hence, the manufacturer's problem is written as

$$\int_0^T \left\{ \left[(p_1(t) - p_2)(a - bp_1(t))eq(t) - \frac{c_1(u_1(t))^2}{2} \right.\right.$$
$$\left.\left. - \frac{d_1(v_1(t))^2}{2} \right] + \lambda\omega\dot{z}(t)N \right\} \times e^{-\lambda z(t)}dt + \theta_1 z(T),$$

s.t.

$$\dot{q}(t) = \sum_{i=1}^{2} \alpha_i u_i(t), \quad q(0) = q_0 > 0,$$

$$\dot{z}(t) = \sum_{i=1}^{2} \beta_i v_i(t), \quad z(0) = z_0 \geq 0.$$

To analyze the game, it is convenient to introduce the state transformation

$$y(t) = e^{-\lambda z(t)},$$

where $y(t) \in [0, 1]$.

Differentiating both sides of this equation with respect to time, we get

$$\dot{y}(t) = -\lambda y(t)\dot{z}(t) = -\lambda y(t) \sum_{i=1}^{2} \beta_i v_i(t).$$

The manufacturer's problem can now be rewritten in terms of the new state variable:

$$\underset{p_1,u_1,v_1}{\text{Max}}\ J^1 = \int_0^T \left\{ \left[(p_1(t) - p_2)(a - bp_1(t))eq(t) - \frac{c_1(u_1(t))^2}{2} - \frac{d_1(v_1(t))^2}{2} \right] \right.$$
$$\left. \times\ y(t) - \omega\dot{y}(t)N \right\} dt - \frac{\theta_1}{\lambda} \ln y(T),$$

s.t.

$$\dot{q}(t) = \sum_{i=1}^{2} \alpha_i u_i(t), \quad q(0) = q_0 > 0,$$

$$\dot{y}(t) = -\lambda y(t) \sum_{i=1}^{2} \beta_i v_i(t), \quad y(0) = y_0 \in [0, 1).$$

3 Noncooperative Game

The game defined above combines linear state and exponential properties [4], which results in an interaction between the state variables. This particular structure precludes any useful characterization of closed-loop equilibrium strategies. In this sense, we confine our interest to an open-loop Nash equilibrium, which offers more tractable and flexible implications.

Skipping the time index for simplicity, the Hamiltonian of the manufacturer is

$$H^1 = \left\{ (p_1 - p_2)(a - bp_1)eq - \frac{c_1 u_1^2}{2} - \frac{d_1 v_1^2}{2} + \lambda[\omega N - \mu_2^1] \sum_{i=1}^{2} \beta_i v_i \right\} y + \mu_1^1 \sum_{i=1}^{2} \alpha_i u_i, \tag{1}$$

where μ_j^1(t) are the costate variables, $j = (1, 2)$.

Assuming interior solutions, necessary conditions for optimality include

$$H_{p_1}^1 = (a - 2bp_1 + bp_2)eqy = 0, \tag{2}$$

$$H_{u_1}^1 = -c_1 u_1 y + \mu_1^1 \alpha_1 = 0, \tag{3}$$

$$H_{v_1}^1 = \{\lambda \beta_1 [\omega N - \mu_2^1] - d_1 v_1\} y = 0. \tag{4}$$

Lemma 3.1. *An optimal noncooperative price satisfies*

$$p_1 = \frac{1}{2}\left[p_2 + \frac{a}{b}\right]. \tag{5}$$

Proof. The proof is obvious from (2). □

This is a myopic rule that is inversely related to the marginal influence of transfer price on demand. The sales price increases with transfer price to the supplier, and this rule amounts to pricing from the transfer price to the supplier.

Substituting (5) for optimal sales price into (1), the costate variables have the following respective dynamics:

$$\dot{\mu}_1^1 = -\frac{e}{4b}(a - bp_2)^2 y, \tag{6}$$

$$\dot{\mu}_2^1 = -\left[\frac{e}{4b}(a - bp_2)^2 q - \frac{c_1 u_1^2}{2} - \frac{d_1 v_1^2}{2}\right] - \lambda[\omega N - \mu_2^1]\sum_{i=1}^{2}\beta_i v_i, \tag{7}$$

$\mu_1^1(T) = 0$ and $\mu_2^1(T) = -\frac{\theta_1}{\lambda y(T)}$ being the transversality conditions.

In equation (6), $\dot{\mu}_1^1 \le 0$ and $\mu_1^1(T) = 0$ both imply $\mu_1^1(t) \ge 0 | t \le T$, which indicates that the improvement activity has a positive influence on the manufacturer's profit. Note that $y(t) \mapsto 1$ if $z(t) \mapsto +\infty$, which is unlikely to be verified over a (short) finite time horizon. Solving for $u_1(t)$ in (3) should then result in $u_1(T) = 0$, that is, $u_1(t)$ should be decreasing over time.

In equation (7), for a nonnegative instantaneous profit, $\dot{\mu}_2^1|_{\mu_2^1 \le 0} < 0$. Given the transversality condition $\mu_2^1(T) = -\frac{\theta_1}{\lambda y(T)}$, the sign of $\mu_2^1(t)$ is ambiguous for $\theta_1 > 0$. However, for $\theta_1 = 0$, $\mu_2^1(T) = 0$, which results in $\mu_2^1(t) \ge 0 | t \le T$. Solving for $v_1(t)$ in (4) yields $v_1(t) = \frac{\lambda\beta_1}{d_1}[\omega N + \frac{\theta_1}{\lambda y(T)}] > 0$, $\theta_1 \ge 0$, which shows that $v_1(t)$ should be increasing over time until it reaches its maximum value at $t = T$.

We now turn to the supplier's problem. Using (5) for sales price, the supplier's Hamiltonian is

$$H^2 = \left\{\frac{p_2}{2}(a - bp_2)eq - \frac{c_2 u_2^2}{2} - \frac{d_2 v_2^2}{2} + \lambda[(1 - \omega)N - \mu_2^2] \times \sum_{i=1}^{2}\beta_i v_i\right\} y + \mu_1^2 \sum_{i=1}^{2}\alpha_i u_i, \tag{8}$$

where $\mu_j^2(t)$ are the costate variables, $j = (1, 2)$.

Assuming interior solutions, necessary conditions for optimality include

$$H_{u_2}^2 = -c_2 u_2 y + \mu_1^2 \alpha_2 = 0, \tag{9}$$

$$H_{v_2}^2 = \{\lambda\beta_2[(1 - \omega)N - \mu_2^2] - d_2 v_2\} y = 0. \tag{10}$$

The costate variables have the following respective dynamics:

$$\dot{\mu}_1^2 = -\frac{ep_2}{2}(a - bp_2)y, \tag{11}$$

$$\dot{\mu}_2^2 = -\left[\frac{p_2}{2}(a - bp_2)eq - \frac{c_2 u_2^2}{2} - \frac{d_2 v_2^2}{2}\right] - \lambda[(1 - \omega)N - \mu_2^2]\sum_{i=1}^{2}\beta_i v_i, \tag{12}$$

$\mu_1^2(T) = 0$ and $\mu_2^2(T) = -\frac{\theta_2}{\lambda y(T)}$ being the transversality conditions. In equation (11), $\dot{\mu}_1^2 \leq 0$ and $\mu_1^2(T)$ both provide $\mu_1^2(t) \geq 0$ whatever $t \in [0, T]$, which means that the improvement process has a positive marginal effect on the supplier's profit. As for the manufacturer, solving for $u_2(t)$ in (9) should result in $u_2(T) = 0$, i.e., $u_2(t)$ should also be decreasing over time.

In equation (12), given a nonnegative instantaneous profit, we obtain $\dot{\mu}_2^2|_{\mu_2^2 \leq 0} < 0$. Given the transversality condition $\mu_2^2(T) = -\frac{\theta_2}{\lambda y(T)}$, the sign of $\mu_2^2(t)$ is unclear. However, for $\theta_2 = 0$, $\mu_2^2(T) = 0$, which results in $\mu_2^2(t) \geq 0 | t \in [0, T]$. Solving for $v_2(t)$ in (10), $v_2(T) = \frac{\lambda\beta_2}{d_2}[(1-\omega)N - \frac{\theta_2}{\lambda y(T)}] > 0$, $\theta_2 \geq 0$, which shows that $v_2(t)$ should be increasing over time until a maximum value is reached at $t = T$.

Proposition 3.1. *Qualitative solutions for the noncooperative improvement effort of the manufacturer and the supplier are respectively,*

$$\dot{u}_1 = \lambda u_1 \sum_{i=1}^{2} \beta_i v_i - \frac{e\alpha_1}{4bc_1}(a - bp_2)^2, \quad u_1(T) = 0, \tag{13}$$

$$\dot{u}_2 = \lambda u_2 \sum_{i=1}^{2} \beta_i v_i - \frac{e\alpha_2 p_2}{2c_2}(a - bp_2), \quad u_2(T) = 0. \tag{14}$$

Proof. Equations (3) and (9) are differentiated totally with respect to time to yield, after elementary manipulations, (13) and (14). □

In (13) and (14), as $\dot{u}_i|_{u_i=0} < 0$ and $u_i(T) = 0$, $i = (1, 2)$, the optimal improvement effort of both players is decreasing over time. The players make an effort to improve the current product quality more during the early period than in the later period since, as time passes, it becomes more likely to successfully complete the new product development. The decline of the players' improvement effort is due to the negative influence of marginal demand with respect to quality. However, the rate of change of the players' improvement effort is positively affected by the speed of NPD. In other words, the innovation effort by both players slows down the decrease of each player's improvement effort. Given $u_i(T) = 0$, a negative influence of both players' innovation effort on each player's improvement effort is obtained (Figure 1).

In the same way, a marginal increase of the supplier's transfer price reduces the manufacturer's improvement effort, i.e., $\dot{u}_{1p_2} = \frac{e\alpha_1}{2c_1}(a - bp_2) > 0$. Conversely, for sufficiently small values of b, the supplier's transfer price has a positive impact on its own improvement effort, i.e., $\dot{u}_{2p_2} = -\frac{e\alpha_2}{2c_2}(a - 2bp_2) < 0$. This result is expected since there is a tension between the manufacturer and the supplier regarding the transfer price, i.e., the higher the transfer price, the smaller the manufacturer's profit and the larger the supplier's profit, and *vice versa*.

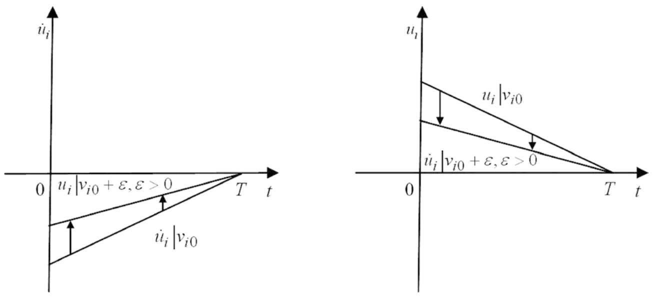


Figure 1: Marginal impact of players' innovation effort on each one's own improvement effort.

Proposition 3.2. *Qualitative solutions for the noncooperative innovation effort of the manufacturer and the supplier are respectively*

$$\dot{v}_1 = \lambda \left\{ v_1 \left[\frac{\beta_1 v_1}{2} + \beta_2 v_2 \right] + \frac{\beta_1}{d_1} \left[\frac{1}{4b}(a - bp_2)^2 eq - \frac{c_1 u_1^2}{2} \right] \right\},$$
$$v_1(T) = \frac{\lambda \beta_1}{d_1} \left[\omega N + \frac{\theta_1}{\lambda y(T)} \right] > 0, \tag{15}$$
$$\dot{v}_2 = \lambda \left\{ v_2 \left[\beta_1 v_1 + \frac{\beta_2 v_2}{2} \right] + \frac{\beta_2}{d_2} \left[\frac{p_2}{2}(a - bp_2) eq - \frac{c_2 u_2^2}{2} \right] \right\},$$
$$v_2(T) = \frac{\lambda \beta_2}{d_2} \left[(1-\omega) N + \frac{\theta_2}{\lambda y(T)} \right] > 0. \tag{16}$$

Proof. Equations (4) and (10) are differentiated totally with respect to time to yield, after elementary manipulations, (15) and (16). □

As expected, the larger the sharing ratio of the new product profit, ω, the higher (lower) the manufacturer's (supplier's) profitability from the new product. In equations (15) and (16), each player's innovation effort increases over time with the speed of knowledge accumulation. We can provide two intuitive explanations for this result. First, as time passes, the probability to have the innovation increases, and thus the dollar spent on the innovation becomes more valuable than that on the existing product's quality improvement. Second, the benefit the firm can get from selling the existing product is limited and will diminish as time passes. Thus, as time progresses, the marginal value of the existing product decreases, whereas that of the innovation increases.

In addition, the speed of each player's innovation effort depends positively on its own instantaneous profit from the current product. In this sense, the higher each player's "current" profit, the faster its own search for "future" profit. In other words, as time passes, it becomes more likely for NPD to be accomplished. As for quality improvement efforts, the speed of innovation effort by the manufacturer (supplier) is negatively (positively) affected by the supplier's transfer price.

4 Cooperative Game

We now look at the cooperative setting in which the manufacturer and its supplier intend to coordinate the two activities (quality improvement and innovation process) as an essential part of a supply chain. The combination of individual players' objective criteria provides the following optimal control problem:

$$\begin{aligned}\underset{p_1,u_1,u_2,v_1,v_2}{\text{Max}} J = \int_0^T \Bigg\{ \Bigg[& p_1(t)(a - bp_1(t))eq(t) - \sum_{i=1}^{2} \frac{c_i(u_i(t))^2}{2} \\ & - \sum_{i=1}^{2} \frac{d_i(v_i(t))^2}{2} \Bigg] y(t) - \dot{y}(t)N \Bigg\} dt - \frac{\theta}{\lambda} \ln y(T),\end{aligned}$$

s.t.

$$\dot{q}(t) = \sum_{i=1}^{2} \alpha_i u_i(t), \quad q(0) = q_0 > 0,$$

$$\dot{y}(t) = -\lambda y(t) \sum_{i=1}^{2} \beta_i v_i(t), \quad y(0) = y_0 \in [0, 1),$$

where $\theta = \sum_{i=1}^{2} \theta_i$.

Skipping the time index for simplicity, the Hamiltonian associated to the coordinated chain is

$$\begin{aligned}H = & \left\{ p_1(a - bp_1)eq - \sum_{i=1}^{2} \frac{c_i u_i^2}{2} - \sum_{i=1}^{2} \frac{d_i v_i^2}{2} + \lambda[N - \mu_2] \sum_{i=1}^{2} \beta_i v_i \right\} \\ & \times y + \mu_1 \sum_{i=1}^{2} \alpha_i u_i, \end{aligned} \tag{17}$$

where μ_j are the costate variables, $j = (1, 2)$.

Assuming interior solutions, necessary conditions for optimality include

$$H_{p_1} = [a - 2bp_1]eqy = 0, \tag{18}$$

$$H_{u_i} = -c_i u_i y + \mu_1 \alpha_i = 0, \tag{19}$$

$$H_{v_i} = \{\lambda \beta_i [N - \mu_2] - d_i v_i\} y = 0. \tag{20}$$

Lemma 4.1. *The optimal cooperative price is lower than the optimal noncooperative price.*

Proof. Solving for p_1 in (18) and comparing with (5) yields

$$p_1^{\Theta} = \frac{a}{2b} < \frac{1}{2}\left[p_2 + \frac{a}{b}\right], \tag{21}$$

where the superscript Θ is for the cooperative solution. □

This result states that the coordination of a supply chain results in a higher consumer's surplus, i.e., a larger demand for the current product, than in a decentralized decision-making system. This reflects a mitigation of the double marginalization practice which is typical in uncoordinated chains [2, 5].

Lemma 4.2. *Player i's relative optimal quality improvement effort is*

$$\frac{u_i}{u_{-i}} = \frac{c_{-i}\alpha_i}{c_i\alpha_{-i}}, \tag{22}$$

where the subscript $-i$ is for the other player.

Proof. The proof is obvious from (19). □

In (22), player i's relative effort has a constant value that depends on the ratio of player i's efficiency (α_i/c_i) to the other player's efficiency (α_{-i}/c_{-i}) in improving the current product quality. What this rule amounts to is that the more efficient player contributes more to the quality improvement activity.

Lemma 4.3. *Player i's relative optimal innovation effort is*

$$\frac{v_i}{v_{-i}} = \frac{d_{-i}\beta_i}{d_i\beta_{-i}}, \tag{23}$$

where the subscript $-i$ is for the other player.

Proof. The proof is obvious from (20). □

Here also, player i's relative effort is constant and depends on the ratio of its own efficiency (β_i/d_i) to the other player's efficiency (β_{-i}/d_{-i}) in NPD. In asymmetric conditions, the more efficient player contributes more to the innovation activity.

Substituting (21) for optimal price, the costate variables have the following dynamics:

$$\dot{\mu}_1 = -\frac{a^2ey}{4b}, \tag{24}$$

$$\dot{\mu}_2 = -\left[\frac{a^2eq}{4b} - \sum_{i=1}^{2}\frac{c_iu_i^2}{2} - \sum_{i=1}^{2}\frac{d_iv_i^2}{2}\right] - \lambda[N-\mu_2]\sum_{i=1}^{2}\beta_iv_i, \tag{25}$$

$\mu_1(T) = 0$ and $\mu_2(T) = -\frac{\theta}{\lambda y(T)}$ being the transversality conditions.

In equation (24), $\dot{\mu}_1 \leq 0$ and $\mu_1(T) = 0$ imply $\mu_1(t) \geq 0|t \leq T$. As for the noncooperative setting, $u_i^{\Theta}(T) = 0$ should result from (19), that is, the cooperative improvement effort of player i, $u_i^{\Theta}(t)$, should be decreasing over time.

In equation (25), for a nonnegative instantaneous profit, $\dot{\mu}_2|_{\mu_2\leq 0} < 0$. If $\theta = 0$, the transversality condition is $\mu_2(T) = 0$, which results in $\mu_2(t) \geq 0|t \leq T$.

From (20), we have $v_i^{\Theta}(T) = \frac{\lambda\beta_i}{d_i}(N + \frac{\theta}{\lambda y(T)}) > 0$, which suggests that v_i^{Θ} is increasing over time until it reaches its maximum value at $t = T$.

For similar abilities of the two partners in the NPD activity, the respective innovation efforts are equal at the end of the planning horizon. Furthermore, *all things being equal*, the terminal innovation effort is larger in the cooperative than in the noncooperative setting.

Proposition 4.1. *Qualitative solutions for player i's cooperative improvement and innovation efforts are respectively*

$$\dot{u}_i^{\Theta} = \lambda u_i \sum_{i=1}^{2} \beta_i v_i - \frac{a^2 e\alpha_i}{4bc_i},$$
$$u_i^{\Theta}(T) = 0, \tag{26}$$
$$\dot{v}_i^{\Theta} = \lambda \left\{ v_i \left[\frac{\beta_i v_i}{2} + \beta_{-i} v_{-i} \right] + \frac{\beta_i}{d_i} \left[\frac{a^2 eq}{4b} - \sum_{i=1}^{2} \frac{c_i u_i^2}{2} - \frac{d_{-i} v_{-i}^2}{2} \right] \right\},$$
$$v_i^{\Theta}(T) = \frac{\lambda\beta_i}{d_i}\left(N + \frac{\theta}{\lambda y(T)}\right). \tag{27}$$

Proof. Equations (19) and (20) are differentiated totally with respect to time to yield, after elementary manipulations, (26) and (27). □

In (26), as $\dot{u}_i^{\Theta}|_{u_i=0} < 0$ and $u_i^{\Theta}(T) = 0$, $i = (1, 2)$, the cooperative improvement effort of both players is decreasing over time. The decline of the players' improvement effort results from the negative influence of marginal demand with respect to quality. As marginal demand with respect to quality is larger, the players' improvement effort decreases faster in the cooperative setting than in the noncooperative setting. All things being equal, this implies a larger initial value of improvement effort in the cooperative case. As for the noncooperative scenario, the speed of knowledge accumulation has a negative influence on each player's improvement effort.

In (27), under nonnegative instantaneous profit, each player's cooperative innovation effort is increasing over time: here we can offer intuitive explanations similar to those for the noncooperative game. The speed of each player's innovation effort is positively affected both by the instantaneous profit from the current product and the speed of knowledge accumulation. However the rate of change of each player's innovation effort is negatively influenced by the cost of the other player's innovation effort. A positive impact of the other player's NPD cost on its own innovation effort then results, which is consistent with the rule according to which the more efficient player contributes more to the innovation activity.

To get further insights on the equilibrium time paths of the decision variables, we now conduct a numerical analysis.

5 Numerical Analysis

The numerical analysis is performed to determine the optimal distribution of players' respective efforts in quality improvement and innovation activities for noncooperative and cooperative equilibria. To do so, we fix the duration of the game to 10 time units, $t \in [0, 10]$, and set the base parameter values corresponding to symmetric players (Table 1). Note that the assumption of asymmetric players is also investigated hereafter.

In addition, we assume a low initial quality level for the current product ($q_0 = 1$) and a starting innovation process ($z_0 = 0$).

The following two figures depict the optimal dynamics of control variables.

Figure 2 shows the optimal path of quality improvement effort. The cooperative solution (u_i^{Θ}, $i = 1, 2$) provides higher levels of such effort than the noncooperative one for both players. Under the cooperative setting, the manufacturer and the supplier commit the same amount of effort. The rule that applies here is that the more efficient player contributes more to the quality improvement activity. Conversely, under the noncooperative setting, the supplier makes much more of an effort than the manufacturer does (i.e., $u_2 \gg u_1$). This result holds whatever the relative efficiency of players. Although there is a dramatically higher efficiency of the manufacturer than the supplier in quality improvement (e.g., $\alpha_1/c_1 = 100\alpha_2/c_2$), its contribution, although greater than in the symmetric case (i.e., $\alpha_1/c_1 = \alpha_2/c_2$), remains to a large extent lower than that of the supplier. Overall, while effort allocation for quality improvement in the cooperative equi-

Table 1: Base case parameter values.

Parameter	α_1	α_2	β_1	β_2	c_1	c_2	d_1	d_2
Value	0.1	0.1	0.1	0.1	1	1	1	1
Parameter	N	a	b	p_2	e	λ	ω	θ
Value	1000	10	0.1	50	1	0.1	0.8	10

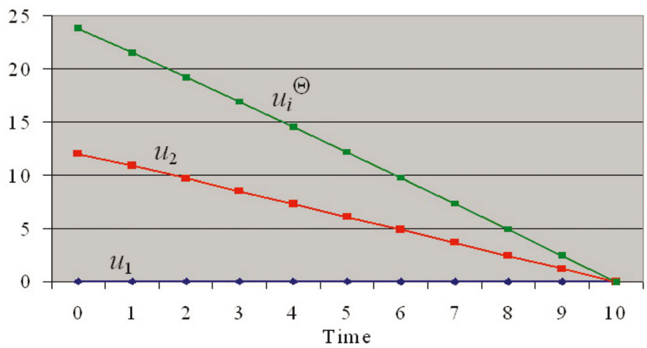


Figure 2: Optimal improvement effort.

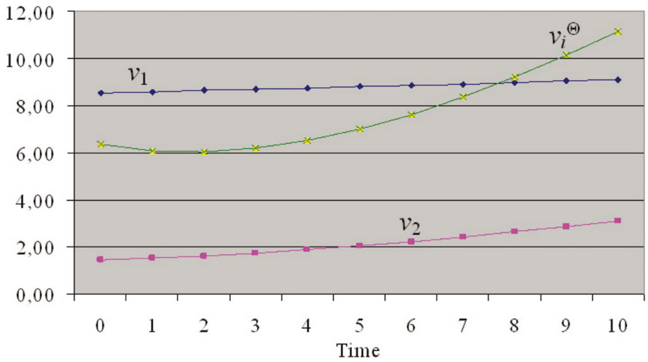


Figure 3: Optimal innovation effort.

librium relies on the pragmatic rule of relative efficiency, it is essentially based on the supplier's commitment in the noncooperative equilibrium.

Figure 3 shows the dynamics of effort for NPD. As expected, the cooperative solution displays the same dynamics for the manufacturer and the supplier, i.e., the relative contribution of each player is proportional to its own relative efficiency in NPD. Furthermore, the individual level of effort for innovation under the cooperative setting (v_i^Θ, $i = 1, 2$) is higher than that of the supplier under the noncooperative setting (v_2) over the entire time horizon, but lower than that of the manufacturer under the noncooperative setting (v_1) over the largest part of the planning horizon. This result, which prevails for a wide range of values for the manufacturer's efficiency in the innovation activity (β_1/d_1), is due to the part of future benefits from the new product that belongs to the manufacturer (ω). For instance, should $\omega = 1$, the noncooperative effort of the manufacturer would be equal to that in the cooperative solution over the entire planning horizon. Note that a larger value of the coefficient for the salvage value of the NPD activity (θ) increases the respective innovation efforts of both players, but more proportionally in the cooperative than in the noncooperative setting.

Comparing Figures 2 and 3, we see that, in the noncooperative setting, the supplier makes a greater effort for quality improvement than the manufacturer, while the manufacturer makes a larger effort for NPD. Overall, the manufacturer's innovation effort remains larger than the supplier's, except for dramatically higher efficiency levels of the supplier (e.g., $\beta_2/d_2 > 6\beta_1 d_1$). This relative stability of the distribution of players' effort shows that the pragmatic rule of relative efficiency is invalid in the noncooperative context. Before making a resource allocation in the noncooperative case, each player compares its implicit benefit in one activity with that in the other activity, rather than comparing its efficiency with that of the other player in the same activity. Here, the supplier gets more from improving the current product quality than the manufacturer does: since we use $p_2 = 50$ and $a/b = 100$ in our numerical analysis, we obtain $p_1 = 75$ according to equation (5).

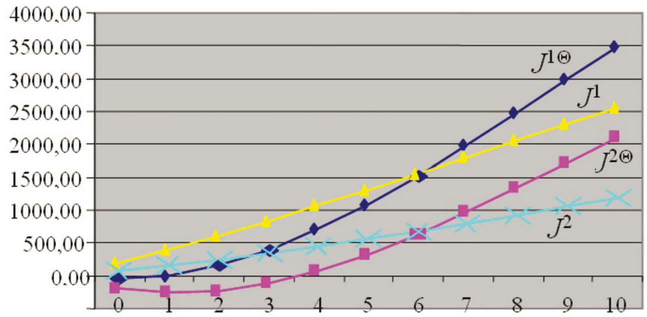


Figure 4: Cooperative and noncooperative cumulative profits.

That is, the supplier gets 50 by selling one unit of the current product, while the manufacturer earns 25. Conversely, the manufacturer gets a higher benefit from the new product than the supplier, i.e., $\omega = 0.8$. Thus, from the supplier's perspective, it is attractive to improve the current product quality for a larger transfer price while, for the manufacturer, it is beneficial to concentrate more on the NPD under a larger rate of revenue from the new product. If, in the cooperative game, one or both players are relatively more efficient in quality improvement than in NPD, their effort in the latter increases only at the very end of the planning horizon. In other words, larger abilities in quality improvement lead cooperative players to postpone their innovation effort. The intuition behind this result is that, since quality improvement offers a better and safer way to make a profit, the players under the cooperative setting consider it far more attractive to capitalize on quality improvement before rushing into NPD than to work on the uncertain activity early.

Figure 4 shows the paths of cumulative profits for the manufacturer and the supplier. For each chain member, the cooperative cumulative profit has been determined using the Nash bargaining scheme. We see that the individual cumulative cooperative profit (resp., $J^{1\Theta}$ and $J^{2\Theta}$) is lower than the noncooperative cumulative profit (resp., J^1 and J^2) at the beginning of the planning horizon, and becomes larger over the last periods. This suggests that the noncooperative setting is more profitable in the short run while the cooperative setting is more beneficial in the longer run.

In Figure 5, the *ex post* optimal transfer price is determined for noncooperative and cooperative settings. It results from the computation of the minimum value of the difference between each player's cumulative cooperative and noncooperative profits (resp., ΔM and ΔP). However, when the supplier (resp., the manufacturer) has larger abilities in the quality improvement activity than its partner, the optimal transfer price and the related optimal sales price should be increasing (resp., decreasing).

Figures 6 and 7 present the dynamics of state variables, i.e., the quality level of current product and the completion probability for NPD. As expected, the quality

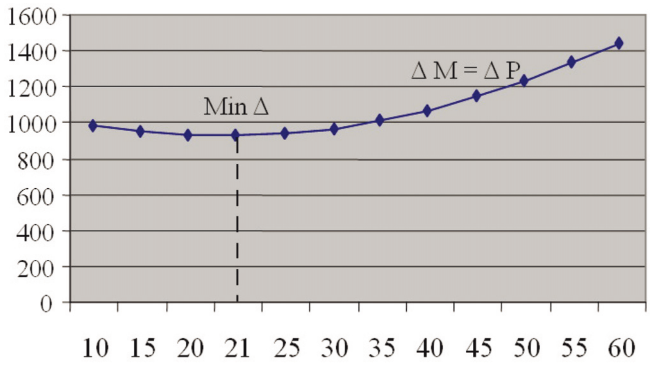


Figure 5: Optimal transfer price.

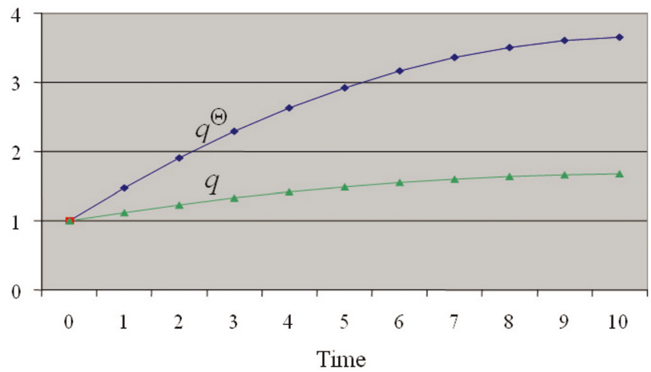


Figure 6: Optimal quality level.

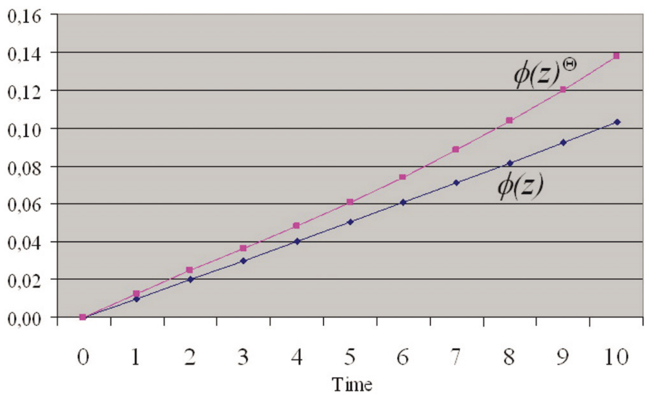


Figure 7: Completion probability for NPD.

Table 2: Game configuration and optimal policy.

		Supplier	
		NPD more effective	Quality more effective
Manufacturer	NPD more effective	Low transfer price and sales price Low quality (both eq) Fast NPD (both eq)	Large transfer price and sales price High quality (higher in coop eq) Fast NPD (slow in coop eq)
	Quality more effective	Low transfer price and sales price Low quality (high in coop eq) Slow NPD (slower in coop eq)	Large transfer price and sales price High quality (higher in coop eq) Slow NPD (slower in coop eq)

level and the completion probability for NPD in the cooperative setting are higher than in the noncooperative setting.

In the case where both players are more efficient in quality improvement than in NPD (i.e., $\alpha_i/c_i > \beta_i/d_i$, $i = 1, 2$), quality improvement of the existing product is greater in the cooperative than in the noncooperative case, while completion probability of the new product is larger in the noncooperative than in the cooperative case. This is due to the fact that greater abilities in quality improvement lead cooperative players to postpone their innovation effort. This observation enlightens a paradox shown in a recent empirical study [12]. According to this study, product innovation is considered as the number 1 factor for revenue growth by a sample of 750 companies located in Europe and North America. However, these companies set quality improvement as the number 1 factor for the supply chains in which they are involved, while product innovation and time-to-market are ranked at their lowest supply chains' priorities. This results in the fact that few of these companies prepare their supply chains for faster introductions. Thus, coordinated chains should perform better in quality improvement while uncoordinated chains should perform better in NPD.

Table 2 summarizes the results obtained in the numerical analysis for both the noncooperative and cooperative equilibria.

6 Conclusion

The issue in this chapter was how the collaborating firms in a supply chain (i.e., a manufacturer and its supplier) allocate resources between improving a marketed product and developing a new product. In this respect, a differential game model for both noncooperative and cooperative settings was designed. Qualitative and numerical analyses were conducted to characterize the optimal path of the decision variables for each player. The following managerial implications emerged.

- Whatever the nature of the equilibrium (noncooperative or cooperative), the optimal strategy implies that both players make initially high and decreasing efforts in quality improvement and initially low and increasing efforts in NPD. The rationale for such behaviors is to achieve maximum profits from the current product at the beginning of the game so as to invest them increasingly in the future product as time goes by.
- In a coordinated chain, the allocation of efforts is determined by the rule according to which, for each activity, the more efficient player contributes more.
- In an uncoordinated chain, the allocation of efforts by each player results from the comparison of its own implicit benefit in one activity with that in the other activity, rather than from the comparison between its own efficiency with that of the other player in the same activity. From the supplier's perspective, the higher the transfer price, the heavier the effort in quality improvement. For the manufacturer, the larger the part of profits from the new product, the heavier the effort in NPD. In our numerical example, we observed a dominant contribution by the supplier in the improvement activity of the existing product and a decisive effort of the manufacturer in the innovation process.
- In most configurations, the uncoordinated chain pays more attention to innovation, while the coordinated chain favors the quality of the existing product.

Although nontrivial observations have been drawn from our model, we believe there is room for extensions. First, the use of a positive discounting rate would provide a broader view of the trade-off among quality improvement and NPD. Furthermore, the assumption of an infinite planning horizon would allow for the study of stationary decisions. Another extension would lie with the analysis of a closed-loop Nash solution. Finally, the exploration of a Stackelberg solution would enlighten the impact of hierarchical plays between the manufacturer and the supplier.

References

[1] Aviv, Y., The effect of collaborative forecasting on supply chain performance, *Management Science*, Vol. 47, No. 10, 2001, pp. 1326–1343.

[2] Bresnahan, T. F. and Reiss, P. C., Dealer and manufacturer margins, *Rand Journal of Economics*, Vol. 16, 1985, pp. 253–268.

[3] Cachon, G. P. and Zipkin, P. H., Competitive and cooperative inventory policies in a two-stage supply chain, *Management Science*, Vol. 45, 1999, pp. 936–953.

[4] Dockner, E., Jørgensen, S., Van Long, N. and Sorger, G., *Differential Games in Economics and Management Science*, Cambridge University Press, Cambridge, 2000.

[5] Durham, Y., An experimental examination of double marginalization and vertical relationships, *Journal of Economic Behavior and Organization*, Vol. 42, No. 2, 2000, pp. 207–229.

[6] Fisher, M., What is the right supply chain for your products?, *Harvard Business Review* (March–April), 1997.

[7] Kamien, M. I. and Schwartz, N. L., Optimal exhaustible resource depletion with endogenous technical change, *Review of Economic Studies*, Vol. 45, 1978, pp. 179–196.

[8] Kamien, M. I. and Schwartz, N. L., *Market Structure and Innovation*, Cambridge University Press, Cambridge, 1982.

[9] Lee, H. L., Padmanabhan, P. and Whang, S., Information distortion in a supply chain: The bullwhip effect, *Management Science*, Vol. 43, No. 4, 1997, pp. 546–558.

[10] Reinganum, J. F., Dynamic games of innovation, *Journal of Economic Theory*, Vol. 25, 1981, pp. 21–41.

[11] Reinganum, J. F., Practical implications of game theoretic models of R&D, *American Economic Review*, Vol. 74, 1984, pp. 61–66.

[12] Smets, R., *Mastering Complexity and Creating Value in Supply Chain Management—A Supply Chain Transformation Roadmap*, Deloitte Consulting Report, 2004.

A Differential Game of a Dual Distribution Channel

Olivier Rubel
HEC Montréal and GERAD
3000 Chemin de la Côte-Sainte-Catherine
Montréal, PQ, H3T 2A7, Canada
olivier.rubel@hec.ca

Georges Zaccour
HEC Montréal and GERAD
3000 Chemin de la Côte-Sainte-Catherine
Montréal, PQ, H3T 2A7, Canada
georges.zaccour@gerad.ca

Abstract

An infinite-horizon differential game between a manufacturer and a retailer is considered. The players control their marketing efforts, and the sales share of the online channel is the state of the system. The manufacturer seeks to maximize her profit made on both the indirect and direct channels and faces, aside from her marketing effort, a logistics cost of selling online. The retailer seeks to keep consumers buying offline through her effort. A feedback Nash equilibrium is identified and results are discussed.

Key words. Differential game, dual distribution channel, electronic commerce.

1 Introduction

During the last decade or so, manufacturers have increasingly been selling directly to consumers through the Internet. At the same time, retailers are competing directly with manufacturers of national brands by offering their private labels to the same pool of shoppers. Therefore, the manufacturers are also becoming retailers and *vice versa*.

The impact of this shift in traditional roles on the relationships between channel's partners and their performance has attracted the attention of marketing scholars. Regarding the introduction of private labels, Mills ([15], [16]), Raju et al. ([20]), and Narasimhan and Wilcox ([19]) have assessed the impact on manufacturers' and retailers' prices and profits. Generally speaking, it seems that retailers would benefit from offering their brands, along with the national ones, to consumers and eventually manufacturers would witness a decline in their profits. Karray and Zaccour ([13]) identified the circumstances under which a cooperative advertising program allows a manufacturer to mitigate, at least partially, the profit losses due to the introduction by the retailer of the store brand. Morton and Zettelmeyer ([18]) investigate whether it is beneficial for the retailer to replace an existing national brand with a private label, and which manufacturer's product should be eliminated from the existing retail assortment.

Turning now to the involvement of manufacturers in retailing, the natural questions are why manufacturers are going direct, and eventually how to manage a dual distribution channel. Although the general answer to the first question is to create a sustainable competitive advantage by increasing the market penetration of their products (Alba et al. ([1])), it seems that the strategy of selling goods and services online could be both performance enhancing and performance destroying (Geyskens et al. ([5])). The negative part is due to the cost of duplicating a distribution channel and the cost of conflict with the other intermediaries involved in the channel, retailers in particular (Sarkar et al. ([21])). As Chiang et al. ([3]) point out, "*while more and more manufacturers are engaging in direct sales, their retailers partners voice the belief that orders placed through a manufacturer's direct channel are orders that should have been placed through them*."

Although a lot has been said about the evolution of the role of manufacturers, the marketing science literature, and more specifically the game theoretic one, is rather sparse. Tsay and Agrawal ([23]) survey contributions in multi-channel distribution systems, and touch the implications of manufacturers going direct on relationships in channels. The survey shows that the main decision variables used in the literature are the transfer and retail prices and/or inventories and that game models have been static. Balasubramanian ([2]) considers a spatial model where one retailer distributes her product online and analyzes its effect on the conventional retail industry. Zettelmeyer ([24]) studies the case of two integrated firms competing in both conventional and direct channels. Decision variables include prices and search costs for consumers. As a result, as the Internet gains ground among consumers, competition tends to shift from search cost towards price competition. Chiang et al. ([3]) consider the situation where a manufacturer is selling both directly (through the Internet) and indirectly (through an independent conventional retailer) to customers. The main result, while somewhat surprising, is that no sales are made online. The manufacturer has actually no interest in selling directly and her only purpose in establishing the new channel is to pressure the retailer to reduce her price and thus boost the demand for the manufacturer's brand. Fruchter and Tapiero ([4]) extend the setting in Chiang et al. ([3]) to a dynamic framework. They account for consumers' heterogeneity on the acceptance of the virtual channel, with prices as strategies in a Stackelberg game.

The purpose of this chapter is to characterize equilibrium marketing strategies in a channel comprised of one manufacturer and one retailer where the former also sells directly to consumers. The model is fully dynamic with the online channel's share as the state variable. We assume that its evolution is governed by marketing expenditures of players and by consumers' channel switching behavior. Our contribution here lies in bringing the differential game methodology to this area and in the explicit analysis of the impact of marketing efforts and store switching on each distribution channel's share.

The remainder of the chapter is organized as follows. In Section 2 we introduce the model and in Section 3 we characterize equilibrium strategies. In Section 4 we conduct some sensitivity analyses on strategies and steady state with respect to the model's parameters. In Section 5 we offer our conclusions.

2 Model

Consider a manufacturer (player M) selling her brand through both a representative retailer (player R) and directly to consumers through, e.g., the Internet. Denote by $x(t)$, $0 \leq x(t) \leq 1$, the share of online total sales of the manufacturer's brand at time $t \in [0, \infty)$. The evolution over time of this state variable will be specified below.

Denote by $E_i(t)$ the marketing efforts of player $i, i \in \{M, R\}$ at time $t \in [0, \infty)$. Player i's marketing effort is aimed at keeping/attracting consumers to "her" preferred channel, i.e., the conventional (offline) channel for the retailer and the direct (online) channel for the manufacturer. Marketing efforts can be interpreted as means to provide relevant information (by advertising and other types of communication actions, i.e., displays of products (virtual or in store), promotions, etc.) to consumers so as to make their search processes less costly (see, e.g., Lal and Sarvary ([14]), Zettelmeyer ([24])). They also intend to influence consumers' choice of purchasing option, as pointed out by Hauser et al. ([6]), "*consumers should seek out formats that enable them to make selections that maximize consumption utility net of price and search costs (…) even if retail formats offer identical merchandise*."

Although the manufacturer is selling in both channels, the assumption here is that she prefers the direct channel. This preference stems from the fact that the manufacturer's margin, measured per point of share of each channel, on sales made directly to consumers (denoted m_D) is higher than the margin obtained by selling through the indirect channel (denoted m_I). Note that this ordering of manufacturer's margins, i.e., $m_I < m_D$, is a necessary condition to avoid having the retailer buying from the online market. Denote by m_R the retailer's margin per point of the offline's channel market share. Note that we do not impose any condition on the magnitude of m_R with respect to m_I and m_D.

We assume that the evolution of the online market share depends on both players' marketing expenditures and on the market share of the online channel, i.e.,

$$\frac{dx(t)}{dt} = \dot{x}(t) = H(E_M(t), E_R(t), x(t)).$$

To keep things simple in this rather exploratory study, we assume that $H(.)$ is separable in the controls and state and is given by[1]

$$\frac{dx}{dt} = \dot{x}(t) = (\mu_M E_M(t) - \mu_R E_R(t)) - (px(t) - q(1 - x(t))),$$
$$x(0) = x_0, \tag{1}$$

where μ_M and μ_R are strictly positive parameters, and p and q are parameters which satisfy $0 \leq p, q \leq 1$.

The first bracketed right-hand-side term of (1) states that the online market share evolution depends on the combined effect of players' marketing efforts; it increases (decreases) with manufacturer's (retailer's) marketing efforts. The parameters μ_M and μ_R can be interpreted as efficiency parameters in transforming marketing efforts into corresponding channel's market share. The second term captures the idea that consumers switch between the two channels, with parameters p and q measuring the intensity with which each channel's customers switch to the alternative one. Thus, the variation in the market share of the online channel depends on a differential in marketing efforts and a differential in the movement of consumers between the two channels. Note that these switches can be explained by a variety of reasons such as convenience, last buying experience, desire of change, etc.

Remark 2.1. Note that if $p = q = 0$, then the market share evolution would be of the "excess-advertising" type, i.e., depending only on the difference in advertising efforts of both players (see Jørgensen and Zaccour ([12]) for a recent review of differential games in advertising competition).

Remark 2.2. If the dynamics of the online channel share could not be at all controlled by the players, i.e.,

$$\dot{x}(t) = -px(t) + q(1 - x(t)), x(0) = x_0,$$

then the equilibrium steady state would be given by

$$0 < x_{ss} = \frac{q}{p+q} < 1.$$

Thus the higher the intensity of switching from the offline channel to the online one, the higher the steady-state share of the latter.

Further, if in this case p and q were functions of the marketing efforts, then the resulting dynamics

$$\dot{x}(t) = -p(E_R)x(t) + q(E_M)(1 - x(t)), x(0) = x_0,$$

would be one à la Lanchester; see Jørgensen and Zaccour ([12]).

[1]Note that nothing guarantees that the state variable will always be in the interval [0, 1]. We shall assume, as is usually done in the advertising dynamics literature, that it will be the case. We shall verify that this assumption is indeed verified in all the numerical simulations.

We assume that the total cost of marketing effort is convex increasing. For simplicity, we adopt, as in, e.g., Jørgensen et al. ([9]–[11]), a quadratic specification, i.e.,

$$C_i(E_i) = \frac{1}{2}a_i E_i^2, \quad a_i > 0, \quad i \in \{M, R\}.$$

Note that the above marketing effort cost functions are asymmetric, which reflects the idea that the two players are not necessarily using the same media to advertise or communicate with their target markets.

The margins m_D and m_I defined above for the manufacturer are net of production cost. Furthermore, the margin m_I is also net of delivery cost to the retailer. In addition, the retailer's margin is net of retailing costs (e.g., inventories, replenishment of shelf space, and merchandising). There is still one cost which has to be accounted for: the cost of selling online by the manufacturer. This cost may include treatment of orders, handling and shipping to consumers, etc. We will assume that these cost items can be captured by a function which depends on the share of the online channel and denoted $F(x)$. We assume that this function is positive and convex increasing and satisfies $F(0) = 0$. One way of justifying convexity is by stating that when the share of the online channel becomes larger, the manufacturer incurs an increasingly high (handling, shipping, etc.) cost because, e.g., of the extra working hours of the staff and the vehicles. Again for mathematical tractability, we will assume a quadratic specification, i.e., $F(x) = 1/2bx^2$, where $b > 0$.

Omitting from now on the time argument when no confusion may arise and assuming that both players discount their stream of profits at the same market rate r, their objective functionals are as follows:

$$\underset{E_M \geq 0}{\text{Max}}\, \Pi_M = \int_0^\infty e^{-rt}\{m_D x + m_I(1-x) - \frac{1}{2}a_M E_M^2 - \frac{1}{2}bx^2\}\, dt, \tag{2}$$

$$\underset{E_R \geq 0}{\text{Max}}\, \Pi_R = \int_0^\infty e^{-rt}\{m_R(1-x) - \frac{1}{2}a_R E_R^2\}\, dt. \tag{3}$$

To recapitulate, by (2) and (3)–(4) we have defined a two-player infinite–horizon differential game with one state variable $x, 0 \leq x(t) \leq 1$, and two controls E_M, E_R constrained to be nonnegative. Note that the structure of our game is of the often used linear-quadratic variety.

To conclude, we wish to highlight the following features of the model:

- Since the manufacturer is selling in both channels, our model treats the issue of how to deal with a dual distribution channel. Indeed, by deciding on the optimal (in an equilibrium sense) marketing effort to attract customers to the online channel, the manufacturer is implicitly proceeding to an arbitrage between the share of sales in the two channels.
- We assume that given the intensities (p and q) at which consumers switch from one channel to another, their values can be different. Therefore, our framework accounts for the idea that these switches may be due to different reasons or underlying processes.

- The two players have clearly conflicting views on the online channel. For the manufacturer it is a complement to the conventional one, for the retailer it is a clear substitute. This raises the question, not answered here, whether coordination of the two players' marketing efforts can still be feasible.

3 Equilibrium

We assume that the players use feedback marketing effort strategies. Since the game is played over an infinite horizon, we shall confine our interest to stationary strategies. The following proposition characterizes the unique feedback Nash equilibrium.

Proposition 3.1. *Feedback Nash marketing equilibrium strategies are given by*

$$E_M(x) = \begin{cases} (Ax+B)\frac{\mu_M}{a_M}, & \textit{for} \quad x \leq -\frac{B}{A} \\ 0, & \textit{otherwise} \end{cases}, \tag{4}$$

$$E_R(x) = -D\frac{\mu_R}{a_R} = cst. \tag{5}$$

Value functions are as follows:

$$V_M(x) = \frac{1}{2}Ax^2 + Bx + C, \tag{6}$$

$$V_R(x) = Dx + E, \tag{7}$$

where

$$A = \frac{a_M}{\mu_M^2}\left(\left(p+q+\frac{r}{2}\right) - \sqrt{\left(p+q+\frac{r}{2}\right)^2 + \frac{b\mu_M^2}{a_M}}\right) \tag{8}$$

$$B = \frac{m_D - m_I + A\left(q + D\frac{\mu_R^2}{a_R}\right)}{\frac{r}{2} + \sqrt{\left(p+q+\frac{r}{2}\right)^2 + \frac{b\mu_M^2}{a_M}}}, \tag{9}$$

$$C = \frac{1}{r}\left(m_I + \frac{(B\mu_M)^2}{2a_M} + \frac{BD\mu_R^2}{a_R} + qB\right), \tag{10}$$

$$D = \frac{-m_R}{\frac{r}{2} + \sqrt{\left(p+q+\frac{r}{2}\right)^2 + \frac{b\mu_M^2}{a_M}}}, \tag{11}$$

$$E = \frac{1}{r}\left(m_R + \frac{(D\mu_R)^2}{2a_R} + D\left(q + B\frac{\mu_M^2}{a_M}\right)\right). \tag{12}$$

Proof. See the Appendix. □

Proposition 3.1 shows that the retailer's value function is linear and thus this player advertises at a constant rate. The manufacturer's marketing effort strategy is state dependent. Note that since A is negative, for the manufacturer's strategy to make sense, B must be positive. Indeed, if B were negative, then the advertising strategy would read $E_M(x) = (Ax + B)\frac{\mu_M}{a_M} \geq 0$, for $x \leq -\frac{B}{A} < 0$, which contradicts the constraint $0 \leq x \leq 1$. We shall thus assume from now on that the parameters are such that B is positive. Since the denominator is always positive, then the necessary condition for having $B > 0$ is that the numerator be positive, i.e., $\left(m_D - m_I + A\left(q + D\frac{\mu_R^2}{a_R}\right)\right) > 0$. Moreover, if the online channel's share x is greater than the threshold level $-\frac{B}{A}$, assuming this threshold is lower than 1, the manufacturer will not invest in marketing efforts. Under such circumstances, the revenues from gaining an additional share will not counterbalance the cost.

Substituting for equilibrium marketing efforts from (4)–(5) in the state dynamics (1), we obtain, after some straightforward manipulations, the following value for the steady state for the online channel:

$$x_{ss} = \left(\frac{m_D - m_I}{Y(Y-r)}\right)\frac{\mu_M^2}{a_M} + \frac{(p+q+r)}{a_R Y^2 (Y-r)}(a_R Y q - m_R \mu_R^2), \tag{13}$$

where

$$Y = \frac{r}{2} + \sqrt{\left(p + q + \frac{r}{2}\right)^2 + \frac{b\mu_M^2}{a_M}}.$$

As is readily seen, the steady state depends on the model's parameters. Conditions on the parameters' values can be derived to ensure that the steady state is bounded between 0 and 1.

Remark 3.1. In some papers dealing with conflicts and cooperation in marketing channels, the authors assume that the game is played à la Stackelberg, with the manufacturer often assuming the role of leader (e.g., Shugan ([22]), Moorthy and Fader ([17]), Jeuland and Shugan ([8]), and Jørgensen et al. ([11])). In our setting, irrespective of who the leader is, a Stackelberg equilibrium coincides with the Nash one derived above. The reason is that the follower's reaction function does not depend on the leader's marketing effort, i.e., the latter cannot influence the former's choice. If the players' decisions were more linked in the dynamics or through their margins (assuming they are endogenous), then one would expect the Nash and Stackelberg equilibria to be different.

The following result applies when the logistics cost of online deliveries is zero.

Corollary 3.1. *If the logistics cost for the online channel is zero, i.e., $b = 0$, then value functions are linear,*

$$V_M(x) = B_1 x + C_1, \quad V_R(x) = D_1 x + E_1,$$

and feedback Nash equilibrium strategies are constant and given by

$$E_M = \frac{\mu_M}{a_M} B_1, \quad E_R = -\frac{\mu_R}{a_R} D_1,$$

where

$$B_1 = \frac{m_D - m_I}{r + p + q} > 0, \quad C_1 = \frac{1}{r}\left[m_I + \frac{B_1 \mu_M^2}{2a_M} + \frac{B_1 D_1 \mu_R^2}{a_R} + q B_1\right],$$
$$D_1 = \frac{-m_R}{r + p + q} < 0, \quad E_1 = \frac{1}{r}\left[m_R + \frac{(D_1 \mu_R)^2}{2a_R} + D_1\left(q + B_1 \frac{\mu_M^2}{a_M}\right)\right].$$

The steady state is given by

$$x_{ss} = \frac{q}{(p+q)} + \frac{1}{(p+q+r)(p+q)}\left(\frac{(m_D - m_I)\,\mu_M^2}{a_M} - \frac{m_R \mu_R^2}{a_R}\right). \tag{14}$$

Proof. The results follow immediately from setting $b = 0$ in the previous proposition. □

The above scenario corresponds to a service or a product which can be delivered, for instance, via the Internet at (almost) zero cost to the service provider or the manufacturer. Software which can be downloaded directly by the customer on her computer and electronic airline tickets are examples of such a context. Note that, in this case, the players will constantly make marketing efforts at positive rates.

To interpret the steady-state value of the online channel, note that (14) can be rewritten as

$$x_{ss} = \frac{q}{(p+q)} + \frac{1}{(p+q)}(\mu_M E_M - \mu_R E_R).$$

Thus, the steady state of the online channel is equal to the steady state when the dynamics are not controlled, plus an “excess-advertising” term. The sign of the latter depends on the relative advertising levels of the two players which in turn depend on all model parameters (margins, intensities of switching, and cost parameters).

Comparing the retailer’s value functions under the two scenarios ($b \neq 0$ and $b = 0$) shows that the slope of the first one is bigger in absolute value than the value function with $b = 0$. Indeed,

$$D_1 = \frac{-m_R}{r + p + q} < D = \frac{-m_R}{\frac{r}{2} + \sqrt{\left(p + q + \frac{r}{2}\right)^2 + \frac{b\mu_M^2}{a_M}}}.$$

Thus, the retailer’s payoff decreases less sharply in x when the manufacturer faces a logistics cost. Although our model does not capture all competitive dimensions of e-commerce, this leads to the conjecture that the competitive standing of offline retailing will heavily depend on the delivery cost of online purchases.

4 Sensitivity Analysis

To get more insight into the players' strategies and the steady-state online market share, we study how they vary with the parameters' values. The following proposition provides the results for margins.

Proposition 4.1. *The manufacturer's marketing strategy is increasing in the online channel and retailer margins and decreasing in the offline channel margin.*

The retailer's marketing strategy is increasing in her margin and is independent of the manufacturer's online and offline margins.

Proof. Straightforward derivations lead to

$$\frac{\partial E_M(x)}{\partial m_D} = -\frac{\partial E_M(x)}{\partial m_I} = \frac{\mu_M}{a_M Y} > 0, \qquad \frac{\partial E_M(x)}{\partial m_R} = -\frac{A\mu_M \mu_R^2}{a_M a_R Y^2} > 0,$$

$$\frac{\partial E_R}{\partial m_R} = \frac{\mu_R}{a_R Y} > 0, \qquad \frac{\partial E_R}{\partial m_I} = \frac{\partial E_R}{\partial m_D} = 0.$$

□

Increasing one's own margin (m_D for M or m_R for R) leads the concerned player to increase her marketing effort for the understandable reason that it becomes more profitable to attract customers to this channel. The independence of the retailer's strategy with respect to the manufacturer's margins is a by-product of the game structure, i.e., the retailer will buy from the manufacturer independently of the latter's margins and she does not intervene in the online market. Now, increasing the retailer's margin will eventually lead the manufacturer to increase her effort, as a reaction to the increase of the retailer's effort to attract customers to her offline channel. Note that although the retailer's margin does not appear explicitly in the manufacturer's problem, it nevertheless affects her profit via its impact on the retailer's strategy and thus on the share of the offline channel.

We now turn to sensitivity analysis of strategies with respect to the switching parameters.

Proposition 4.2. *The retailer's marketing effort strategy is decreasing in both channels' switching intensities parameters.*

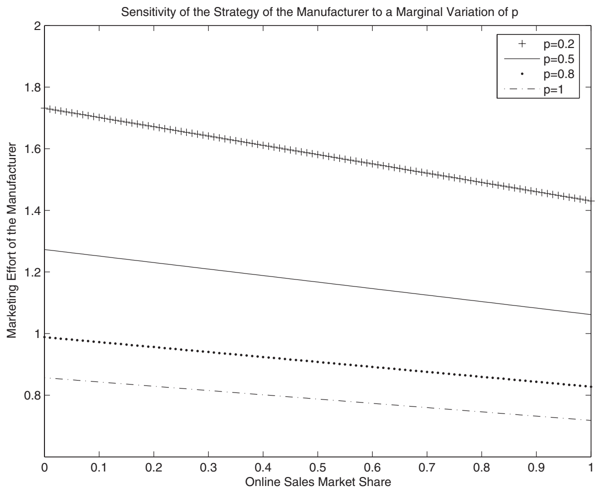
Proof. Straightforward computations lead to the result

$$\frac{\partial E_R}{\partial p} = \frac{\partial E_R}{\partial q} = -\frac{\mu_R}{a_R}\frac{m_R}{Y^2}\frac{(p+q+\frac{r}{2})}{(Y-r)} \leq 0,$$

where

$$Y = \frac{r}{2} + \sqrt{\left(p+q+\frac{r}{2}\right)^2 + \frac{b\mu_M^2}{a_M}}.$$

□



	m_D	m_I	m_R	a_M	a_R	b	r	μ_M	μ_R	p	q
q	4.5	1.6	1.3	1.4	1.1	0.1	0.1	0.6	0.9		0.3

Figure 1: Sensitivity of the manufacturer's strategy to p.

Proposition 4.2 states that a retailer's response to a change in either p or q leads to the same impact on marketing efforts, i.e., $\frac{\partial E_R}{\partial p} = \frac{\partial E_R}{\partial q} \leq 0$. The first result $\frac{\partial E_R}{\partial p} \leq 0$ is rather intuitive; indeed the higher the (uncontrollable) propensity of the switch from the online channel to the offline one, the less the retailer needs to spend on marketing efforts to attract customers to her outlet. Thus p and E_R are substitutes. The second result is surprising; it says that the higher the intensity of customers leaving the offline channel to the online one, the less the retailer invests in marketing efforts.

To shed light on this result, we examine the impact of these parameters on the manufacturer's strategy. We could not determine analytically the signs of $\frac{\partial E_M}{\partial q} \leq 0$ and $\frac{\partial E_M}{\partial p} \leq 0$ (the expressions are very long, and do not admit any apparent interpretation). We shall thus assess numerically the impact of these parameters on the manufacturer's strategy. We computed these derivatives for different values of p and q.[2]

Figure 1 shows that increasing p (or q) also has a decreasing effect on the manufacturer's strategy. The result for q is intuitive; indeed, q and E_M are substitutes for the manufacturer. The result for p, namely saying that the higher the intensity of consumers leaving her "own" channel, the lower the marketing effort, is also a surprise. One possible explanation of $\frac{\partial E_R}{\partial q} \leq 0$ and $\frac{\partial E_M}{\partial p} \leq 0$ may lie in the dynamics of the model. Indeed, the equation describing the evolution of the

[2]Although we conducted numerical experiments for all values of p and q lying between 0.1 and 0.9, with a step size of 0.1, we have only printed a few curves for clarity.

online channel market share states implicitly that the higher the switch from one channel to another at any given instant in time, the higher the potential for a comeback at a later instant in time. In other words, given that the flows of switchers are uncontrollable, the best a player can do is adopt a *surf-the-wave strategy*.[3]

The following proposition gives the results of the sensitivity analysis of the retailer's strategy with respect to the cost parameters.

Proposition 4.3. *Increasing any cost parameter leads the retailer to reduce her marketing effort.*

Proof. Straightforward computations give

$$\frac{\partial E_R}{\partial a_R} = \frac{D\mu_R}{a_R^2} \leq 0, \frac{\partial E_R}{\partial a_M} = -\frac{m_R\mu_R b\mu_M^2}{2a_R a_M^2 Y^2(Y - \frac{r}{2})} \leq 0,$$
$$\frac{\partial E_R}{\partial b} = -\frac{\mu_M^2}{a_M}\frac{\mu_R}{a_R}\frac{m_R}{(Y-r)Y^2} \leq 0.$$

□

The impact of a_R is intuitive; indeed, the marketing effort cost is convex increasing and thus an upward shift in the value of the parameter leads, quite naturally, to a decrease in this activity. To attempt to interpret the other results, it is insightful to consider the impact of varying these parameters on the manufacturer's strategy. We claim the following.

Claim 4.1. Increasing any cost parameter, i.e., a_M, a_R, or b, leads to a decrease in the manufacturer's marketing effort.

This claim is based on the following observations. First, note that

$$\frac{\partial E_M(x)}{\partial a_R} = \frac{-AD\mu_R^2}{a_R^2 Y}\frac{\mu_M}{a_M} \leq 0.$$

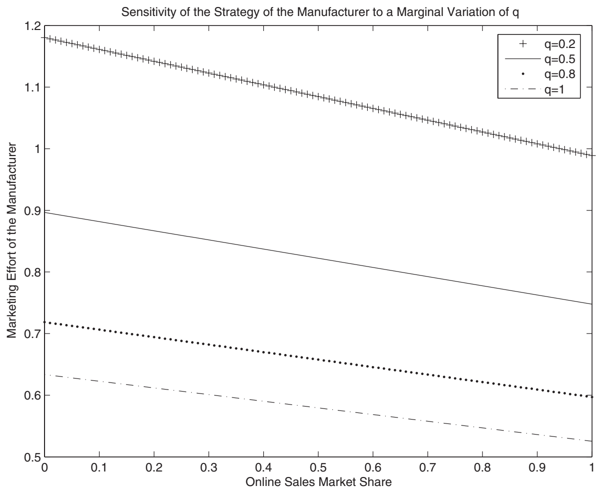
Next, although we could not assess analytically the signs of $\frac{\partial E_M(x)}{\partial a_M}$ and $\frac{\partial E_M(x)}{\partial b}$, numerical and other evidence seem to show that these derivatives are negative. Indeed, in the simple case where there is no logistics cost ($b = 0$), the manufacturer's equilibrium strategy is given by

$$E_M = \frac{\mu_M}{a_M}\frac{m_D - m_I}{r + p + q},$$

and hence

$$\frac{\partial E_M}{\partial a_M} = -\frac{\mu_M}{a_M^2}\frac{m_D - m_I}{r + p + q} \leq 0.$$

[3]This type of phenomenon is also present in the Lanchester model, one of the most celebrated differential game models of advertising competition (see, e.g., Jarrar et al. ([7]) or Jørgensen and Zaccour ([12])).



	m_D	m_I	m_R	a_M	a_R	b	r	μ_M	μ_R	p	q
q	4.5	1.6	1.3	1.4	1.1	0.1	0.1	0.6	0.9	0.7	

Figure 2: Sensitivity of the manufacturer's strategy to q.

Furthermore, the numerical results (see Figure 2) also show that the manufacturer's strategy is, in the general case ($b > 0$), a decreasing function in her cost parameters a_M and b. Hence, we believe that Proposition 4.3 and the other indications allow us to state that increasing the cost of either player leads to a decrease in both players' marketing effort. A simple explanation is as follows. It is straightforward to admit that increasing any cost parameter implies a decrease in the activity by the concerned player. This means a reduction in the competitive pressure on the other player to defend "her" channel, so that it is optimal to decrease her marketing expenditures. This result can also be interpreted in terms of strategic dependency. Marketing efforts are strategic complements (substitutes) when an increase in the expenditures by one player leads the other player to increase (decrease) her expenditures. Here, we have strategic complementarity, which is due to the "excess-advertising" term in the dynamics.

Proposition 4.4. *Increasing either marketing effort efficiency parameter leads to an increase in the retailer's marketing effort.*

Proof. It suffices to note that

$$\frac{\partial E_R}{\partial \mu_R} = -\frac{D}{a_R} \geq 0, \qquad \frac{\partial E_R}{\partial \mu_M} = \frac{m_R \mu_R b \mu_M}{a_R a_M Y^2 \left(Y - \frac{r}{2}\right)} \geq 0.$$

□

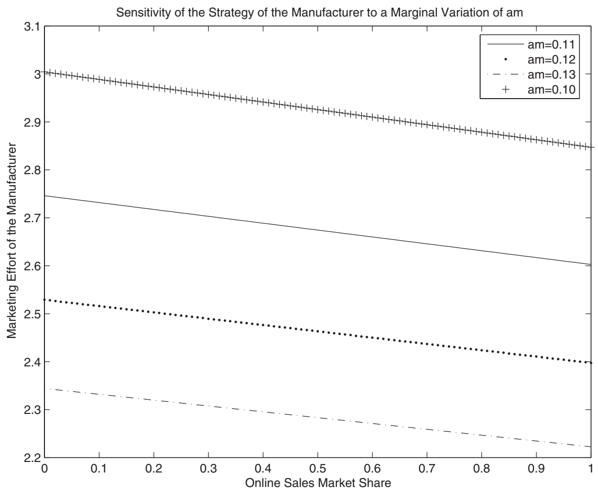
The proposition is stated for the retailer, but the result can also be claimed for the manufacturer. Indeed, differentiating the manufacturer's marketing effort equilibrium strategy with respect to the retailer's efficiency parameter gives

$$\frac{\partial E_M(x)}{\partial \mu_R} = \frac{2AD\mu_R}{a_R Y}\frac{\mu_M}{a_M} \geq 0.$$

We cannot characterize analytically the sign of $\frac{\partial E_M(x)}{\partial \mu_M}$. Again, looking at the special case where the logistics cost is zero, we have

$$\frac{\partial E_M}{\partial \mu_M} = \mu_M \left(\frac{m_D - m_I}{r + p + q} \right) \geq 0.$$

Further, the numerical results reported in Figure 3 for the general case (i.e., $b > 0$) show that the manufacturer's marketing effort decreases with μ_M. These results and the ones in the preceding propositions provide sufficient evidence to claim that increasing the value of any efficiency parameter positively affects both players' marketing efforts. Actually, this claim mirrors the claim stated for the cost parameters. Indeed, if one player finds it advantageous to increase her marketing effort because she is more efficient, then the other player follows in order to defend her channel's market share. Here again we see that marketing efforts are strategic complements.



	m_D	m_I	m_R	a_M	a_R	b	r	μ_M	μ_R	p	q
a_M	5	4	1		0.4	0.1	0.1	0.5	1.3	0.8	0.7

Figure 3: Sensitivity of the manufacturer's strategy to a_M.

The following proposition provides the results of the sensitivity analysis of the steady state with respect to the different parameters.

Proposition 4.5. *The steady state increases in the online margin and decreases in offline ones.*

Proof. Recall that the steady state is given by

$$x_{ss} = \left(\frac{m_D - m_I}{Y(Y-r)}\right)\frac{\mu_M^2}{a_M} + \frac{(p+q+r)}{a_R Y^2 (Y-r)}(a_R Y q - m_R \mu_R^2),$$
$$\text{with } Y = \frac{r}{2} + \sqrt{\left(p+q+\frac{r}{2}\right)^2 + \frac{b\mu_M^2}{a_M}}.$$

It suffices to compute the following derivatives to get the result:

$$\frac{\partial x_{ss}}{\partial m_D} = -\frac{\partial x_{ss}}{\partial m_I} = \frac{\mu_M^2}{Y(Y-r)a_M} \geq 0, \quad \frac{\partial x_{ss}}{\partial m_R} = \frac{-\mu_R^2(p+q+r)}{Y^2(Y-r)a_R} \leq 0.$$

□

The steady state of the online market share increases in the margin that the manufacturer gets in this channel and decreases in the margin of the alternative channel. This result is expected. Note that the sum of variations of the steady state with respect to the manufacturer's margins is zero, i.e., $\left(\frac{\partial x_{ss}}{\partial m_D} + \frac{\partial x_{ss}}{\partial m_I} = 0\right)$. This "trade-off" rule is due to the linearity in x of the revenues' terms in the manufacturer's optimization problem.

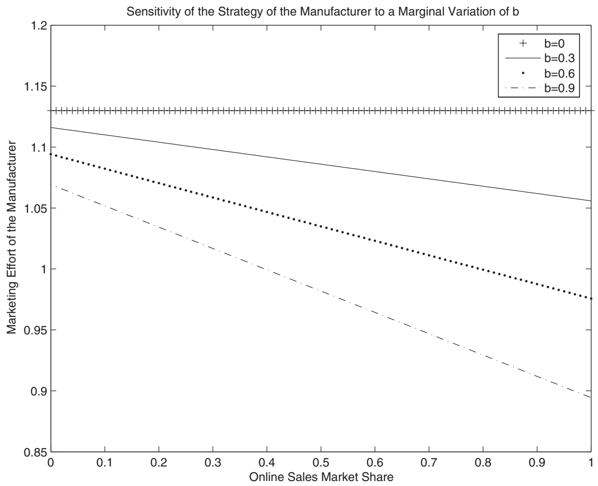
The impact of p and q on the steady state cannot be assessed analytically. It is however easy to establish that $\frac{\partial x_{ss}}{\partial p} < \frac{\partial x_{ss}}{\partial q}$. Indeed, straightforward computations lead to

$$\frac{\partial x_{ss}}{\partial p} - \frac{\partial x_{ss}}{\partial q} = -\frac{(p+q+r)}{Y(Y-r)} < 0 \Leftrightarrow \frac{\partial x_{ss}}{\partial p} < \frac{\partial x_{ss}}{\partial q}. \quad (15)$$

Since we expect p and q to play opposite roles, the inequality in (15) provides a basis to conjecture that increasing the propensity of consumers leaving the online channel hurts its steady-state value, i.e., $\frac{\partial x_{ss}}{\partial p} < 0$, and that it is the other way around for q. We conducted a series of numerical simulations for $p, q \in [0.1, 0.9]$, and in all cases the results confirm this rather intuitive conjecture (see Figure 4).

Turning to the impact of cost parameters, the results are summarized in the following.

Proposition 4.6. *The online channel steady-state market share increases in the retailer's marketing effort cost parameter and decreases in the efficiency of marketing effort.*



	m_D	m_I	m_R	a_M	a_R	b	r	μ_M	μ_R	p	q
b	4.5	1.6	1.3	1.4	1.1		0.1	0.6	0.9	0.7	0.3

Figure 4: Sensitivity of the manufacturer's strategy to *b*.

Proof. The derivatives are given by

$$\frac{\partial x_{ss}}{\partial a_R} = \frac{(p+q+r)}{Y^2(Y-r)} \left(\frac{m_R \mu_R^2}{a_R^2} \right) \geq 0,$$
$$\frac{\partial x_{ss}}{\partial \mu_R} = \frac{-2 m_R \mu_R (p+q+r)}{a_R Y^2 (Y-r)} \leq 0.$$

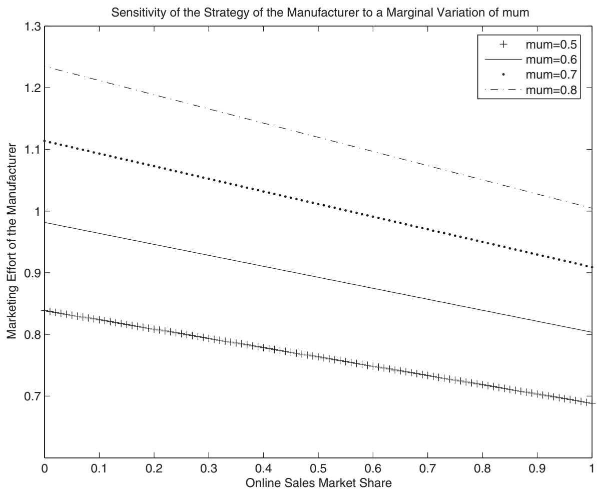
□

Regarding the manufacturer, numerical simulations show that the online channel steady state market share decreases in her marketing effort and logistics cost parameters and increases in the efficiency parameter μ_M (see Figures 5 and 6). These results are not surprising. If the cost parameter of the player is increased, then the level of the marketing effort is reduced and so is the share of this player channel. The efficiency parameter plays an opposite role.

Table 1 recapitulates the results of the sensitivity analysis.

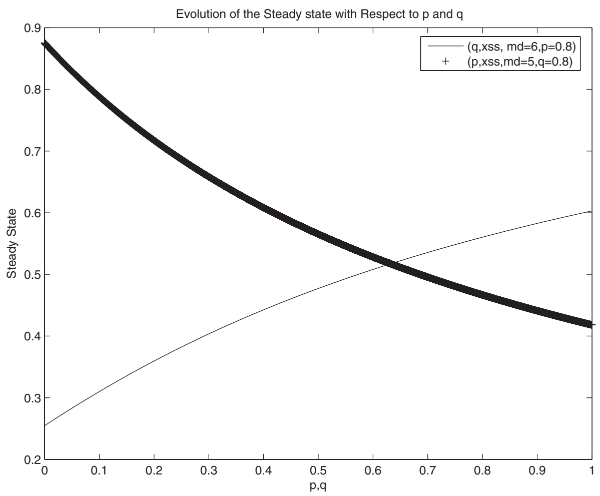
5 Conclusion

This rather exploratory work on marketing efforts in a dual distribution channel game could be extended in numerous ways. We have considered margins as given. Since pricing issues are of crucial importance in e-commerce, introducing them



	m_D	m_I	m_R	a_M	a_R	b	r	μ_M	μ_R	p	q
μ_M	4.5	1.6	1.3	1.4	1.1	0.1	0.1		0.9	0.7	0.7

Figure 5: Sensitivity of the manufacturer's strategy to μ_M.

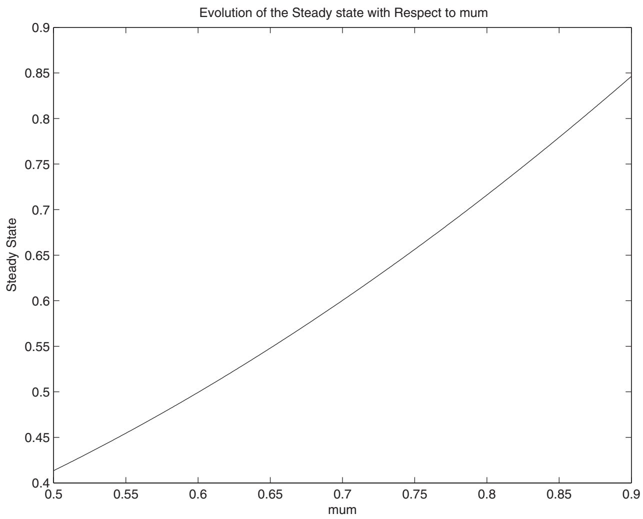


	m_D	m_I	m_R	a_M	a_R	b	r	μ_M	μ_R	p	q
p	5	1.6	1.3	1.4	1.1	0.1	0.1	0.6	0.9		0.8
q	6	1.6	1.3	1.4	1.1	0.1	0.1	0.6	0.9	0.8	

Figure 6: Sensitivity of the steady state to p and q.

Table 1: Signs of variations of strategies and steady state with respect to parameters.

	m_D	m_I	m_R	a_M	a_R	b	μ_M	μ_R	p	q
E_M	+	−	+	-	-	-	+	+	-	-
E_R	0	0	+	-	-	-	+	+	-	-
x_{ss}	+	-	-	-	+	-	+	-	-	+

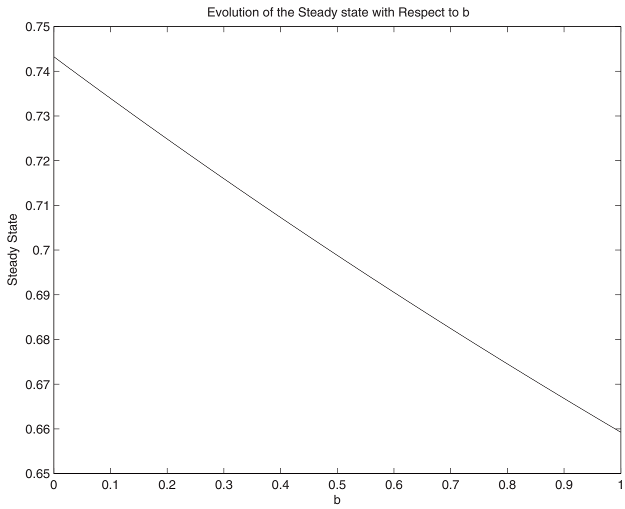


	m_D	m_I	m_R	a_M	a_R	b	r	μ_M	μ_R	p	q
μ_M	5	1.6	1.3	1.4	1.1	0.1	0.1		0.9	0.8	0.8

Figure 7: Sensitivity of the steady state to μ_M.

in the model would lead to a more realistic one and provide valuable insights on competitive issues in dual channels. This is an ongoing research by the authors. A second restriction in our model is the assumption that the switching intensities between the two channels are exogenous. One possible extension is to let them be a function of marketing effort or prices. In any event, this would lead to a structure which is not of the tractable linear-quadratic variety. Here, one would have to resort to numerical methods to obtain an equilibrium. A third limitation in our model is in the number of players. Introducing upstream and downstream competitions will certainly have an impact on the results.

We wish to conclude by raising the question of coordination/cooperation in the case of e-commerce between a manufacturer and her retailer. The manufacturer's decision to compete with her retailers by adding a new distribution channel is likely to generate some conflicts between the partners, and possibly also lead to inefficiencies (managerial, economic and strategic). The natural question is thus what kind of mechanisms one can think of that could improve a channel's profitability as well as individual profits. This is an interesting open-ended question.



	m_D	m_I	m_R	a_M	a_R	b	r	μ_M	μ_R	p	q
b	5	1.6	1.3	1.4	1.1		0.1	0.7	0.9	0.8	0.8

Figure 8: Sensitivity of the steady state to b.

Acknowledgments

This research was supported by NSERC Canada. We wish to thank an anonymous reviewer for helpful comments.

Appendix: Proof of Proposition 3.1

We apply a standard sufficient condition for a stationary Markov perfect Nash equilibrium and wish to find bounded and continuously differentiable functions $V_i(x)$, $i \in \{M, R\}$, satisfying, for all $0 \leq x(t) \leq 1$, the Hamilton–Jacobi–Bellman (HJB) equations

$$rV_M(x) = \underset{E_M \geq 0}{\text{Max}} \left\{ m_D x + m_I(1-x) - \frac{1}{2} a_M E_M^2 - \frac{1}{2} b x^2 + \frac{dV_M}{dx}(\mu_M E_M - \mu_R E_R - px + q(1-x)) \right\} \tag{16}$$

$$rV_R(x) = \underset{E_R \geq 0}{\text{Max}} \left\{ m_R(1-x) - \frac{1}{2} a_R E_R^2 + \frac{dV_R}{dx}(\mu_M E_M - \mu_R E_R - px + q(1-x)) \right\}. \tag{17}$$

Differentiating the right-hand sides and equating to zero gives

$$E_M = V_M' \frac{\mu_M}{a_M}, \tag{18}$$

$$E_R = -V_R' \frac{\mu_R}{a_R}. \tag{19}$$

Note that in (16) and (17), the maximands are concave in E_M and E_R respectively, thus yielding unique stationary feedback marketing effort rates.

Substituting for E_M and E_R from (18)–(19) in (16)–(17) leads to

$$rV_M(x) = m_D x + m_I(1-x) + \frac{1}{2a_M}\left(V_M' \mu_M\right)^2 \tag{20}$$

$$-\frac{1}{2}bx^2 + V_M'\left(V_R' \frac{\mu_R^2}{a_R} - px + q(1-x)\right),$$

$$rV_R(x) = m_R(1-x) + \frac{1}{2a_R}(V_R' \mu_R)^2 \tag{21}$$

$$+ V_R'\left(V_M' \frac{\mu_M^2}{a_M} - px + q(1-x)\right). \tag{22}$$

We conjecture that solutions to (16) and (17) will be quadratic:

$$V_M(x) = \frac{1}{2}Ax^2 + Bx + C, \tag{23}$$

$$V_R(x) = Dx + E, \tag{24}$$

in which A, B, C, D, E are constants. Substitute $V_M(x)$ and $V_R(x)$ from (20) and (21), as well as their derivatives $V_M'(x) = Ax + B$, $V_R'(x) = D$ into (18) and (19) and collect terms to obtain

$$r\left(\frac{1}{2}Ax^2 + Bx + C\right) = x^2\left(\frac{\mu_M^2 A^2}{2a_M} - \frac{b}{2} - A(p+q)\right)$$

$$+ x\left(m_D - m_I + \left(\frac{A\mu_M^2}{a_M} - p - q\right)B + A\left(q + D\frac{\mu_R^2}{a_R}\right)\right)$$

$$+ m_I + \frac{\mu_M^2 B^2}{2a_M} + \frac{DB\mu_R^2}{a_R} + qB,$$

$$r(Dx + E) = x\left(-m_R + D\left(\frac{A\mu_M^2}{a_M} - p - q\right)\right) + m_R$$

$$+ \frac{D^2\mu_R^2}{2a_R} + \frac{BD\mu_M^2}{a_M} + Dq.$$

By identification, we obtain

$$
\begin{aligned}
A &= \frac{a_M}{\mu_M^2}\left(\left(p+q+\frac{r}{2}\right)-\sqrt{\left(p+q+\frac{r}{2}\right)^2+\frac{b\mu_M^2}{a_M}}\right),\\
B &= \frac{m_D - m_I + A\left(q + D\frac{\mu_R^2}{a_R}\right)}{\frac{r}{2}+\sqrt{\left(p+q+\frac{r}{2}\right)^2+\frac{b\mu_M^2}{a_M}}},\\
C &= \frac{1}{r}\left(m_I + \frac{(B\mu_M)^2}{2a_M} + \frac{BD\mu_R^2}{a_R} + qB\right),\\
D &= \frac{-m_R}{\frac{r}{2}+\sqrt{\left(p+q+\frac{r}{2}\right)^2+\frac{b\mu_M^2}{a_M}}},\\
E &= \frac{1}{r}\left(m_R + \frac{(D\mu_R)^2}{2a_R} + D\left(q + B\frac{\mu_M^2}{a_M}\right)\right).
\end{aligned}
$$

To obtain an asymptotically stable steady state, choose for A the negative solution.

References

[1] Alba J., Lynch J., Weitz B., Janiszewski C., Lutz R., Sawyer A. and Wood S., Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces, *Journal of Marketing*, 61, 38–53, 1997.

[2] Balasubramanian S., Mail versus Mall: A strategic analysis of competition between direct marketers and conventional retailers, *Marketing Science*, 7(3), 181–195, 1998.

[3] Chiang W., Chhajed D. and Hess J., Direct marketing, indirect profits: A strategic analysis of dual-channel supply-chain design, *Management Science*, 49(1), 1–21, 2003.

[4] Fruchter G. and Tapiero C., Dynamic online and offline channel pricing for heterogeneous customers in virtual acceptance, to appear in *International Game Theory Review*.

[5] Geyskens I., Katrijn G. and Dekimpe M., The market valuation of Internet channel additions, *Journal of Marketing*, 66, 102–119, 2000.

[6] Hauser J., Urban G. and Weinberg B., How consumers allocate their time when searching for information, *Journal of Marketing Research*, 30, 452–466, 1993.

[7] Jarrar R., Martín-Herrán G. and Zaccour G., Markov perfect equilibrium advertising strategies of Lanchester duopoly model: A technical note, *Management Science*, 50(7), 995–1000, 2004.

[8] Jeuland A.P. and Shugan S.M., Coordination in marketing channels, in D.A. Gautchi (ed.), *Productivity and Efficiency in Distribution Systems*, Elsevier Publishing Co., New York, 1983.

[9] Jørgensen S., Sigué S.P. and Zaccour G., Dynamic cooperative advertising in a channel, *Journal of Retailing*, 76, 71–92, 2000.

[10] Jørgensen S., Sigué S.P. and Zaccour G., Stackelberg leadership in a marketing channel, *International Game Theory Review*, 3, 1–14, 2001.

[11] Jørgensen S., Taboubi S. and Zaccour G., Cooperative advertising in a marketing channel, *Journal of Optimization Theory and Applications*, 110, 145–158, 2001.

[12] Jørgensen S. and Zaccour G., *Differential Games in Marketing*, International Series in Quantitative Marketing, Kluwer Academic Publishers, Dordrecht, 2004.

[13] Karray S. and Zaccour G., Co-op advertising: A manufacturer's counterstrategy to store brand, *Les Cahiers du GERAD*, G–2004–27, HEC Montréal, 2004.

[14] Lal R. and Sarvary M., When and how is the Internet likely to decrease price competition? *Marketing Science*, 18(4), 485–503, 1999.

[15] Mills D., Private labels and manufacturer counterstrategies, *European Review of Agricultural Economics*, 26, 125–145, 1999.

[16] Mills D., Why retailers sell private labels? *Journal of Economics and Management Strategy*, 4(3), 509–528, 1995.

[17] Moorthy S. and Fader P., Strategic interaction within a channel, in Pellegrini and Reddy (eds), *Retail and Marketing Channels*, Routledge, London, 1989.

[18] Morton F. and Zettelmeyer F., The strategic positioning of store brands in retailer-manufacturer negotiations, to appear in *Review of Industrial Organization*, 2004.

[19] Narasimhan C. and Wilcox R., Private labels and the channel relationship: A cross-category analysis, *Journal of Business*, 71(4), 573–600, 1998.

[20] Raju J., Sethuraman R. and Dhar S., The introduction and performance of store brands, *Management Science*, 41(6), 957–978, 1995.

[21] Sarkar M., Butler B. and Steinfield C., Intermediaries and cyberintermediaries: a continuing role for mediating players in electronic marketplace, *Journal of Computer-Mediated Communications*, 1, 3, 1996.

[22] Shugan, S.M., Implicit understandings in channels of distribution, *Management Science*, 31(4), 435–460, 1985.

[23] Tsay A. and Agrawal N., Modeling conflict and coordination in multi-channel distribution systems: A review, in D. Simchi-Levi, D. Wu and M. Shen (eds.), *Handbook of Supply Chain Analysis in E-Business Era*, International Series in Operations Research and Management Science, Kluwer Academic Publishers, Dordrecht, 557–606, 2004.

[24] Zettelmeyer F., Expanding to the Internet: Pricing and communications when firms compete on multiple channels, *Journal of Marketing*, 38, 292–308, 2000.

Design Imitation in the Fashion Industry

Steffen Jørgensen
Department of Business and Economics
University of Southern Denmark
Campusvej 55, DK-5230, Denmark
stj@sam.sdu.dk

Andrea Di Liddo
Department of Economics, Mathematics and Statistics
University of Foggia
Via Gramsci 89-91
71100 Foggia, Italy
a.diliddo@unifg.it

Abstract

This chapter deals with the imitation of fashion products, an issue that attracts considerable interest in practice. Copying of fashion originals is a major concern of designers and, in particular, their financial backers. Fashion firms are having a hard time fighting imitations, but legal sanctions are not easily implemented in this industry. We study an alternative strategy that has been used by designers. Instead of fighting the imitators in the courtroom, designers fight them in the market. The designer markets her products in two separate markets. Typically, there is a "high-class" market in which the products are sold in exclusive stores at high prices. Customers in this market seek exclusivity and their utility diminishes when they see an increasing number of copies around. Their perception of the brand tends to dilute which poses a serious threat to a fashion company. The second market is a "middle-class" market in which there are many more buyers, and the fashion firm competes directly with the imitators in this market. This market can be used to practice price discrimination, to sell off leftover inventories, and to get a spin-off from the design. This chapter models the decision problems of the fashion firm and the imitators as a two-period game in which firms make pricing decisions and decisions on when to introduce their products in the markets. In addition, the fashion firm decides how much effort to spend to increase its brand image in the two markets.

1 Introduction

This chapter deals with product imitation, an issue that attracts considerable interest in practice. It occurs in many industries (e.g., furniture, toys, music, and fashion). Here we are concerned with *fashion*. A quick search of the World Wide Web suggests that a broad range of fashion designs are copied and imitations offered for sale at low prices: jewelry, handbags, leather belts and wallets, perfumes, watches, sunglasses, and mens' and womens' apparel.[1] This chapter considers one particular designer product.

A specific feature of fashion goods is the fact that a consumer is willing to pay a very high price for a product that she probably could buy much cheaper elsewhere. What makes this possible is the "brand integrity" or "brand image" that lies behind the fashion firm and its products. The natural enemy of brand integrity is "brand dilution," which occurs if too many people can be seen using the product, be it the original or a copy. Brand dilution leads to decreasing demand for the original product, and can be hard to reverse once the process is in motion.

The imitation of design originals is a major concern of designers and, in particular, their financial backers. Indeed, imitations of designer products can be produced and marketed very shortly after (and even before) the original appears on the market. Fashion firms wish to fight imitations, but legal actions are often infeasible. One reason is that in most countries, fashion designs cannot be patented. Various counter-strategies have been seen. Among these is the practice that a fashion firm buys up copy products and destroys them, usually in a public event witnessed by the press. (This has been done in the case of Cartier watches.) Another strategy is that the fashion firm offers consumers who have bought a copy to have it replaced, at no cost, with an original. This has been seen in the case of designer T-shirts and caps of the brand Von Dutch. The strategy is usable for items produced in smaller numbers ("limited editions").

This chapter studies the effects of a different strategy that has been employed by designers. Instead of fighting the imitators in the courtroom or buying up copies, designers fight imitators in the *market*. A typical instance of this practice can be described as follows. The designer firm markets its products in two separate markets. One market, A, is a "high-class" market in which the designer product is sold in exclusive stores at high prices and limited numbers. Typically, stores are owned or franchised by the fashion house. Customers in market A seek *exclusivity*, that is, their utility is highest when they know that they are the only ones who possess the product. If they see an increasing number of "similar" products around, their perception of the brand tends to dilute.

Another factor that can decrease the strength of a fashion brand name is the use of the name on products other than the fashion items themselves. One often sees that fashion houses use their name on a wide range of products (e.g., sunglasses and perfumes). In this case the fashion firm itself may contribute to the dilution of the brand.

Knowing that it cannot prevent imitations, the fashion firm enters the "middle-class" market, B. Thus, after having sold the product in market A for a limited

[1]The fashion imitation industry seems to be particularly flourishing in India and Southeast Asia.

time, the firm "moves the product" to market B. Typically this is accomplished by introducing a cheaper "diffusion line." (One example is the expensive Armani Via Borgo Nuevo line which diffuses into the less expensive Armani and Emporio Armani lines). In market B there are many more buyers and the products are sold at considerably lower prices than in market A. The stores selling the fashion product can be licensees of the fashion firm or manufacturer-owned "factory outlets."

The fashion firm can have several objectives for entering market B. One is to fight the imitations in the market, a second is to dispose of leftover inventory, and a third is to benefit from a spin-off of the original design.

This chapter gives a stylized image of the marketing of original designs and imitations in the fashion industry. We consider a fashion firm, henceforth denoted by E, and confine our interest to one specific product of firm E. The firm spends considerable amounts of money in developing new designs of the product. Although the manufacturing costs may not be that high, the total costs of the firm typically include sizeable expenditures on its creative staff and on advertising.

In market B there is a competing firm (or a fringe of firms), henceforth denoted by P, producing an imitation of E's product. The aim of P is to have a product which "looks like" that of E, with the purpose of free-riding on the brand image of the fashion firm. Firm P does not spend anything on developing new designs of its own; the firm simply adjusts the design of their product to imitate that of E. The costs of firm P are substantially lower than those of firm E and the products of firm P sell at considerably lower prices in nonexclusive stores. The imitating firm does nothing to create a demand for its product. Demand for the imitation is generated by consumers who wish to imitate the consumption of wealthy, trend-setting consumers. Thus, consumption imitation creates a demand for the imitator's product.

Studies of fashion imitation problems using game theoretic methods are very few. In fact, we only know of the one in Caulkins et al. [1]. These authors are interested in explaining the occurrence of fashion cycles and suggest an optimal control model in which state variables are the positions of the fashion firm and the imitator along a one-dimensional product space. However, the imitator is not a decision maker. She always chases the innovator and her position is determined by the position of the innovator (which, in turn, is determined by the innovator's design efforts). Thus, our work can be seen as a first exploration of the (realistic) situation in which a fashion firm faces imitators in the market place. An account of fashion imitation in practice is found in the article in *The Economist*, March 6, 2004.

The chapter proceeds as follows. Section 2 develops a two-period game model, incorporating pricing decisions of both firms and image building expenditures of the fashion firm E. The extension to a general number of periods is straightforward, but does not add to our understanding of the imitation problem, at least in the present setup. Section 3 contains the analysis of the game, and Section 4 presents our conclusions.

2 A Two-Period Game Model

Firm E is the leading fashion firm and P is a firm that imitates the design of the product of firm E.[2] The product of E is sold in markets A ("high-class") and B ("middle-class"), while P sells its imitations in market B only. Thus, the two firms compete in market B, but not in market A. However, the sales volume in market B may influence the image of the designer brand in markets A (for details, see Section 2.3). For the particular product under consideration, we assume that firm E is a monopolist in market A.

2.1 Timing of Decisions

Firm E introduces a new design in market A at each of the time instants $n = 0, 1, 2$. One incentive for the continual introduction of new designs is the fact that designs will be copied by imitating firms, but this is not the main objective. Coming up continually with new designs is the *raison d'être* of fashion firms.

The game terminates at time $n = 2$. When a new design is introduced, the existing design is withdrawn from markets A and B. A design having been introduced in market A at time $n \in \{0, 1\}$ is removed from market A and introduced in market B at time t^n. Thus, at any instant of time there is only one design of firm E in the market.

Time instants t^n are decision variables of firm E and must satisfy the constraints

$$n + \kappa_E \le t^n \le n + 1, \tag{1}$$

where $\kappa_E \in]0, 1[$ is a constant. The interpretation of the left-hand inequality in (1) is that the design original will be exclusive to the buyers in market A from time n to time t^n, and that firm E guarantees its buyers in market A that the product will be sold exclusively in that market at least until time $n + \kappa_E$. (Clearly, an imitation may appear in market B before time t^n.) Note that we have assumed that a new design remains in market A at least for a time κ_E, in both periods.

Firm P observes instantaneously a new design introduced by E and starts to prepare the production and marketing of its imitation. Let τ^n denote the time of the introduction of P's product in market B in period n and impose the constraint

$$n + \kappa_P \le \tau^n \le n + 1,$$

where $\kappa_P \in [0, 1[$ is a constant. Thus, the earliest time at which the imitation can be introduced is at $n + \kappa_P$. If $\kappa_P > 0$, the interpretation is that the imitator needs some time to prepare to market its product. If $\kappa_P = 0$, the imitator has the option

[2] The letters designating the firms are borrowed from pursuit-evasion games: "E" for evader, "P" for pursuer. Firm P can also represent a fringe of homogeneous firms.

of introducing its product simultaneously with the new design.[3] Note that we have assumed that the earliest time to introduce the imitation is the same in both periods.

For the earliest introduction times κ_E and κ_P we assume $\kappa_P < \kappa_E$. This means that the imitator could introduce its product before the time where the design original can be moved from market A to B.[4]

2.2 Demand Functions

Let "period n" be the time interval from time n to time $n+1$. In what follows, a superscript n indicates a period number. To specify the demand functions in the three markets we need some notation. All demands defined below are deterministic and should be interpreted as "demand rates per unit of time."

Demand	Price	Market	Product	Period
Q_A^n	p_A^n	A	original	monopoly
Q_{BE}^n	p_{BE}^n	B	original	duopoly
Q_{BMP}^n	p_{BMP}^n	B	imitation	monopoly
Q_{BP}^n	p_{BP}^n	B	imitation	duopoly

We assume that the price p_A^n of the original in market A is fixed, at some (high) level, $\bar{p}$. This level is the standard profit-maximizing price of a monopolist firm.

Let X_A^n and X_B^n represent the brand image of the design original, as assessed by consumers in markets A and B, respectively. Our hypothesis is that consumers in these markets have different perceptions of the brand and are influenced by the prices charged by the two firms. Demand rates are specified as follows:

$$Q_A^n = \eta X_A^n \tag{2}$$

$$Q_{BE}^n = \alpha_{BE} X_B^n - \theta_E p_{BE}^n - \gamma\left(p_{BE}^n - p_{BP}^n\right) \tag{3}$$

$$Q_{BMP}^n = \alpha_{BMP} X_B^n - \delta p_{BMP}^n \tag{4}$$

$$Q_{BP}^n = \alpha_{BP} X_B^n - \theta_P p_{BP}^n + \gamma\left(p_{BE}^n - p_{BP}^n\right) \tag{5}$$

in which $\eta, \alpha_{BE}, \theta_E, \gamma, \alpha_{BMP}, \alpha_{BP}, \theta_P$, and δ are positive constants.

The assumption in (2) is that demand in market A, in which firm E is a monopolist, depends on the brand image only. Buyers in this market are willing to pay, and do pay the high price $\bar{p}$, but their demand diminishes with their assessment of the brand image.

Equations (3) and (4) mean that the demand of both firms increase with the brand image during the duopoly period in market B. The α's being different means that

[3]Copies could actually be on the market *before* the original is for sale. Magdo writes that "more advanced technology makes it possible to see high-quality copies appear in stores before the original has even hit the market" (Magdo [5], p. 1). Note that such a situation is not possible in our framework.

[4]Otherwise, firm E could have a monopoly period in market B which seems less plausible, in view of the speed at which imitations of new designs appear on the market.

buyers of the original and the imitation may react differently to their perceptions of the brand image. The second term on the right-hand side reflects the direct influence of a firm's own price on its demand. Here, a plausible assumption could be $\theta_E < \theta_P$, that is, those customers who buy the fashion product are less sensitive to a change in product price than those who buy the imitation. The third term reflects that consumers also react on the "price differential" $p_{BE}^n - p_{BP}^n$ between the two products. Thus, ceteris paribus, consumers buy the product with the lowest price.

The specification in (5) means that the demand in market B, during the monopoly period of firm P, depends on the consumers' assessment of the brand and the price of the imitation.

Using the demand functions introduced above, we define the sales functions

$$S_A^n = (t^n - n)Q_A^n \tag{6}$$

$$S_{BE}^n = (n + 1 - t^n)Q_{BE}^n \tag{7}$$

$$S_{BP}^n = (t^n - \tau^n)Q_{BMP}^n + (n + 1 - t^n)Q_{BP}^n. \tag{8}$$

Equation (6) says that sales of firm E in market A are the sales rate Q_A^n over the time interval from time n to time t^n where the designer product is removed from market A. Equation (7) means that sales of firm E in market B are the sales rate Q_{BE}^n over the remaining period from t^n to $n + 1$. The specification in (8) says that the sales of firm P in market B are the sales rate Q_{BMP}^n during the monopoly period (from τ^n to t^n), plus the sales rate Q_{BP}^n over the duopoly period.

2.3 Brand Images

Brand images X_A^n and X_B^n can be seen as the combined set of beliefs (perceptions) about the brand of firm E, held by buyers in markets A and B, respectively. These perceptions influence demand in markets A and B. Brand images are supposed to evolve over time according to

$$X_A^{n+1} = X_A^n + \beta_A u_A^n - \rho_A\big[S_{BE}^n + S_{BP}^n\big]; \quad X_A^0 \text{ fixed} \tag{9}$$

$$X_B^{n+1} = X_B^n + \beta_B u_B^n; \quad X_B^0 \text{ fixed.} \tag{10}$$

In (9) and (10), $\beta_A > 0$, $\beta_B \geq 0$, $\rho_A > 0$ are constants and $u_A^n \geq 0$, $u_B^n \geq 0$ are efforts (e.g., image advertising) made by firm E to improve its goodwill in markets A and B.

The interpretation of (9) is as follows. The brand image in market A dilutes in proportion to the sales in market B. As customers in market A observe an increasing number of imitations appearing, their image of the brand is negatively affected. This effect can be modified, however, if the fashion firm spends image building efforts ($u_A^n > 0$).

A driving force behind the demand for highly priced fashion products is the "snob" effect (Leibenstein [4]). This effect refers to the phenomenon that some consumers' demand for a product decreases because others are consuming it. In our model, the snob effect is not modeled directly in the demand function in market A, but affects demand indirectly due to the brand image dynamics in market A.

The assumption in (10) is that the sales volume in market B does not affect (negatively) market B consumers' perception of the brand; consumers in this market do not care about the number of products (original or fake) being around.

If the fashion firm wishes to have a strong brand image, the dynamics in (9) suggest that this aim can be accomplished by (i) trying to reduce the sales of imitations and (ii) advertising to build up the brand image. Equation (9) shows that the fashion firm's *own sales* in market B are also detrimental to brand integrity (as assessed by market A customers). Firm E faces a dilemma here: It knows that its designs will be imitated, but since E cannot fight the imitator P in market A (where P's product is not sold), firm E faces its competitor in market B. This will, however, tend to dilute the brand image in market A.

Inserting from the sales functions into (9) and (10) yields

$$\begin{aligned} X_A^{n+1} &= X_A^n + \beta_A u_A^n - \rho_A(t^n - \tau^n)\big[\alpha_{BMP} X_B^n - \delta p_{BMP}^n\big] \qquad (11)\\ &\quad - \rho_A(n+1-t^n)\big[(\alpha_{BE}+\alpha_{BP})X_B^n - \theta_E p_{BE}^n - \theta_P p_{BP}^n\big]. \\ X_B^{n+1} &= X_B^n + \beta_B u_B^n. \qquad (12) \end{aligned}$$

Equations (11) and (12) show that having a good image in market B actually hurts the image in market A: the larger X_B^n, the smaller X_A^{n+1}. On the other hand, having a good brand image in market B stimulates demand for the product of firm E in this market—but also the demand of the imitator.

The reader should note the implicit assumption made in (11) and (12): X_B is not affected by X_A. One may, more realistically, suppose that X_B is increasing with X_A, which means that consumers in market B follow consumers in market A in their regard for the fashion product. Then one should add some increasing function of X_A on the right-hand side of (12). We are indebted to an anonymous reviewer for making this observation. However, the suggested change in the dynamics (12) makes the problem considerably more complicated. In an analysis of the modified problem we did not succeed in obtaining any interpretable analytical results.

2.4 Cost Functions

Before the game is played, firm E is assumed to have decided to introduce new designs at the start of each of the two time periods. The implication is that E's costs of developing new designs are sunk and can be disregarded. The variable production costs of firm E are denoted C_E^n and depend on production (sales) volume only. Assuming a linear cost function we have

$$C_E^n = c_E\big[S_A^n + S_{BE}^n\big],$$

in which c_E is a positive constant. For simplicity it is assumed that goods produced for markets A and B have the same unit production cost.

The production costs of firm P are denoted C_P^n. These costs depend on the total production (sales) volume:

$$C_P^n = c_P S_{BP}^n,$$

in which c_P is a positive constant.

Denote by K_E^n the costs of image building efforts. These costs are given by

$$K_E^n = \frac{1}{2}\big[k_A (u_A^n)^2 + k_B (u_B^n)^2\big]$$

in which k_A and k_B are positive constants. The quadratic costs reflect a hypothesis that brand image building activities are subject to diminishing marginal returns.

2.5 Profit Functions

To save a bit on notation, define the constant $\bar{\eta} \triangleq (\bar{p} - c_E)\eta$. The profit of firm E in period n is given by

$$\begin{aligned}\pi_E^n = &-\frac{1}{2}\big[k_A (u_A^n)^2 + k_B (u_B^n)^2\big] + \bar{\eta}(t^n - n) X_A^n \\ &+ \big(p_{BE}^n - c_E\big)(n + 1 - t^n)\big[\alpha_{BE} X_B^n - \theta_E p_{BE}^n - \gamma (p_{BE}^n - p_{BP}^n)\big],\end{aligned}$$

and that of firm P is

$$\begin{aligned}\pi_P^n = &\big(p_{BMP}^n - c_P\big)(t^n - \tau^n)\big[\alpha_{BMP} X_B^n - \delta p_{BMP}^n\big] \\ &+ \big(p_{BP}^n - c_P\big)(n + 1 - t^n)\big[\alpha_{BP} X_B^n - \theta_P p_{BP}^n + \gamma (p_{BE}^n - p_{BP}^n)\big].\end{aligned}$$

In period n, firm E has decision variables u_A^n, u_B^n, t^n, p_{BE}^n and firm P has decision variables τ^n, p_{BMP}^n, p_{BP}^n. Decisions of the two firms are made simultaneously and independently.

It is straightforward to demonstrate that firm P will always choose its entry time τ^n as the minimal value, $n + \kappa_P$. The reason is that the imitator gains nothing by postponing the introduction of its product. Hence, in what follows we set $\tau^n = n + \kappa_P$ for all n.

2.6 Behavioral Assumptions

The imitator's decisions influence the brand image X_A, but the latter does not affect the payoff of the imitator. For this reason the imitator can disregard X_A (and therefore also its dynamics). On the other hand, although the brand image X_B is payoff relevant for the imitator, the firm knows that its decisions have no influence upon X_B. Consequently, the imitator can disregard X_B (and its dynamics). The upshot is that the imitating firm will act *myopically*, i.e., the firm makes its decisions on a period-by-period basis. The reader should note that the imitator's myopism is a result of the structure of the game and hence we cannot foresee what would happen if the imitator did not behave myopically. To do so one would need a model in which the rational behavior of the imitator would not be myopic. Then one could suppose that the imitator was irrational and behaved myopically.

State variables X_A and X_B are payoff relevant to firm E, and the firm's decisions influence the dynamics of both variables. Thus, if it wishes, firm E can condition its actions upon the state variables. Since the strategy of the imitator at most will depend on time, the fashion firm can without loss restrict itself to a strategy that depends on time only (Fudenberg and Tirole ([2], p. 530).

The profit function of the fashion firm E is given by

$$\Pi_E = \sum_{n=0}^{N-1} (i_E)^n \pi_E^n + (i_E)^N \big[\sigma_A X_A^N + \sigma_B X_B^N\big],$$

where $N = 2$, $i_E \in (0, 1]$ is a discount factor and σ_A, σ_B are positive constants. The term in square brackets is a salvage value function which assesses the value to the fashion firm of having brand images X_A^N and X_B^N at the end of the planning horizon.

The imitating firm P maximizes, period by period, its profit π_P^n for $n \in \{0, 1\}$.

3 Analysis of the Game

An equilibrium of the two-period game is defined by a pair of decision paths

$$\big\{\hat{u}_A^n, \hat{u}_B^n, \hat{t}^n, \hat{p}_{BE}^n\big\}_{n\in\{0,1\}}, \quad \big\{\hat{p}_{BMP}^n, \hat{p}_{BP}^n\big\}_{n\in\{0,1\}}$$

such that each firm's decisions maximize its objective while taking the decision path of the other firm as given.

Firm E determines at time zero its decision path so as to maximize

$$\begin{aligned}
\Pi_E = &-\frac{1}{2}\big[k_A(u_A^0)^2 + k_B(u_B^0)^2\big] + \bar{\eta} t^0 X_A^0 \\
&+ \big(p_{BE}^0 - c_E\big)(1 - t^0)\big[\alpha_{BE} X_B^0 - \theta_E p_{BE}^0 - \gamma(p_{BE}^0 - p_{BP}^0)\big] \\
&+ i_E \Bigg\{-\frac{1}{2}\big[k_A(u_A^1)^2 + k_B(u_B^1)^2\big] + \bar{\eta}(t^1 - 1) X_A^1 \\
&+ \big(p_{BE}^1 - c_E\big)(2 - t^1)\big[\alpha_{BE} X_B^1 - \theta_E p_{BE}^1 - \gamma(p_{BE}^1 - p_{BP}^1)\big]\Bigg\} \\
&+ (i_E)^2\{\sigma_A X_A^2 + \sigma_B X_B^2\},
\end{aligned}$$

subject to the brand image dynamics

$$\begin{aligned}
X_A^{n+1} &= X_A^n + \beta_A u_A^n - \rho_A\big(t^n - k_P^n\big)\big[\alpha_{BMP} X_B^n - \delta p_{BMP}^n\big] \\
&\quad - \rho_A(n + 1 - t^n)\big[(\alpha_{BE} + \alpha_{BP}) X_B^n - \theta_E p_{BE}^n - \theta_P p_{BP}^n\big], \\
X_B^{n+1} &= X_B^n + \beta_B u_B^n,
\end{aligned}$$

where $n \in \{0, 1\}$ and X_A^0, X_B^0 are given.

At time zero, firm P determines its decisions $\{\hat{p}^0_{BMP}, \hat{p}^0_{BP}\}$ so as to maximize

$$\pi^0_P = \left(p^0_{BMP} - c_P\right)(t^0 - \kappa_P)\left[\alpha_{BMP} X^0_B - \delta p^0_{BMP}\right] + \left(p^0_{BP} - c_P\right)(1 - t^0)\left[\alpha_{BP} X^0_B - \theta_P p^0_{BP} + \gamma(p^0_{BE} - p^0_{BP})\right],$$

in which X^0_A, X^0_B are given. At time one, firm P determines its decisions $\{\hat{p}^1_{BMP}, \hat{p}^1_{BP}\}$ to maximize

$$\pi^1_P = \left(p^1_{BMP} - c_P\right)(t^1 - (1 + \kappa_P))\left[\alpha_{BMP} X^1_B - \delta p^1_{BMP}\right] + \left(p^1_{BP} - c_P\right)(2 - t^1)\left[\alpha_{BP} X^1_B - \theta_P p^1_{BP} + \gamma\left(p^1_{BE} - p^1_{BP}\right)\right],$$

in which X^1_A, X^1_B are given.

3.1 Period 1 Equilibrium Decisions

Prices of the imitation in market B are

$$\hat{p}^1_{BMP} = \frac{1}{2\delta}\left[\alpha_{BMP} X^1_B + \delta c_P\right], \tag{13}$$

$$\hat{p}^1_{BP} = \frac{1}{2(\gamma + \theta_P)}\left[\alpha_{BP} X^1_B + \gamma\left(c_P + p^1_{BE}\right) + c_P \theta_P\right]. \tag{14}$$

Advertising efforts of the fashion firm are

$$\hat{u}^1_A = \frac{i_E \sigma_A \beta_A}{k_A}, \quad \hat{u}^1_B = \frac{i_E \sigma_B \beta_B}{k_B}. \tag{15}$$

The results in (15) say that the marginal advertising cost in period 1 equals the (one-period) discounted value of the marginal increase in the salvage value of the brand image.[5] Efforts decrease if costs increase and increase if efficiency β and/or the salvage value σ increases. Efforts also increase if the firm becomes more farsighted ($i_E \to 1$). These results are as expected.

The price of the fashion product in market B is

$$\hat{p}^1_{BE} = \frac{1}{2(\gamma + \theta_E)}\left[\alpha_{BE} X^1_B + \gamma\left(c_E + p^1_{BP}\right) + c_E \theta_E + i_E \theta_E \rho_A \sigma_A\right]. \tag{16}$$

Equations (14) and (16) show that the price of one product increases if the price of the other increases. This is the standard result in a pricing game.

Solving (14) and (16) yields the duopoly prices, expressed as functions of X^1_B :

$$\hat{p}^1_{BE} = \frac{1}{3\gamma^2 + 4\gamma(\theta_E + \theta_P) + 4\theta_E\theta_P}\left\{2c_E(\gamma + \theta_E)(\gamma + \theta_P) + c_P\gamma(\gamma + \theta_P) + 2i_E\theta_E\rho_A\sigma_A(\gamma + \theta_P) + [2\alpha_{BE}(\gamma + \theta_P) + \gamma\alpha_{BP}]X^1_B\right\} \tag{17}$$

$$\hat{p}^1_{BP} = \frac{1}{3\gamma^2 + 4\gamma(\theta_E + \theta_P) + 4\theta_E\theta_P}\left\{2c_P(\gamma + \theta_E)(\gamma + \theta_P) + c_E\gamma(\gamma + \theta_E) + i_E\theta_E\rho_A\sigma_A\gamma + [2\alpha_{BP}(\gamma + \theta_E) + \gamma\alpha_{BE}]X^1_B\right\}.$$

[5] If there were no salvage values, firm E would have no incentive to increase the brand images in period 1, and no advertising should be done in that period.

The higher the brand image X_B^1 at the start of period 1, the higher the prices. The idea is that the negative impact on demand of a high price can be offset by a strong brand image. We note that although the imitating firm does not care about the brand image, it enjoys the benefits of a strong image.

The time to move the fashion original from market A to B is determined by the sign of the derivative

$$\begin{aligned}\frac{\partial \Pi_E}{\partial t^1} &= i_E\{\bar{\eta}X_A^1 - (\hat{p}_{BE}^1 - c_E)[\alpha_{BE}X_B^1 - \theta_E\hat{p}_{BE}^1 - \gamma(\hat{p}_{BE}^1 - \hat{p}_{BP}^1)]\}\\ &\quad + (i_E)^2\sigma_A\rho_A[-(\alpha_{BMP}X_B^1 - \delta p_{BMP}^1) + (\alpha_{BE} + \alpha_{BP})X_B^1\\ &\quad - (\theta_E\hat{p}_{BE}^1 + \theta_P p_{BP}^1)] \qquad (18)\end{aligned}$$

and we have

$$\frac{\partial \Pi_E}{\partial t^1}\left\{\begin{array}{c} > \\ = \\ < \end{array}\right\} 0 \Longleftarrow \hat{t}^1 \left\{\begin{array}{c} = 2 \\ \in [1+\kappa_E, 2] \\ 1+\kappa_E \end{array}\right\}. \qquad (19)$$

The result in (19) means that the fashion original should be moved from market A to market B either as early or as late as possible. Clearly, this result is an extreme one and excludes a fine-tuning of the decision t^1. It is due to the linearity of the profit function with respect to t^1.

The derivative in (18) exhibits three effects of an increase in t^1 :

(1) $\bar{\eta}X_A^1$ is the increased profit earned on market A in period 1
(2) $-(\hat{p}_{BE}^1 - c_E)[\alpha_{BE}X_B^1 - \theta_E\hat{p}_{BE}^1 - \gamma(\hat{p}_{BE}^1 - \hat{p}_{BP}^1)]$ is the loss of profit in market B in period 1
(3) $(i_E)^2\sigma_A\rho_A[-(\alpha_{BMP}X_B^1 - \delta p_{BMP}^1) + (\alpha_{BE} + \alpha_{BP})X_B^1 - (\theta_E\hat{p}_{BE}^1 + \theta_P p_{BP}^1)]$ is the present value of the gain (or loss) of salvage value of X_A^2 at the end of period 1.

For a sufficiently large value of the brand image X_A^1, the derivative in (18) is positive and the product is never introduced in market B. This is intuitive since when the brand has a strong image in market A it is more profitable to keep the product in this market as long as possible. On the other hand, if the lost profit in market B is sufficiently large, the product should be moved to market B as early as possible.

Note that

$$\rho_A[-(\alpha_{BMP}X_B^1 - \delta p_{BMP}^1) + (\alpha_{BE} + \alpha_{BP})X_B^1 - (\theta_E\hat{p}_{BE}^1 + \theta_P p_{BP}^1)] = \frac{\partial X_A^2}{\partial t^1}$$

is the effect on the final brand image X_A^2 of keeping the product in market A for a marginally longer time. This effect can be positive or negative: If there is a sufficiently large gain of salvage value, it pays to keep the product in market A throughout period 1. On the other hand, if there is a large loss of salvage value, it pays to move the product to market B as soon as possible. If this effect is positive and sufficiently large, the product should be in market A throughout period 1.

Remark 3.1. For another interpretation, suppose that $\theta_E = \theta_P = 0$.[6] Then, if $\alpha_{BE} + \alpha_{BP} > \alpha_{BMP}$, we have

$$-\left(\alpha_{BMP} X_B^1 - \delta p_{BMP}^1\right) + (\alpha_{BE} + \alpha_{BP}) X_B^1 > 0,$$

that is, there is a gain of salvage value.

3.2 Period 0 Equilibrium Decisions

Prices of the imitation are structurally equivalent to that in period 1:

$$\hat{p}_{BMP}^0 = \frac{1}{2\delta}\left[\alpha_{BP} X_B^0 + \delta c_P\right] \tag{20}$$

$$\hat{p}_{BP}^0 = \frac{1}{2(\gamma + \theta_P)}\left[\alpha_{BP} X_B^0 + \gamma\left(p_{BE}^0 + c_P\right) + c_P \theta_P\right]. \tag{21}$$

The advertising effort in market A is

$$\hat{u}_A^0 = \frac{i_E \beta_A [(\hat{t}^1 - 1)\bar{\eta} + i_E \sigma_A]}{k_A}. \tag{22}$$

The result says that the marginal advertising cost $k_A \hat{u}_A^0$ is the sum of the present value of the marginal revenue obtained in market A in period 1, $i_E \beta_A (\hat{t}^1 - 1)\bar{\eta}$, and the discounted marginal salvage value, $(i_E)^2 \beta_A \sigma_A$, of the terminal brand image X_A^2.

Sensitivity results are, by and large, the same as for $\hat{u}_A^1$, but we note the additional term $(\hat{t}^1 - 1)\bar{\eta}$. This means that there is a positive relationship between advertising effort in market A in period 0 and the length $\hat{t}^1 - 1$ of the period during which the product is sold in market A in period 1. If the product is sold for a longer time in market A in period 1, more advertising effort should be used in period 0. The intuition is that then it pays to build up the brand image X_A^1 on which demand in period 1 will depend.

The advertising effort in market B is

$$\begin{aligned}\hat{u}_B^0 = {} & \frac{i_E \beta_B (2 - \hat{t}^1)(\hat{p}_{BE}^1 - c_E)\alpha_{BE}}{k_B} \\ & + \frac{(i_E)^2 \beta_B}{k_B}\{\sigma_B - \sigma_A \rho_A [(\hat{t}^1 - (1 + \kappa_P))\alpha_{BMP} + (2 - \hat{t}^1)(\alpha_{BE} + \alpha_{BP})]\}.\end{aligned} \tag{23}$$

Using (12) and (17) shows that $\hat{u}_B^0$ is linearly increasing in X_B^0. What drives this result is the fact that brand image has a positive impact on demand. A similar result as that in (23) occurs in, e.g., Jørgensen et al. [3] where advertising also increases with the brand image (although at a decreasing rate).

[6] This means that the demand functions in the duopoly period in market B depend on the price differential $\gamma(p_{BE}^1 - p_{BP}^1)$ only.

Remark 3.2. Note that brand image X_B^0 has two impacts: a high level of X_B^0 provides, by the brand image dynamics in (12), a high level of X_B^1, but it also implies a high level of effort u_B^0 (cf. (23) which increases X_B^1.

The price of the fashion original is

$$\hat{p}_{BE}^0 = \frac{1}{2(\gamma + \theta_E)}\big[\alpha_{BE}X_B^0 + \gamma\big(c_E + p_{BP}^0\big) + \theta_E c_E \\ + i_E\theta_E\rho_A(\bar{\eta}(t^1 - 1) + i_E\sigma_A)\big]. \tag{24}$$

Solving (21) and (24) provides the duopoly prices, expressed as functions of X_B^0 :

$$\hat{p}_{BE}^0 = \frac{1}{3\gamma^2 + 4\gamma(\theta_E + \theta_P) + 4\theta_E\theta_P}\big\{2c_E(\gamma + \theta_E)(\gamma + \theta_P) + c_P\gamma(\gamma + \theta_P) \\ + 2(\gamma + \theta_P)i_E\theta_E\rho_A[\bar{\eta}(\hat{t}^1 - 1) + i_E\sigma_A] + [2\alpha_{BE}(\gamma + \theta_P) + \gamma\alpha_{BP}]X_B^0\big\}$$

$$\hat{p}_{BP}^0 = \frac{1}{3\gamma^2 + 4\gamma(\theta_E + \theta_P) + 4\theta_E\theta_P}\big\{2c_P(\gamma + \theta_E)(\gamma + \theta_P) + c_E\gamma(\gamma + \theta_E) \\ + \gamma i_E\theta_E\rho_A[\bar{\eta}(\hat{t}^1 - 1) + i_E\sigma_A] + [2\alpha_{BP}(\gamma + \theta_E) + \gamma\alpha_{BE}]X_B^0\big\}. \tag{25}$$

The time to move the fashion original to market B is determined by the sign of the derivative

$$\frac{\partial\Pi_E}{\partial t^0} = \bar{\eta}X_A^0 - \big(\hat{p}_{BE}^0 - c_E\big)\big[\alpha_{BE}X_B^0 - \theta_E\hat{p}_{BE}^0 - \gamma\big(\hat{p}_{BE}^0 - p_{BP}^0\big)\big] \\ + i_E\rho_A[\bar{\eta}(\hat{t}^1 - 1) + i_E\sigma_A]\big[-(\alpha_{BMP}X_B^0 - \delta p_{BMP}^0) \\ + (\alpha_{BE} + \alpha_{BP})X_B^0 - \big(\theta_E\hat{p}_{BE}^0 + \theta_P p_{BP}^0\big)\big] \tag{26}$$

and we have

$$\frac{\partial\Pi_E}{\partial t^0}\left\{\begin{array}{c} > \\ = \\ < \end{array}\right\} 0 \Longleftarrow \hat{t}^0\left\{\begin{array}{c} = 1 \\ \in [\kappa_E, 1] \\ \kappa_E \end{array}\right\}.$$

As in period 1, the fashion product should be moved from market A to market B as early or as late as possible. The derivative in (26) exhibits four effects of an increase in t^0 (that is, by extending marginally the period during which the product is kept in market A in period 0):

(1) $\bar{\eta}X_A^0$ is the increased profit earned in market A in period 0
(2) $-(\hat{p}_{BE}^0 - c_E)[\alpha_{BE}X_B^0 - \theta_E\hat{p}_{BE}^0 - \gamma(\hat{p}_{BE}^0 - \hat{p}_{BP}^0)]$ is the loss of profit in market B in period 0
(3) $i_E\rho_A\bar{\eta}(\hat{t}^1 - 1)[-(\alpha_{BMP}X_B^0 - \delta p_{BMP}^0) + (\alpha_{BE} + \alpha_{BP})X_B^0 - (\theta_E\hat{p}_{BE}^0 + \theta_P p_{BP}^0)]$ is the (present value of the) effect on profit in market A in period 1
(4) $i_E^2\sigma_A\rho_A[-(\alpha_{BMP}X_B^0 - \delta p_{BMP}^0) + (\alpha_{BE} + \alpha_{BP})X_B^0 - (\theta_E\hat{p}_{BE}^0 + \theta_P p_{BP}^0)]$ is the present value of the gain (or loss) of salvage value of X_A^2.

Effects 1, 2, and 4 have the same interpretations as in period 1. As to the third effect, note that

$$\frac{\partial X_A^1}{\partial t^0} = i_E \rho_A \big[- \big(\alpha_{BMP} X_B^0 - \delta p_{BMP}^0\big) + (\alpha_{BE} + \alpha_{BP}) X_B^0 \\ - \big(\theta_E \hat{p}_{BE}^0 + \theta_P p_{BP}^0\big)\big]$$

is the effect on brand image X_A^1 of keeping the product in market A for a (marginally) longer time during period 0. If this effect is positive and sufficiently large, the product should be in market A throughout period 0.

3.3 Comparing Decisions over Time

We shall need the following

Assumption 1. $\theta_E = 0$.

The assumption means that buyers of the fashion product in market B react on price differences only (and not on the absolute price of the fashion product).

Using (17) and (25), invoking Assumption 1, shows that

$$\hat{p}_{BE}^1 \geq \hat{p}_{BE}^0, \quad \hat{p}_{BP}^1 \geq \hat{p}_{BP}^0, \tag{27}$$

which means that duopoly prices are nondecreasing over time. Actually, for $\hat{u}_B^0 > 0$ they increase and for $\hat{u}_B^0 = 0$ they are constant over time. Thus, by its advertising expenditure in period 0, the fashion firm can influence the price development in the duopoly period in market B. The result is driven by the fact that $X_B^1 = X_B^0 + \beta_B \hat{u}_B^0$ and that duopoly prices in period 1 depend on X_B^1.

Since the brand image X_B is nonincreasing, one obtains from (13) and (20)

$$\hat{p}_{BMP}^1 \geq \hat{p}_{BMP}^0.$$

The imitator increases its price [keeps it constant] in the monopoly period in market B if $\hat{u}_B^0 > 0$ [$\hat{u}_B^0 = 0$].

To compare the fashion firm's advertising efforts over time we need the following.

Assumption 2. $i_E = 1$.

The assumption says that the fashion firm does not discount the future. It is easy to see from (15) and (22) that

$$\hat{u}_A^0 > \hat{u}_A^1.$$

Advertising efforts are decreasing over time in market A. The reason is that advertising in period 0 has two impacts: to increase the image X_A^1 and to increase the salvage value. Advertising in period 1 only affects the salvage value.

Using (15) and (23) provides

$$\hat{u}_B^0 \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} \hat{u}_B^1 \Longleftrightarrow (2 - \hat{t}^1)\big(\hat{p}_{BE}^1 - c_E\big)\alpha_{BE} \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} \tag{28}$$
$$\sigma_A \rho_A [(\hat{t}^1 - (1 + \kappa_P))\alpha_{BMP} + (2 - \hat{t}^1)(\alpha_{BE} + \alpha_{BP})].$$

Suppose that $\hat{t}^1 = 2$, i.e., the product is kept in market A throughout period 1. Then (28) yields $\hat{u}_B^0 < \hat{u}_B^1$. Thus, advertising in market B is *increased* in period 1, despite the fact that the product is kept in market A throughout period 1. This may seem counterintuitive. What provides the result is the positive term $\sigma_A \rho_A (2 - (1 + \kappa_P))\alpha_{BMP}$ which is the loss of salvage value of X_A^2 caused by the imitator having a monopoly in market B throughout period 1.

The time to move the fashion product from market A to B is given by the signs of the derivatives

$$\begin{aligned} \frac{\partial \Pi_E}{\partial t^0} &= \bar{\eta} X_A^0 - \big(\hat{p}_{BE}^0 - c_E\big)\big[\alpha_{BE} X_B^0 - \theta_E \hat{p}_{BE}^0 - \gamma\big(\hat{p}_{BE}^0 - p_{BP}^0\big)\big] \\ &\quad + i_E \rho_A [\bar{\eta}(\hat{t}^1 - 1) + i_E \sigma_A]\big[- \big(\alpha_{BMP} X_B^0 - \delta p_{BMP}^0\big) \\ &\quad + (\alpha_{BE} + \alpha_{BP}) X_B^0 - \big(\theta_E \hat{p}_{BE}^0 + \theta_P p_{BP}^0\big)\big] \end{aligned} \tag{29}$$

$$\begin{aligned} \frac{\partial \Pi_E}{\partial t^1} &= i_E \big\{\bar{\eta} X_A^1 - \big(\hat{p}_{BE}^1 - c_E\big)\big[\alpha_{BE} X_B^1 - \theta_E \hat{p}_{BE}^1 - \gamma\big(\hat{p}_{BE}^1 - \hat{p}_{BP}^1\big)\big]\big\} \\ &\quad + (i_E)^2 \sigma_A \rho_A \big[- \big(\alpha_{BMP} X_B^1 - \delta p_{BMP}^1\big) + (\alpha_{BE} + \alpha_{BP}) X_B^1 \\ &\quad - \big(\theta_E \hat{p}_{BE}^1 + \theta_P p_{BP}^1\big)\big]. \end{aligned} \tag{30}$$

The right-hand side of (29) shows that the derivative $\partial \Pi_E / \partial t^0$ depends on the time t^1 at which the product is moved from A to B in period 1. The larger the t^1, the larger the value of $\partial \Pi_E / \partial t^0$. This means that if the product is moved late (or never) to market B in period 1, the likelihood increases that the product is moved late (or never) to market B in period 0.

The signs of the derivatives in (29) and (30) provide four cases:

	$\frac{\partial \Pi_E}{\partial t^1} > 0$	$\frac{\partial \Pi_E}{\partial t^1} < 0$
$\frac{\partial \Pi_E}{\partial t^0} > 0$	(a): $t^0 = 1,\ t^1 = 2$	(b): $t^0 = 1,\ t^1 = 1 + \kappa_E$
$\frac{\partial \Pi_E}{\partial t^0} < 0$	(c): $t^0 = \kappa_E,\ t^1 = 2$	(d): $t^0 = \kappa_E,\ t^1 = 1 + \kappa_E$

In Case (a), the fashion product is never introduced in market B. In Case (d), the product is always introduced in market B as early as possible. In Case (b), the product is not introduced in market B in period 0, but is put in market B as early as possible in period 1. In Case (c), the situation is the opposite: the product is put in market B as early as possible in period 0, but is not introduced in market B in period 1.

Next, we study the derivative in (30) as a function of X_B^1. For this purpose we need the following.

Assumption 3. $\theta_E = \theta_P = 0$.

Define
$$\frac{\partial \Pi_E}{\partial t^1} = f\left(X_B^1\right)$$
and recall that equilibrium prices p_{BE}^1, p_{BP}^1, and p_{BMP}^1 are functions of X_B^1. Differentiation with respect to X_B^1 yields
$$f'(X_B^1)\left\{\begin{array}{c}>\\=\\<\end{array}\right\}0 \text{ if } X_B^1\left\{\begin{array}{c}<\\=\\>\end{array}\right\}\Gamma, \tag{31}$$
where
$$\Gamma = \frac{1}{2\alpha_{BE}+\alpha_{BP}}\left[2\gamma(c_E - c_P) + \frac{9\gamma i_E \sigma_A \rho_A(\alpha_{BE}+\alpha_{BP}-0.5\alpha_{BMP})}{2(2\alpha_{BE}+\alpha_{BP})}\right]$$
is a constant which is positive under the following.

Assumption 4. *(i)* $c_E > c_P$, *(ii)* $\alpha_{BE}+\alpha_{BP} \geq 0.5\alpha_{BMP}$.

Item (i) of the assumption means that the fashion firm has the larger unit production cost. This is very likely. Item (ii) states that the joint impact of brand image X_B on the duopoly demands is not smaller than half the impact of the image on the imitator's monopoly demand in market B. This is also very likely.

Function $f(X_B^1)$ is strictly concave and it holds that
$$f(0) = i_E\left[\bar{\eta}X_A^1 - \frac{\gamma(c_P - c_E)^2}{9}\right] + (i_E)^2\sigma_A\rho_A\frac{\delta}{2}c_P. \tag{32}$$

The sign of $f(0)$ is not unique and there are three possibilities:

(1) $f(0) > 0$, $f(\Gamma) > 0 \Longrightarrow \exists\alpha_1 > \Gamma$ such that $f(\alpha_1) = 0$. We have $f(X_B^1) > 0$ for $X_B^1 \in [0, \alpha_1)$, implying $\hat{t}^1 = 2$, and $f(X_B^1) < 0$ for $X_B^1 > \alpha_1$, implying $\hat{t}^1 = 1+\kappa_E$. Qualitatively speaking, for X_B^1 sufficiently large, it pays to move the fashion product as early as possible to market B. The intuition is clear: The brand image X_B^1 has a large value which the firm exploits by moving the product to market B as soon as possible.

(2) $f(0) < 0$, $f(\Gamma) > 0 \Longrightarrow \exists\alpha_{21} < \Gamma$ and $\alpha_{22} > \Gamma$ such that $f(\alpha_{21}) = f(\alpha_{22}) = 0$. We have $f(X_B^1) < 0$ for $X_B^1 \in [0, \alpha_{21})$, implying $\hat{t}^1 = 1+\kappa_E$, $f(X_B^1) > 0$ for $X_B^1 \in (\alpha_{21}, \alpha_{22})$, implying $\hat{t}^1 = 2$, and $f(X_B^1) < 0$ for $X_B^1 > \alpha_{22}$, implying $\hat{t}^1 = 1+\kappa_E$. Here it pays to move the product to market B if the brand image X_B^1 has an intermediate value; if the value is small or large, the product should remain on market A.

(3) $f(0) < 0$, $f(\Gamma) < 0 \Longrightarrow f(X_B^1) < 0$ for all $X_B^1 \geq 0$, implying $\hat{t}^1 = 1+\kappa_E$ always. It pays to move the product to market B as early as possible, irrespective of the value of the brand image.

Consider (32). A key determinant of the value of $f(0)$ is the brand image X_A^1. For a sufficiently large value of X_A^1 we are in Case (1). Then, if the brand image in market B is "small" (i.e., below the threshold α_1), the product will not be moved to this market because it is more profitable to exploit the brand image in market A. Note that Case (1) also emerges if $c_P \approx c_E$.

3.4 Comparing Pricing Decisions Across Firms

The only comparison to be made is between the duopoly prices in market B. We invoke Assumptions 1 and 4. Using (17) and (25) then yields

$$\hat{p}_{BE}^i > \hat{p}_{BP}^i \quad \text{for } i \in \{0, 1\},$$

that is, the fashion original should be higher priced than the imitation. The result seems to confirm what can be observed in real life. What drives the result are Assumption 3 and item (ii) of Assumption 4: When the brand image has a "high" impact on the demand for the fashion product, and the latter is not affected by the product's absolute price, the product can be higher priced.

4 Conclusions

This chapter has identified equilibrium strategies for price, advertising, and entry decisions in a two-period game. A fashion company introduces new designs regularly and competes in one market, B, with a myopic imitator who copies the design of the fashion company. The fashion firm has a monopoly market, A, and the imitator may have a monopoly in market B throughout a period if the fashion firm chooses not to move its product from market A to B. Demand in markets A and B depends on the respective brand images. Among our findings are the following.

- Equilibrium prices in the duopoly period in market B are proportional to the brand image in that market: the stronger the image, the higher the prices.
- The time of entry of the fashion firm into market B depends on the strength of its image in markets A and B.
- Image advertising efforts of the fashion firm in the first period depends on the time at which it moves its product from market A to B in period 1. Efforts depends positively on the brand image in the first period.
- Equilibrium prices in the duopoly market decrease over time.
- Image advertising efforts in market A are decreasing over time.
- The fashion product is higher priced in market B than the imitation.

A possible avenue for future research is to include product quality. Here one could assume that the quality of the fashion product is fixed, but the imitator can decide which quality to manufacture. The better the quality, the longer it will take before copy can be introduced in market B. On the other hand, a better quality stimulates the demand for the imitation. This problem is currently investigated by the authors. Due to the increased complexity of the model, results are derived by numerical simulations.

References

[1] Caulkins, J., R.F. Hartl, P.M. Kort and G. Feichtinger, Explaining Fashion Cycles: Chasing Innovators by Imitators. Working Paper. Carnegie Mellon University/University of Vienna/Tilburg University/Vienna University of Technology, 2003.

[2] Fudenberg, D. and J. Tirole, *Game Theory*. The MIT Press, Cambridge, MA, 1992.

[3] Jørgensen, S., S. P. Sigue and G. Zaccour, Stackelberg Leadership in a Marketing Channel, *International Game Theory Review* 3, 2001, 13–26.

[4] Leibenstein, H., Bandwagon, Snob and Veblen Effects in the Theory of Consumers' Demand, *Quarterly Journal of Economics* 64, 1950, 183–207.

[5] Magdo, C., Protecting Works of Fashion from Design Piracy, Working Paper, LEDA, Harvard Business School, 2000.

[6] Rags and Riches, *The Economist*, March 6, 2004, 3–16.

Formulating and Solving Service Network Pricing and Resource Allocation Games as Differential Variational Inequalities

T. L. Friesz
Department of Industrial and Manufacturing Engineering
The Pennsylvania State University
University Park, PA 16802, USA
tfriesz@psu.edu

R. Mookherjee
Department of Industrial and Manufacturing Engineering
The Pennsylvania State University
University Park, PA 16802, USA
reeto@psu.edu

M. A. Rigdon
Department of Industrial and Manufacturing Engineering
The Pennsylvania State University
University Park, PA 16802, USA
marigdon@psu.edu

Abstract

In this chapter we show how certain noncooperative Cournot–Nash dynamic games arising in service pricing and resource allocation may be articulated as differential variational inequalities. We show how such problems may be solved using a fixed point algorithm having continuous time optimal control subproblems.

Key words. Differential variational inequalities, dynamic games, service pricing, and resource allocation.

1 Introduction

Noncooperative nonlinear game-theoretic models have been successfully employed to study economic competition in the marketplace, highway and transit traffic in the presence of congestion, regional and global wars, and both intra- and interspecies biological competition. One of the key developments that has made such diverse applications practical is our ability to compute static game-theoretic equilibria as the limit of a sequence of well-defined mathematical programs or complementarity problems.

In many applications, a static perspective is not appropriate. In particular, intermediate disequilibrium states of mathematical games can be intrinsically important. When this is the case, disequilibrium adjustment mechanisms must be articulated, thereby forcing explicit consideration of time. The archetypal example of a disequilibrium adjustment mechanism is the tatonnement story told to explain how equilibrium prices assuring market clearing arise in microeconomic theory. In still other applications, the game-theoretic system of interest displays a *moving equilibrium*, wherein equilibrium is enforced at each instant of time, although state and control variables will generally fluctuate with time. These fluctuations with respect to time are exactly those needed to maintain the balance of behavioral and economic circumstances defining the equilibrium of interest. An example of a moving equilibrium is the dynamic user equilibrium problem studied in the field of dynamic traffic assignment.

In this chapter we study a class of service pricing and resource application problems that are dynamic Cournot–Nash games expressible as infinite-dimensional variational inequalities with clearly distinguished state and control variables. The state variables obey state dynamics expressed as ordinary differential equations and the control variables obey either pure control constraints or mixed state-control constraints. Each game agent expresses its strategy through subsets of the control variables that may or may not be shared with other agents, depending on whether collusion occurs. We refer to variational inequalities with such structures as *differential variational inequalities*, just as some dynamic games are referred to as differential games. We advocate this terminology, despite the fact that Aubin and Cellina [1] use the name *differential variational inequalities* for variational inequalities whose principal operators depend on derivatives of decision variables, since our usage seems very natural, and we have discovered no literature derived from Aubin and Cellina that adopts the Aubin usage.

2 Differential Variational Inequalities Defined

We follow Friesz et al. [2] in referring to the following differential variational inequality problem as $DVI(F, f, U, x^0)$:

$$\text{find } u^* \in U \text{ such that}$$

$$\langle F(x(u^*), u^*, t), u - u^* \rangle \geq 0 \quad \text{for all } u \in U, \tag{1}$$

where

$$u \in U \subseteq (L^2[t_0, t_f])^m$$

$$x(u,t) = \arg\left\{ \frac{dx}{dt} = f(x,u,t), x(0) = x^0, \Gamma[x(L), L] = 0 \right\} \in (\mathcal{H}^1[t_0, t_f])^n \tag{2}$$

$$x^0 \in \Re^n \tag{3}$$

$$F : (\mathcal{H}^1[t_0, t_f])^n \times (L^2[t_0, t_f])^m \times \Re^1_+ \longrightarrow (L^2[t_0, t_f])^m \tag{4}$$

$$f : (\mathcal{H}^1[t_0, t_f])^n \times (L^2[t_0, t_f])^m \times \Re^1_+ \longrightarrow (L^2[t_0, t_f])^n \tag{5}$$

$$\Gamma : (\mathcal{H}^1[t_0, t_f])^n \times \Re^1_+ \longrightarrow (\mathcal{H}^1[t_0, t_f])^r \tag{6}$$

Note that $(L^2[t_0, t_f])^m$ is the m-fold product of the space of square-integrable functions $L^2[t_0, t_f]$, while $(\mathcal{H}^1[t_0, t_f])^n$ is the n-fold product of the Sobolev space $\mathcal{H}^1[t_0, t_f]$.[1] We would like to point out here that Friesz and Mookherjee [3] have generalized (1) involving the non-differentiability of the principal operator $F(.,.,.)$ as well as differential-difference equations arising from state-dependent time shifts of control variables. Such time-shifted DVIs are known to arise as a result of the forward-looking or anticipatory perspective of individual, extremizing game-theoretic agents. Note that

$$\langle F(x(u^*), u^*, t), u - u^* \rangle \equiv \int_{t_0}^{t_f} [F(x(u^*), u^*, t)]^T (u - u^*) dt \geq 0$$

defines the inner product in (1).

Naturally there is a fixed point formulation of (1), namely

$$u = P_U[u - \alpha F(x(u), u, t)] \tag{7}$$

where $P_U[.]$ is the minimum norm projection operator relative to the set U. The following fixed point algorithm is immediate from (7):

The Fixed Point Algorithm

Step 0. Initialization. Identify an initial feasible solution $u^0 \in U$ and set $k = 0$.
Step 1. Optimal control subproblem. Solve

$$\min_v J^k(v) = \gamma^T \Gamma[x(t_f), t_f] + \int_{t_0}^{t_f} \frac{1}{2}[u^k - \alpha F(x^k, u^k, t) - v]^2 dt \tag{8}$$

$$\text{subject to } \frac{dx}{dt} = f(x, v, t) \tag{9}$$

$$v \in U \tag{10}$$

$$x(0) = x^0 \tag{11}$$

[1] Vector space $\mathcal{H}^1\left[t_0, t_f\right]$ is a Sobolev space if

$$\mathcal{H}^1[t_0, t_f] = \left\{ v \,\middle|\, v \in (L^2[t_0, t_f])^m;\ \frac{\partial v}{\partial x_i} \in L^2[t_0, t_f] \text{for all } i = 1, \ldots, m \right\}.$$

and call the solution u^{k+1}. Note that in this step it is advantageous to "unfold" and explicitly state the constraints embedded in the operator $x(u, t)$ of statement (2) in order to facilitate computation.

Step 2. Stopping test. If

$$\left\| u^{k+1} - u^k \right\| \leq \varepsilon,$$

where $\varepsilon \in \Re^1_{++}$ is a preset tolerance, stop and declare $u^* \approx u^{k+1}$. Otherwise set $k = k + 1$ and *go to Step 1*.

Convergence of this algorithm can be guaranteed in certain circumstances that are not explored here. Please refer to Friesz and Mookherjee [12] for such details.

3 Some Service Enterprise and Engineering Applications

There are numerous service enterprise and engineering areas where DVIs are an obvious choice to model the behavior of Cournot–Nash agents in a dynamic setting. Some service enterprise and engineering applications that have been placed in the form of variational inequalities include, but are not limited to, revenue management, competition in dynamic supply chains, spatial oligopolistic network competition, electric power networks, and predictive models of data flows over the Internet. Many, if not all, of these applications may be modeled to involve time shifts which can also be accommodated in a DVI formulation.

In particular, this chapter examines a revenue management problem involving oligopolistic competition in joint pricing and resource allocation under demand uncertainty. Two different versions of the model are inspected, one involving bidirectional pricing and the other involving markdown pricing. In order to illustrate the application of these approaches, a detailed numerical example is provided.

4 Service Pricing and Resource Allocation

Service pricing and resource allocation are recognized subdisciplines of revenue management (RM) was pioneered by the airline industry in the 1970s after the market deregulation. Since then, many other service industries have successfully adapted the art of RM to increase profitability. McGill and van Ryzin [4] provide a detailed survey of the research advancements in this field since 1970. Although service pricing and resource allocation are widely argued to be both dynamic and game theoretic in the real world, relatively few mathematical models with these properties have been developed. The DVI perspective offers a way to rapidly create a model that is both dynamic and game theoretic, as illustrated by the model presented below.

4.1 Overview of the Model

We study a joint pricing and resource allocation problem in a network with applications to production planning and airline RM. This model is similar to that of

Perakis and Sood [5], though we consider multiple products which leads to the network structure of the problem. We also approach the problem directly, whereas Perakis and Sood address the competitive aspect of the problem by using ideas from robust optimization.

Suppose a set of service providers has a finite supply of resources which it can use to produce multiple products with different resource requirements. The firms do not have the ability to restock their resources during the planning horizon and any unsold resource does not have any salvage value at the end of the planning horizon. Further, if realized demand for a service type at any time is more than the rate of provision of service, excess demand is lost. All service providers can set the price for each of their services which are bounded from above and below. The demand for each product is uncertain and depends on own service prices as well as non-own service prices. Each service provider has to decide how to allocate its resources for providing different services and how to price its services to maximize its expected revenue. This research is motivated by the airline pricing problem where the services are combinations of origin, destination, and fare class and the resources are seats on flights.

4.2 Notation

Before we proceed with the details of the model, we overview the notation we will be using in rest of the chapter.

Parameters

$\mathcal{F}$	set of firms
$\mathcal{S}$	set of services each firm provides
$\mathcal{C}$	set of resources that firms use to provide services
$\mathcal{C}_i$	set of resources that firms use to provide service $i \in \mathcal{S}$
$\mathcal{A}$	resource-service incidence matrix
t_0	starting time of the booking period (planning horizon)
t_1	end time of the booking period (planning horizon)
$t \in [t_0, t_1]$	instant of time
$p^f_{i,\min}$	minimum price that firm f can charge for service $i \in \mathcal{S}$
$p^f_{i,\max}$	maximum price that firm f can charge for service $i \in \mathcal{S}$
$u^f_{i,\min}$	minimum rate of provision of service $i \in \mathcal{S}$ offered by firm f
$u^f_{i,\max}$	maximum rate of provision of service $i \in \mathcal{S}$ offered by firm f
K^f_j	firm f's capacity for resource type $j \in \mathcal{C}$.

From this notation it is evident that

$$\cup_{i=1}^{|\mathcal{S}|} \mathcal{C}_i = \mathcal{C}.$$

The variables we use are primarily the controls (prices and allocation of resources) and the states (cumulative allocated).

Variables

$p_i^f(t)$	price for the service $i \in \mathcal{S}$ charged by firm $f \in \mathcal{F}$ at time t
$u_i^f(t)$	firm f's service level of type $i \in \mathcal{S}$ at time t
$x_j^f(t)$	firm f's total allocated resource of type $j \in \mathcal{C}$ until time t
$d_i^f(p, t)$	mean demand for service $i \in \mathcal{S}$ from firm $f \in \mathcal{F}$ at time t when prevailing price is p
$D_i^f(p, t)$	firm f's realized demand of service $i \in \mathcal{S}$ at time t
$z_i^f(t)$	random component associated with demand faced by the firm f for service i at time t.

In vector notation, decision variables for firm f are

$$p^f = \left(p_i^f : i \in \mathcal{S}\right)$$
$$u^f = \left(u_i^f : i \in \mathcal{S}\right).$$

Decision variables of firm f's competitors are

$$p^{-f} = (p^{-g} : g \in \mathcal{F}\backslash f)$$
$$u^{-f} = (u^{-g} : g \in \mathcal{F}\backslash f).$$

State variables for firm f are the cumulative allocation of resources

$$x^f = \left(x_j^f : j \in \mathcal{C}\right),$$

which can be concatenated as

$$x = (x^f : f \in \mathcal{F}).$$

4.2.1 Resource-Service Incidence Matrix

The network we are interested in has $|\mathcal{C}|$ resources and the firm provides $|\mathcal{S}|$ different services. Each network product is a bundle of the $|\mathcal{C}|$ resources sold with certain purchase terms and restrictions at a given price. The resource-service *incidence matrix*, $\mathcal{A} = [a_{ij}]$ is a $|\mathcal{C}| \times |\mathcal{S}|$ matrix where

$$\begin{aligned} a_{ij} &= 1 \text{ if resource } i \text{ is used by service } j \\ &= 0 \text{ else.} \end{aligned}$$

Thus the jth column of $\mathcal{A}$, denoted $\mathcal{A}_j$, is the *incidence vector* for the service j; the ith row, denoted $\mathcal{A}^i$, has an entry of one in the column j corresponding to a service j that utilizes the resource i. Note that there will be multiple identical columns if there are multiple ways of selling a given bundle of resources, probably with different restrictions. Each would have an identical column in the matrix $\mathcal{A}$, but they could have different revenue values and different demand patterns (Talluri and van Ryzin [6]).

4.2.2 Demand for Services

Demand for each service is uncertain. Its expectation only depends on the current market price of services. Firm f's realized demand at time t for service i is $D_i^f(p,t)$ when the prevailing market price is p. Demands in consecutive periods are independent and nonnegative. Two types of demand models are predominant in the supply chain and news vendor pricing literature:

(1) additive form:
$$D_i^f(p,t) = d_i^f(p,t) + z_i^f$$
(2) multiplicative form:
$$D_i^f(p,t) = d_i^f(p,t) \cdot z_i^f,$$

where $z_i^f \geq 0$ is a continuous random variable with known distribution and $d_i^f(p,t)$ is the expected or average demand faced by the firm f for its service i when the price combination is p at time t. Nevertheless, to keep the exposition simple, we focus on the multiplicative demand form in the rest of the paper. The probability density function and the cumulative probability distribution function of the random variable z_i^f are denoted by $g(z_i^f)$ and $G(z_i^f)$ respectively. In the RM literature it is common to make two major assumptions regarding the nature of the random variable z_i^f and the average demand d_i^f.

Assumption A.1. *The random variable z_i^f is independent of price, p and time, t, and identically distributed with $E(z_i^f) = 1$, $Var\,(z_i^f) < \infty$.*

Assumption A.2. *For any firm $f \in F$ and the service type $i \in S$, the mean demand $d_i^f(p,t)$ has the following properties:*

(1) d_i^f depends only on the prices charged by firm f and its competitors for service type i only
(2) $d_i^f(p,t)$ is continuous, bounded, and differentiable in p_i on the strategy space $[p_{i,\min}, p_{i,\max}]$ where
$$p_i = \left\{p_i^g : g \in \mathcal{F}\right\}$$
(3) average demand decreases with own service price, i.e., $\frac{\partial d_i^f(p,t)}{\partial p_i^f} < 0$ and
(4) average demand increases with non-own service prices, i.e., $\frac{\partial d_i^f(p,t)}{\partial p_i^g} > 0$ for all $g \neq f$

Note that
$$E\big(D_i^f(p,t)\big) = d_i^f(p,t).$$
The multiplicative model implies that the coefficients of variation of the one-period demands are constants, independent of the price vectors. This type of multiplicative model was originally proposed by Karlin and Carr [7] and has since been used

frequently in the supply chain literature. However, it is also possible to consider a more general demand model in which the standard deviation of demand may fail to be proportional to the mean, i.e., the coefficient of variation may vary as service prices and mean demand volumes vary. One such demand model is

$$D_i^f(p,t) = d_i^f(p,t) + \sigma_i^f(d_i^f(p,t)) \cdot z_i^f,$$

where $\sigma_i^f(\cdot)$ is a general increasing function with $\sigma_i^f(0) = 0$ and $z_i^f \geq 0$ is a continuous random variable (see Bernstein and Federgruen [8] for such nonmultiplicative demand models). It can be easily verified that most of the commonly used demand functions satisfy the requirements in Assumption A2; in particular, the

(1) *linear function*

$$d_i^f(p,t) = \rho_i^f(t) - \sigma_i^f(t) \cdot p_i^f + \sum_{g \in \mathcal{F}\setminus f} \gamma_i^g(t) \cdot p_i^g,$$

where $\rho_i^f(t), \sigma_i^f(t), \gamma_i^f(t) \in \mathbb{R}^1_{++}$ for all $f \in \mathcal{F}$ and $i \in \mathcal{S}$,

(1) *logit*

$$d_i^f(p,t) = \frac{a_i^f(t)\exp(-b_i^f(t) \cdot p_i^f)}{\sum_{g\in\mathcal{F}} a_j^f(t)\exp(-b_j^f(t) \cdot p_j^f)},$$

where $a_i^f(t), b_i^f(t) \in \mathbb{R}^1_{++}$ for all $f \in \mathcal{F}$ and $i \in \mathcal{S}$,

(2) *Cobb–Douglas*

$$d_i^f(p,t) = a_i^f(p_i^f(t))^{-\beta_i^f} \prod_{g \in \mathcal{F}\setminus i} (p_i^g(t))^{\beta_i^{fg}},$$

where $a_i^f > 0$, $\beta_i^f > 1$, $\beta_i^{fg} > 0$ for all $f \in \mathcal{F}$, $i \in \mathcal{S}$, and $t_0 \leq t \leq t_1$.

In what follows we now describe the revenue optimization problem for the service providers and articulate the Cournot–Nash equilibrium of the dynamic pricing and resource allocation game as a DVI.

4.3 Service Providers' Optimal Control Problem

The time scale (booking period) we consider here is short enough so that the time value of money does not need to be considered. With competitors' service prices

$$p^{-f} \equiv \{p^g : g \in \mathcal{F}\setminus f\}$$

taken as exogenous to the firm $f \in \mathcal{F}$'s optimal control problem and yet endogenous to the overall equilibrium problem, firm f computes its prices p^f and allocation of resources u^f in order to maximize revenue generated throughout the booking period

$$\max_{p^f, u^f} J(p^f, u^f; p^{-f}) = E\left[\int_{t_0}^{t_1} p^f \cdot \min(u^f, D^f(p,t))dt\right] \tag{12}$$

subject to

$$\frac{dx^f}{dt} = \mathcal{A} \cdot u^f \tag{13}$$

$$x^f(t_0) = 0 \tag{14}$$

$$x^f(t_1) \leq K^f \tag{15}$$

$$p^f_{\min} \leq p^f \leq p^f_{\max} \tag{16}$$

$$u^f_{\min} \leq u^f \leq u^f_{\max}. \tag{17}$$

As in a typical RM industry, there is no salvage value of unsold inventory at the end of the planning horizon; all unfulfilled demand is lost. At the same time, firms are restricted from accepting backorders. Our model does not make decisions regarding overbooking, which is a standard practice done in anticipation of cancellations and no-shows. Mookherjee and Friesz [12] show how this model can be adapted to a model of overbooking. The definitional dynamics (13) describe total commitment of resources derived from the rate at which services are provided. The initial condition (14) articulates that at the start of the booking period no resources are committed. In the absence of overbooking, the total committed resource at the end of the booking period cannot exceed the actual available resource as expressed in constraint (15). Constraints (16) are simply the upper and lower bounds on the prices that the firm is allowed to charge due to some regulation. The lower bounds ($u^f_{\min} > 0$) on the allocation variables u^f in (17) ensure that each service provider participates in each period with a strictly positive provision of service. The implication, if this were not true, would be that a firm with nothing to sell in a period could influence the demand seen by its competitors by setting a price. In other words, setting a price would make sense only if there is a nonzero sale in that period. Note that we have not imposed any constraint on the direction of price change. Clearly, as firms' optimal control problems are coupled, this gives rise to a Corunot–Nash dynamic game setting.

4.4 Analysis of the Optimal Control Problem and DVI Formulation of the Game

We need to study firm f's best response optimal control problem given its competitors' service prices p^{-f}. The expected instantaneous revenue function for firm f at time t is

$$\begin{aligned} R_f(p^f, u^f; p^{-f}) &= E[p^f(t) \cdot \min(u^f(t), D^f(p,t))] \\ &= E[p^f(t) \cdot \min(u^f(t), d^f(p,t) \cdot z^f)] \\ &= \sum_{i \in \mathcal{S}} \left(p^f_i(t) \cdot u^f_i(t) - p^f_i(t) \cdot d^f(p,t) \int_0^{\frac{u^f_i(t)}{d^f(p,t)}} G(\tau) d\tau \right). \end{aligned} \tag{18}$$

The component $\sum_{i \in \mathcal{S}} p_i^f \cdot u_i^f$ in the instantaneous revenue function (18) is called the *riskless component* of the revenue for firm $f \in \mathcal{F}$. Further, depending on the component $\frac{u_i^f}{d^f(p,t)}$, the expected revenue is reduced as the term $\frac{u_i^f}{d^f(p,t)}$ increases. There is a deterministic optimal control problem that is equivalent to the stochastic optimal control problem of Section 4.3 because of the following properties: (a) the dynamics are deterministic and (b) a closed-form expression of the objective expectation can be found for multiplicative demand as shown above. Therefore, the equivalent deterministic Hamiltonian associated with firm f's problem is

$$\begin{aligned} &H_f(p^f, u^f; p^{-f}; \lambda^f; t) \\ &\quad = \sum_{i \in \mathcal{S}} \left(p_i^f(t) \cdot u_i^f(t) - p_i^f(t) \cdot d^f(p,t) \int_0^{\frac{u_i^f(t)}{d^f(p,t)}} G(\tau) d\tau \right) \\ &\qquad + (\lambda^f(t))^T \cdot (\mathcal{A} \cdot u^f(t)), \end{aligned}$$

where λ^f is the vector of adjoint variables which has the interpretation of *shadow price of resources* in the current context. Let us define

$$\lambda^f = (\lambda_j^f : j \in \mathcal{C}).$$

Therefore, the Hamiltonian can be rewritten as

$$\begin{aligned} H_f(p^f, u^f; p^{-f}; \lambda^f) &= \sum_{i \in \mathcal{S}} \left(p_i^f(t) \cdot u_i^f(t) - p_i^f(t) \cdot d_i^f(p,t) \int_0^{\frac{u_i^f(t)}{d_i^f(p,t)}} G(\tau) d\tau \right) \\ &\quad + (\lambda^f(t))^T \cdot (\mathcal{A} \cdot u^f(t)) \\ &= \sum_{i \in \mathcal{S}} [p_i^f(t) + (\lambda^f(t))^T \mathcal{A}_i] \cdot u_i^f(t) \\ &\quad - \sum_{i \in \mathcal{S}} p_i^f(t) \cdot d_i^f(p,t) \int_0^{\frac{u_i^f(t)}{d_i^f(p,t)}} G(\tau) d\tau. \end{aligned} \tag{19}$$

Now, assume

$$c_i^f(t) \equiv -(\lambda^f(t))^T \mathcal{A}_i$$

to be firm f's "shadow price" of providing per unit service type $i \in \mathcal{S}$, which is endogenously determined, at each time $t \in [t_0, t_1]$. Thus, (19) can be rewritten as

$$\begin{aligned} H_f(p^f, u^f; p^{-f}; \lambda^f) &= \sum_{i \in \mathcal{S}} \{p_i^f(t) - c_i^f(t)\} \cdot u_i^f(t) \\ &\quad - \sum_{i \in \mathcal{S}} p_i^f(t) \cdot d_i^f(p,t) \int_0^{\frac{u_i^f(t)}{d_i^f(p,t)}} G(\tau) d\tau. \end{aligned} \tag{20}$$

Under this condition this becomes a *multi-product newsvendor pricing problem* for firm f with endogenously determined prices.

The adjoint dynamics associated with the optimal control problem of firm f is

$$\frac{d\lambda_j^f}{dt} = -\frac{\partial H_f}{\partial x_j^f} = 0.$$

This is so because Hamiltonian (19) is free from the state, hence, λ_j^f remains stationary throughout the trajectory

$$\lambda_j^f(t) = \lambda_j^f \quad \forall t \in [t_0, t_1], \; j \in \mathcal{C}.$$

From the terminal time state condition $x^f(t_1) \leq K^f$ we can form the following complementarity conditions (see Sethi and Thompson [9]):

$$\lambda^f \cdot [x^f(t_1) - K^f] = 0 \tag{21}$$

and

$$\lambda^f \leq 0 \tag{22}$$

$$x^f(t_1) - K^f \leq 0. \tag{23}$$

Pontryagin's maximum principle [10] tells us that an optimal solution to (12)–(17) is a quadruplet

$$\{p^{f*}(t), u^{f*}(t); x^{f*}(t); \lambda^{f*}(t)\}$$

that, given $H_f(p^f, u^f; p^{-f}; \lambda^f)$, must satisfy at each time $t \in [t_0, t_1]$

$$\begin{pmatrix} p^{f*}(t) \\ u^{f*}(t) \end{pmatrix} = \arg\left\{ \max_{(p^f(t), u^f(t)) \in \mathcal{K}_f} H_f(p^f, u^f; p^{-f}; \lambda^f) \right\}, \tag{24}$$

where

$$\mathcal{K}_f = \{(p^f, u^f) : \text{ (13)–(17) hold}\}.$$

$x^f(u^f, t)$ is the solution of

$$x^f(u^f, t) = \arg\left\{ \frac{dx^f}{dt} = \mathcal{A} \cdot u^f, \; x^f(t_0) = 0, \; x^f(t_1) \leq K^f \right\}. \tag{25}$$

By regularity (in particular, U_f is convex and compact), the necessary condition for (24) can be expressed in the following variational form (see Minoux's Theorem 10.6 [11]):

$$\left\{ \begin{array}{l} [\nabla_{p^f} H_f(p^{f*}, u^{f*}; p^{-f*}; \lambda^{f*})]^T (p^f - p^{f*}) \\ +[\nabla_{u^f} H_f(p^{f*}; u^{f*}; p^{-f*}; \lambda^{f*})]^T (u^f - u^{f*}) \end{array} \right\} \leq 0 \quad \forall \begin{pmatrix} p^f \\ u^f, \end{pmatrix} \in \mathcal{K}_f, \tag{26}$$

where

$$H_f^* = H_f(p^{f*}, u^{f*}; p^{-f*}; \lambda^{f*}).$$

Now consider the following DVI which can be shown to have solutions that are Cournot–Nash equilibria for the joint network pricing and resource allocation game described above in which individual service providers maximize expected revenue in light of current information about competitors' prices:

$$\text{find } \begin{pmatrix} p^* \\ u^* \end{pmatrix} \in \mathcal{K} \text{ such that}$$

$$\sum_{f \in \mathcal{F}} \int_{t_0}^{t_1} (\nabla_{p^f} H_f^* \cdot (p^f - p^{f*}) + \nabla_{u^f} H_f^* \cdot (u^f - u^{f*}))dt \leq 0 \tag{27}$$

$$\text{for all} \begin{pmatrix} p \\ u \end{pmatrix} \in \mathcal{K} = \prod_{f \in \mathcal{F}} \mathcal{K}_f,$$

where

$$\begin{pmatrix} p \\ u \end{pmatrix} = \begin{pmatrix} [p^f]_{f \in \mathcal{F}} \\ [u^f]_{f \in \mathcal{F}} \end{pmatrix}.$$

We omit the formal proof of this equivalence relationship between solution of the above DVI and Cournot–Nash equilibrium of the game for the sake of brevity. Curious readers are encouraged to refer to Mookherjee and Friesz [12].

4.5 Cooperative Equilibrium

We are also interested in comparing pricing and allocation strategies of the firms in the event of perfect collusion where all firms collaborate to maximize their aggregate expected revenue throughout the booking period. The single optimal control problem that the firms collectively seek to solve is the following:

$$\max_{p,u} J_C(p, u) = \sum_{f \in \mathcal{F}} E\left[\int_{t_0}^{t_1} p^f \cdot \min(u^f, D^f(p, t))dt\right] \tag{28}$$

subject to

$$\frac{dx^f}{dt} = \mathcal{A} \cdot u^f \quad \forall f \in \mathcal{F} \tag{29}$$

$$x^f(t_0) = 0 \quad \forall f \in \mathcal{F} \tag{30}$$

$$x^f(t_1) \leq K^f \quad \forall f \in \mathcal{F} \tag{31}$$

$$p_{\min}^f \leq p^f \leq p_{\max}^f \quad \forall f \in \mathcal{F} \tag{32}$$

$$u_{\min}^f \leq u^f \leq u_{\max}^f \quad \forall f \in \mathcal{F}. \tag{33}$$

The Hamiltonian of the above optimal control problem is

$$H_1(p,u,\lambda) = \sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{S}}[p_i^f + (\lambda^f)^T\mathcal{A}_i]\cdot u_i^f$$
$$- \sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{S}} p_i^f \cdot d_i^f(p,t)\int_0^{\frac{u_i^f}{d_i^f(p,t)}} G(\tau)d\tau.$$

Let $(\tilde{p},\tilde{u})$ be the cooperative service prices and resource allocations respectively for the firms where

$$\tilde{p} = (\tilde{p}^f)_{f\in\mathcal{F}}$$
$$\tilde{u} = (\tilde{u}^f)_{f\in\mathcal{F}}.$$

Cooperative service prices and resource allocations can be computed by solving the following DVI:

$$\text{find } \begin{pmatrix}\tilde{p}\\ \tilde{u}\end{pmatrix} \in \mathcal{K} \text{ such that}$$
$$\int_{t_0}^{t_1} (\nabla_p \tilde{H}_1 \cdot (p-\tilde{p}) + \nabla_u \tilde{H}_1 \cdot (u^f - \tilde{u}))dt \le 0 \tag{34}$$
$$\text{for all } \begin{pmatrix}p\\ u\end{pmatrix} \in \mathcal{K},$$

where

$$\tilde{H}_1 = H_1(\tilde{p},\tilde{u};\tilde{\lambda}).$$

Note that $\tilde{\lambda}$ is the cooperative shadow price of the resources owned by the firms.

5 Competitive Model of Markdown Pricing and Resource Allocation

So far we have considered bidirectional price changes where the service prices charged by firms can be changed arbitrarily from period to period. In what follows we consider the setting where only markdowns are permitted. Let us define $r_i^f(t)$ as the rate of change of price for the service $i \in \mathcal{S}$ charged by firm $f \in \mathcal{F}$ at time $t \in [t_0, t_1]$ which is the control in the hands of the service provider. Service prices, $p_i^f(t)$, become the states. To keep the exposition simple, we consider the service prices at time t_0 exogenous, i.e.,

$$p(t_0) = \breve{p} \in \mathbb{R}_{++}^{|\mathcal{F}|\times|\mathcal{S}|}.$$

However, this initial condition may also be set endogenously.

5.1 Price Dynamics

The relationship between the service price and the rate of change of the price can be articulated through the following definitional relationship:

$$\frac{dp_i^f(t)}{dt} = r_i^f(t) \tag{35}$$

$$p_i^f(t_0) = \breve{p}_i^f. \tag{36}$$

Since prices can only go down or stay flat throughout the booking period, for the company f,

$$r_{\min}^f \leq r^f \leq 0, \tag{37}$$

where

$$r^f = (r_i^f : i \in \mathcal{S})$$

and $r_{\min}^f < 0$ is the maximum allowable markdown between two consecutive periods. In the event of markup pricing, the bounds are changed to

$$0 \leq r^f \leq r_{\max}^f,$$

where $r_{\min}^f < 0$. We are now in a position to state the coupled optimal control problem faced by each service provider to set its service prices and allocation of resources in the event of markdown pricing.

5.2 Service Providers' Optimal Control Problem Considering Markdown

In the markdown model, each firm must be able to completely and perfectly observe the rate of change of service prices as set by its competitors. We argue that this is no more difficult than observing the service price set by its competitors as we postulate that with the advent of information technology it is becoming increasingly easier for the companies to obtain pricing information about their competitors through support tools like aggregators and comparators. With this information a firm requires only a bookkeeping effort to monitor its competitors' change of prices and the frequency of such a change by looking at the price trajectories. With the competitors' rate of change of service prices

$$r^{-f} \equiv \{r^g : g \in \mathcal{F} \backslash f\}$$

taken as exogenous to firm f's optimal control problem and yet endogenous to the overall equilibrium problem, firm f computes its markdown decisions r^f (hence the service prices) and allocation of resources u^f in order to maximize revenue generated throughout the booking period,

$$\max_{r^f, u^f} J(r^f, u^f; r^{-f}) = E\left[\int_{t_0}^{t_1} p^f \cdot \min(u^f, D^f(p,t))dt\right] \tag{38}$$

subject to

$$\frac{dx^f}{dt} = \mathcal{A} \cdot u^f \tag{39}$$

$$x^f(t_0) = 0 \tag{40}$$

$$x^f(t_1) \leq K^f \tag{41}$$

$$\frac{dp^f}{dt} = r^f \tag{42}$$

$$p^f(t_0) = \check{p}^f \tag{43}$$

$$r^f_{\min} \leq r^f \leq 0 \tag{44}$$

$$p^f_{\min} \leq p^f \tag{45}$$

$$u^f_{\min} \leq u^f \leq u^f_{\max} \tag{46}$$

Note that this formulation is very similar to that for the bidirectional price change in Section 4.3 except for equations (42)–(45). The definitional dynamics (42) and the initial conditions (43) were explained in Section 5.1. Constraint (44) is the bound on the rate of markdown of service prices. Since the service price cannot increase between two consecutive periods, an upper bound on the service price is unnecessary in this scenario and is hence omitted in (45). Decision variables at hand for firm f are the rate of markdown (r^f) and the rate of provision of services (u^f). The states are the service prices (p^f) and cumulative allocation of resources (x^f). Once again, as firms' optimal control problems are coupled, this gives rise to a Cournot–Nash dynamic games setting.

5.3 DVI Formulation of the Cournot–Nash Game

For brevity we suppress details of the analysis of each service provider's optimal control problem (see Mookherjee and Friesz for a detailed discussion [12]). Instead we state the DVI which has solutions that are Cournot–Nash equilibria for the markdown pricing and resource allocation based network RM game described above in which each individual service provider maximizes expected revenue in light of current information about its competitors' rate of change of prices:

$$\text{find} \begin{pmatrix} r^* \\ u^* \end{pmatrix} \in \tilde{\mathcal{K}} \text{ such that}$$

$$\sum_{f \in \mathcal{F}} \int_{t_0}^{t_1} (\nabla_{r^f} \tilde{H}^*_f \cdot (r^f - r^{f*}) + \nabla_{u^f} \tilde{H}^*_f \cdot (u^f - u^{f*}))dt \leq 0 \tag{47}$$

$$\text{for all} \begin{pmatrix} r \\ u \end{pmatrix} \in \tilde{\mathcal{K}} = \prod_{f \in \mathcal{F}} \tilde{\mathcal{K}}_f,$$

where

$$\begin{pmatrix} r \\ u \end{pmatrix} = \begin{pmatrix} [r^f]_{f\in\mathcal{F}} \\ [u^f]_{f\in\mathcal{F}} \end{pmatrix}$$

$$\tilde{H}_f^* = \tilde{H}_f(r^{f*}, u^{f*}; r^{-f*}; \lambda^{f*})$$

$$\tilde{H}_f(r^f, u^f; r^{-f}; \lambda^f) = \sum_{i\in\mathcal{S}} \{p_i^f + (\lambda^f)^T \mathcal{A}_i\} \cdot u_i^f(t) - \sum_{i\in\mathcal{S}} p_i^f \cdot d_i^f(p,t) \int_0^{\frac{u_i^f(t)}{d_i^f(p,t)}} G(\tau)d\tau + \sum_{i\in\mathcal{S}} r_i^f$$

and

$$\tilde{\mathcal{K}}_f = \{(r^f, u^f) : \text{ (39)–(46) hold}\}.$$

5.4 Cooperative Equilibrium with Markdown Pricing

Similar to the bidirectional pricing case (see Section 4.5), we are also interested in comparing pricing and allocation strategies of the firms in the event of perfect collusion where all firms collaborate to maximize their aggregate expected revenue throughout the booking period. The single optimal control problem that the firms collectively seek to solve is the following:

$$\max_{r,u} J_{\mathcal{C}}(r,u) = \sum_{f\in\mathcal{F}} E\left[\int_{t_0}^{t_1} p^f \cdot \min(u^f, D^f(p,t))dt\right] \tag{48}$$

subject to

$$\frac{dx^f}{dt} = \mathcal{A} \cdot u^f \quad \forall f \in \mathcal{F} \tag{49}$$

$$x^f(t_0) = 0 \forall f \in \mathcal{F} \tag{50}$$

$$x^f(t_1) \le K^f \forall f \in \mathcal{F} \tag{51}$$

$$\frac{dp^f}{dt} = r^f \forall f \in \mathcal{F} \tag{52}$$

$$p^f(t_0) = \breve{p}^f \forall f \in \mathcal{F} \tag{53}$$

$$r_{\min}^f \le r^f \le 0 \quad \forall f \in \mathcal{F} \tag{54}$$

$$p_{\min}^f \le p^f \forall f \in \mathcal{F} \tag{55}$$

$$u_{\min}^f \le u^f \le u_{\max}^f \forall f \in \mathcal{F}. \tag{56}$$

The Hamiltonian of this optimal control problem is

$$H_2(r, u, \lambda) = \sum_{f\in\mathcal{F}} \sum_{i\in\mathcal{S}} [p_i^f + (\lambda^f)^T \mathcal{A}_i] \cdot u_i^f$$

$$- \sum_{f\in\mathcal{F}} \sum_{i\in\mathcal{S}} p_i^f \cdot d_i^f(p,t) \int_0^{\frac{u_i^f}{d_i^f(p,t)}} G(\tau)d\tau + \sum_{f\in\mathcal{F}} \sum_{i\in\mathcal{S}} r_i^f.$$

Let $(\hat{r}, \hat{u})$ be the cooperative markdown of service prices and resource allocations respectively for the firms where

$$\hat{r} = (\hat{r}^f)_{f\in\mathcal{F}}$$
$$\hat{u} = (\hat{u}^f)_{f\in\mathcal{F}}.$$

The cooperative markdown of service prices and resource allocations can be obtained by solving the following DVI:

$$\text{find } \begin{pmatrix} \hat{r} \\ \hat{u} \end{pmatrix} \in \tilde{\mathcal{K}} \text{ such that}$$

$$\int_{t_0}^{t_1} (\nabla_r \hat{H}_2 \cdot (r - \hat{r}) + \nabla_u \hat{H}_2 \cdot (u - \hat{u}))dt \leq 0 \tag{57}$$

$$\text{for all } \begin{pmatrix} r \\ u \end{pmatrix} \in \tilde{\mathcal{K}},$$

where

$$\hat{H}_2 = H_2(\hat{r}, \hat{u}, \hat{\lambda}).$$

6 Numerical Example

6.1 Problem Generation and Parameters

Our numerical example is motivated by the airline RM problem. We have considered a 2-hub, 6-node, and 12-leg airlines network as shown in Figure 1, where nodes 2 and 3 are the "hub" nodes.

Three firms are competing over this network. To keep the exposition simple we assume that each firm uses the same network; however, this can be relaxed with a more general setting where each firm has a different service network. We consider 18 different origin-destination (O-D) pairs of this network and for simplicity we

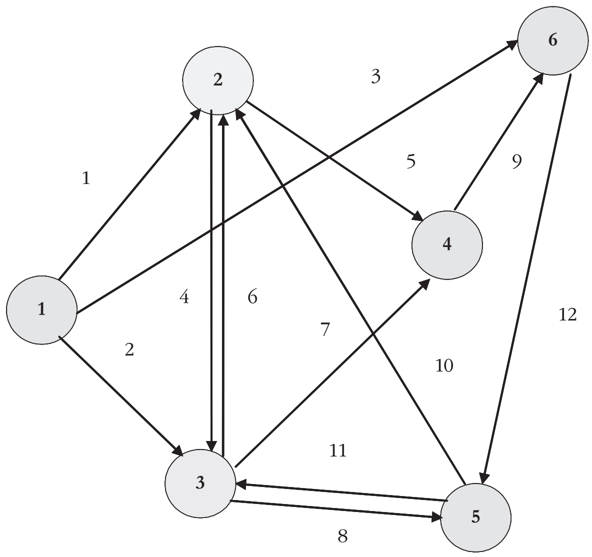


Figure 1: A 6-node, 12-leg airlines network.

assume that each O-D pair is connected by only one path (itinerary) which are the different services that each firm offers, as shown below.

Service ID	O-D pair	Itinerary	Service ID	O-D pair	Itinerary
1	$1-2$	$1-2$	10	$3-2$	$3-2$
2	$1-3$	$1-3$	11	$3-4$	$3-4$
3	$1-4$	$1-2-4$	12	$3-5$	$3-5$
4	$1-5$	$1-3-5$	13	$3-6$	$3-4-6$
5	$1-6$	$1-6$	14	$4-6$	$4-6$
6	$2-3$	$2-3$	15	$5-3$	$5-2$
7	$2-4$	$2-4$	16	$5-3$	$5-3$
8	$2-5$	$2-3-5$	17	$5-4$	$5-3-4$
9	$2-6$	$2-4-6$	18	$5-6$	$5-3-4-6$

We assume that the average demands for different services follow a linear model. Furthermore, average demand for a particular type of service provided by a firm depends only on the own and non-own prices for that service, i.e.,

$$d_i^f(p_i, t) = \alpha_i^f(t) - \beta_i^f(t) \cdot p_i^f(t) + \sum_{g \neq f} \gamma_i^{f,g}(t) \cdot p_i^g(t)$$

for all $f \in \mathcal{F}, i \in \mathcal{S}, t \in [t_0, t_1]$ and $\alpha_i^f(t), \beta_i^f(t), \gamma_i^{f,g}(t) \in \mathbb{R}_{++}^1$. The random variables associated with the stochastic demand are assumed to follow a uniform distribution with range [0.3, 1.7] which is independent of service types and firms. The booking period runs from clock time 0 to 10. Lower and upper bounds on the service prices are set independent of the identity of firms depending on the number of legs used in a particular itinerary, as shown below.

No. of legs in the itinerary	**1**	**2**	**3**
$\mathbf{p}_{i,\min}^f$	75	150	225
$\mathbf{p}_{i,\max}^f$	500	650	725

Each service provider has different, yet mostly comparable capacities of 12 legs, as listed below.

Legs	**1**	**2**	**3**	**4**	**5**	**6**
Firm1	300	200	200	400	800	400
Firm2	320	360	180	320	700	440
Firm3	280	300	220	380	820	380

Legs	**7**	**8**	**9**	**10**	**11**	**12**
Firm1	400	200	600	200	200	240
Firm2	500	210	720	300	400	320
Firm3	440	220	480	280	600	340

6.2 Bidirectional Pricing

6.2.1 Competitive Pricing and Resource Allocation

In Figure 2, we plot competitive bidirectional equilibrium price trajectories grouped by firms over the planning horizon. The equilibrium rate of provision of services by different firms is plotted vs. time in Figure 3. The average demand for 18 different services (grouped by firms) is plotted vs. time in Figure 4.

6.2.2 Cooperative Pricing and Resource Allocation

In this section we discuss pricing and allocation strategies for the firms when they are in complete collusion. Figure 5 plots collusive bidirectional equilibrium price trajectories grouped by firms over the planning horizon. The equilibrium rate of provision of services by different firms is plotted vs. time in Figure 6. The average demand for 18 different services (grouped by firms) is plotted vs. time in Figure 7.

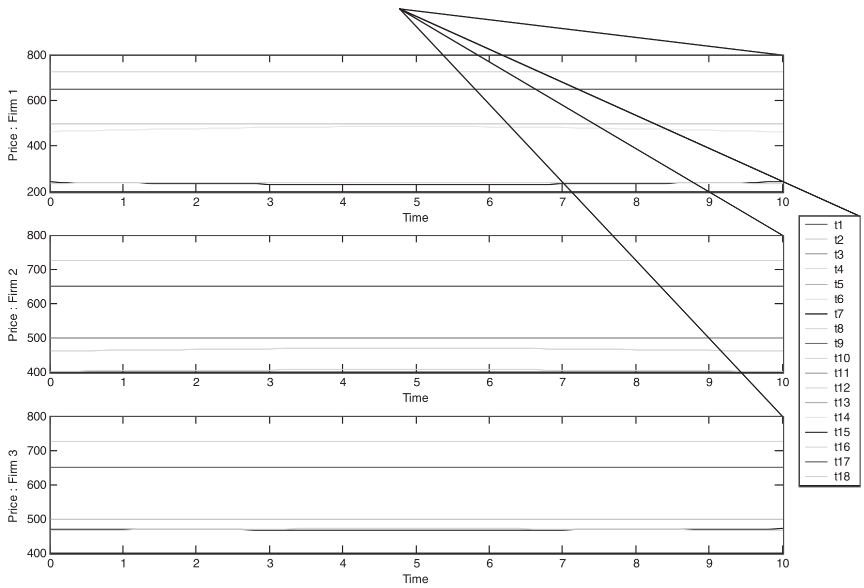


Figure 2: Plot of noncooperative equilibrium service prices (grouped by firms) vs. time.

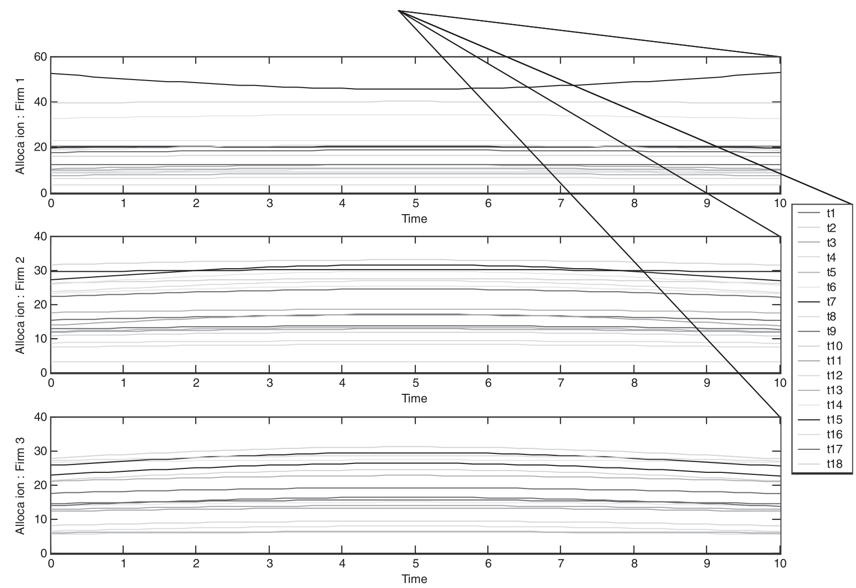


Figure 3: Plot of noncooperative equilibrium rate of provision of services (grouped by firms) vs. time.

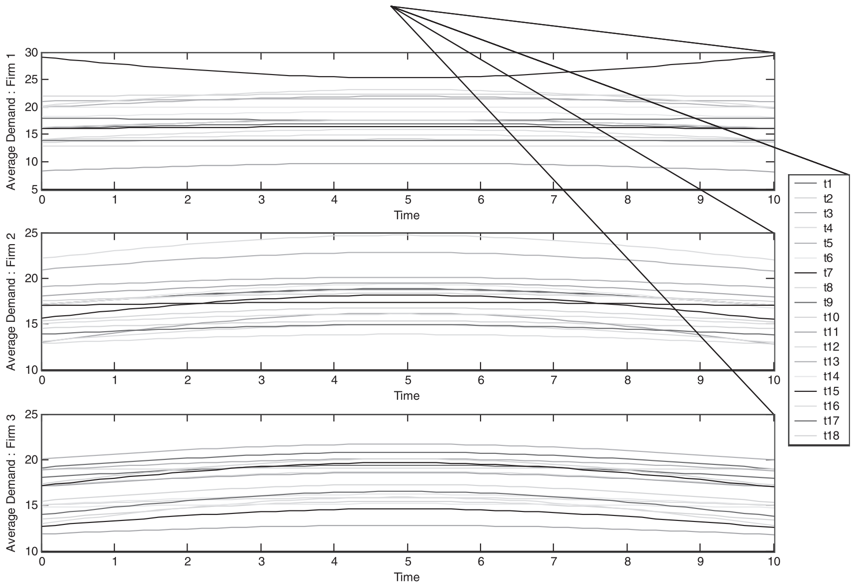


Figure 4: Average demand for 18 services (grouped by the firms) over the booking period.

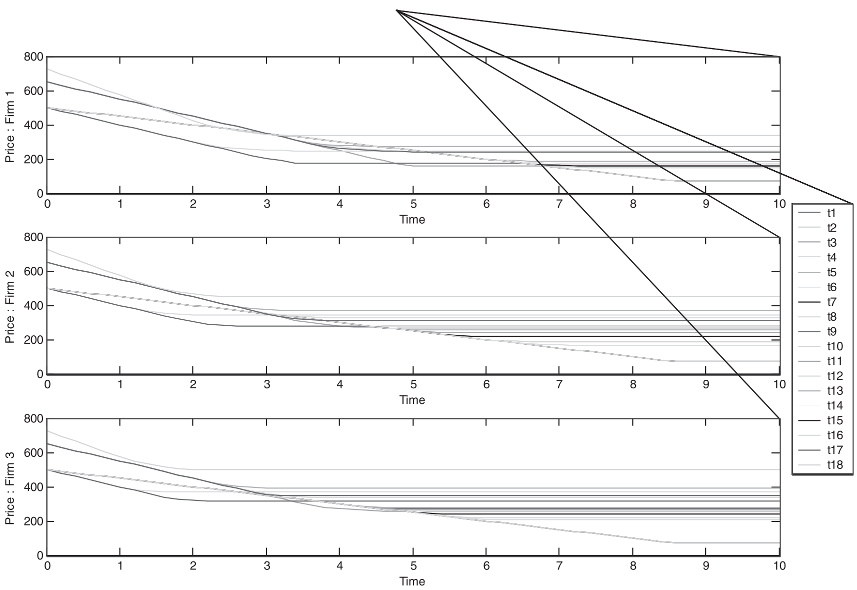


Figure 5: Plot of cooperative equilibrium service prices (grouped by firms) vs. time.

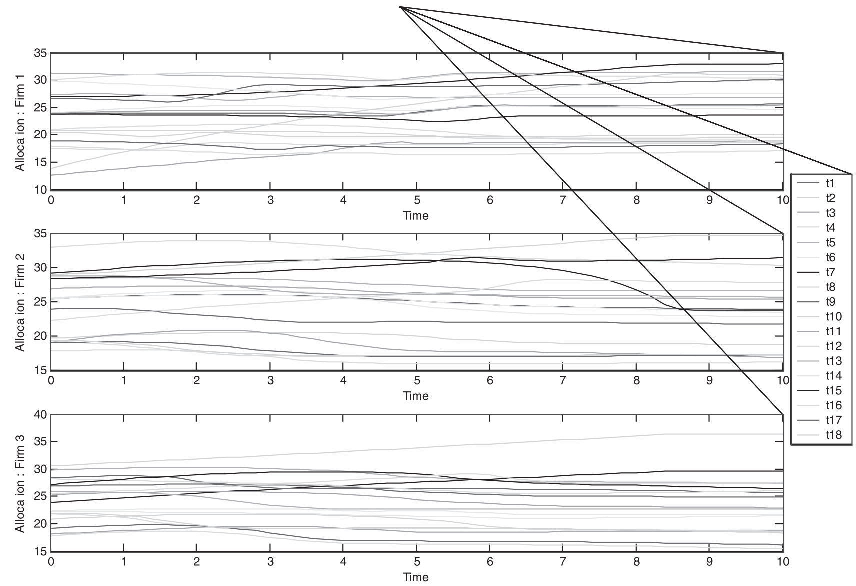


Figure 6: Plot of rate of provision of services (grouped by firms) vs. time when firms form collusion.

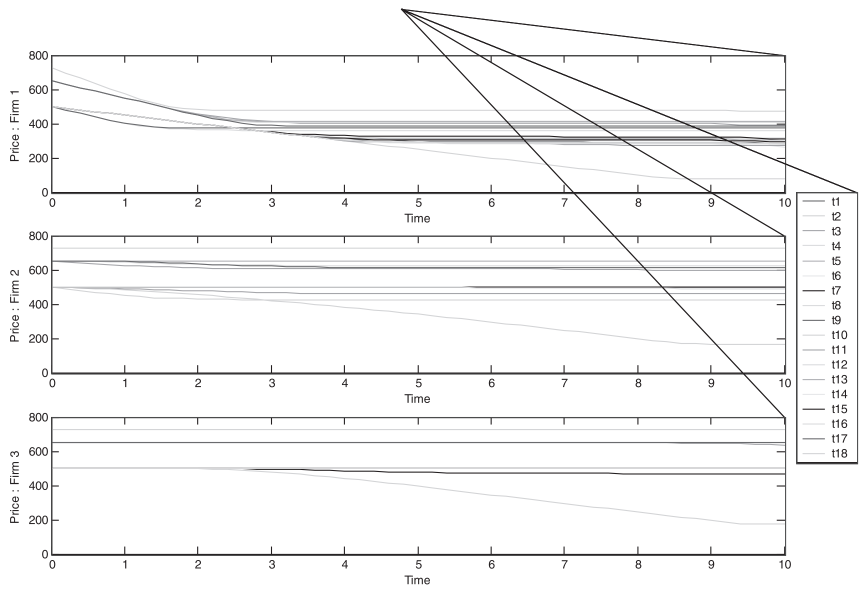


Figure 7: Average demand for 18 services (grouped by the firms) over the booking period when all the firms collude.

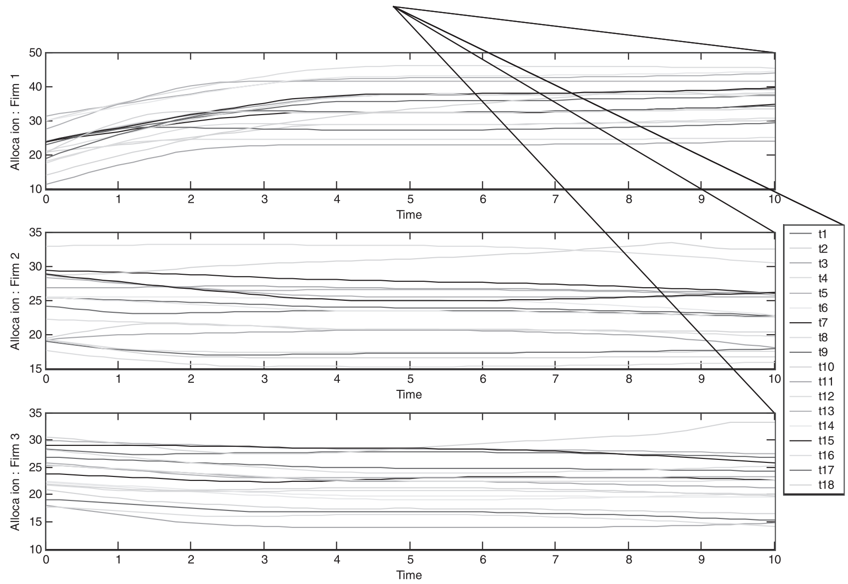


Figure 8: Comparison of cooperative service price with competitive service price for 18 different services provided by 3 firms.

6.2.3 Comparison of Competitive vs. Collusive Pricing and Allocations

In Figure 8 we compare competitive pricing with that under collusion by plotting $(\tilde{p}_i^f(t) - p_i^{f*}(t))$ vs. time for all 18 services and 3 firms where $\tilde{p}$ and p^* are the cooperative and competitive service prices respectively. It is evident from the plot that competition leads to underpricing of services which is, for certain services, as high as \$500 per unit of service (an airline seat in our example). Similarly, we plot in Figure 9 the difference between $\tilde{u}_i^f(t) - u_i^{f*}(t)$ over time where $\tilde{u}$ and u^* are the rates of provision of different services by the firms when they compete with each other and collude respectively. We observe that for most of the service types competition leads to overallocation of resources.

6.2.4 Comparison of Expected Revenue

When the firms are in competition, they generate a total revenue of \$2,643,900 with the following breakdown of individual's expected revenue.

Firms	Expected Revenue
Firm 1	\$780,200
Firm 2	\$1,013,900
Firm 3	\$849,800

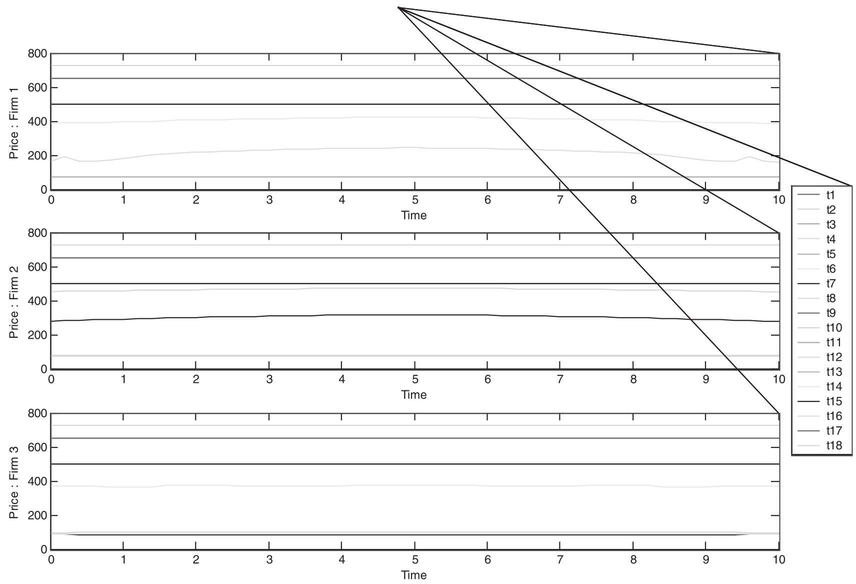


Figure 9: Comparison of cooperative service provision rate with competitive service provision rate for 18 different services provided by 3 firms.

However, when they collude, their total expected revenue increases to \$2,926,400. Therefore in this example competition causes an efficiency loss of around 11%.

6.3 Markdown Pricing

6.3.1 Competitive Pricing and Resource Allocation

When we consider markdown pricing the rate of service provision trajectories show remarkably different qualitative properties compared to bidirectional pricing. In Figure 10 we plot competitive markdown price trajectories grouped by firms over the planning horizon. The equilibrium rate of provision of services by different firms is plotted vs. time in Figure 11.

6.3.2 Cooperative Pricing and Resource Allocation

Figure 12 describes cooperative markdown price trajectories grouped by firms over the planning horizon. The equilibrium rate of provision of services by different firms is plotted vs. time in Figure 13.

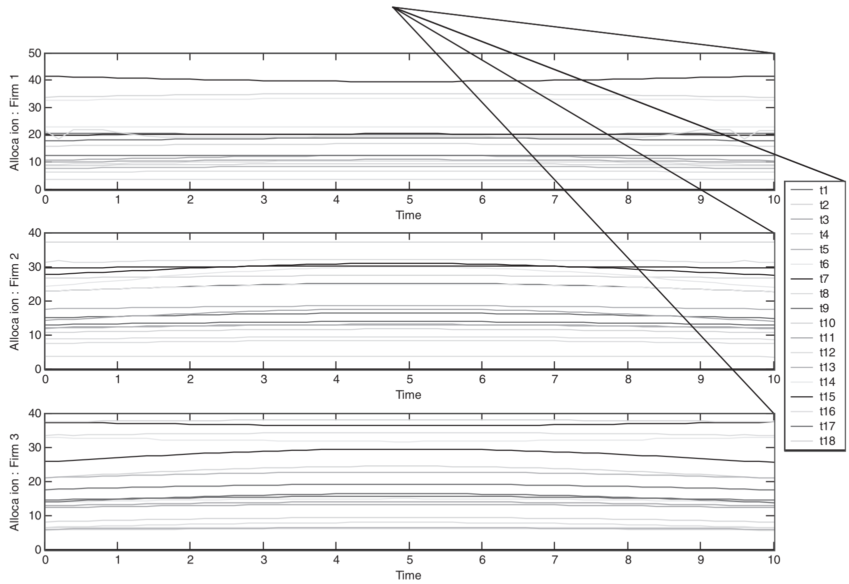


Figure 10: Plot of noncooperative equilibrium service prices (grouped by firms) vs. time under markdown pricing.

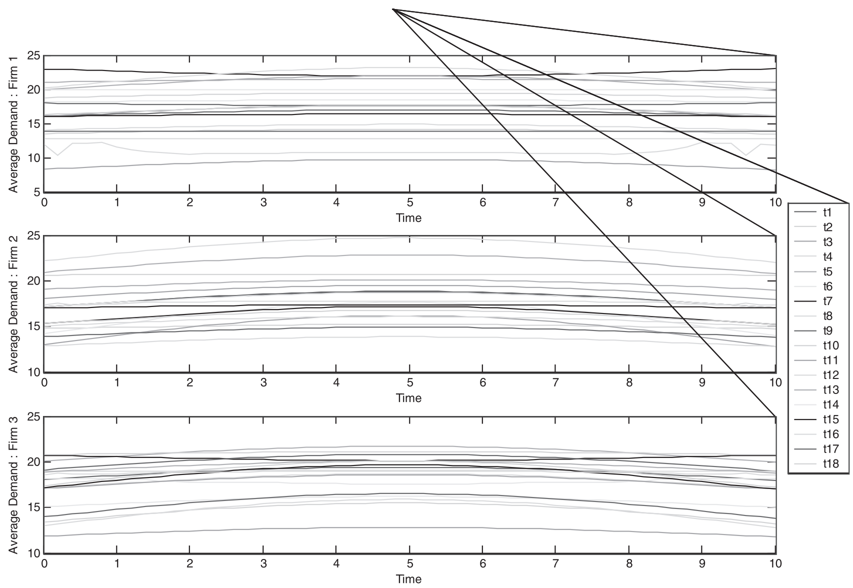


Figure 11: Plot of noncooperative equilibrium rate of provision of services (grouped by firms) vs. time under markdown pricing scheme.

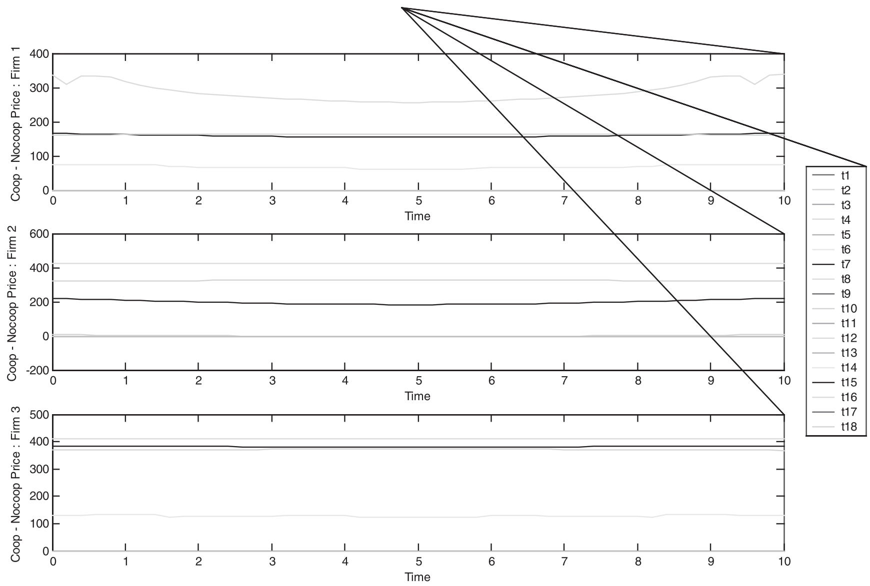


Figure 12: Plot of cooperative equilibrium service prices (grouped by firms) vs. time under markdown pricing.

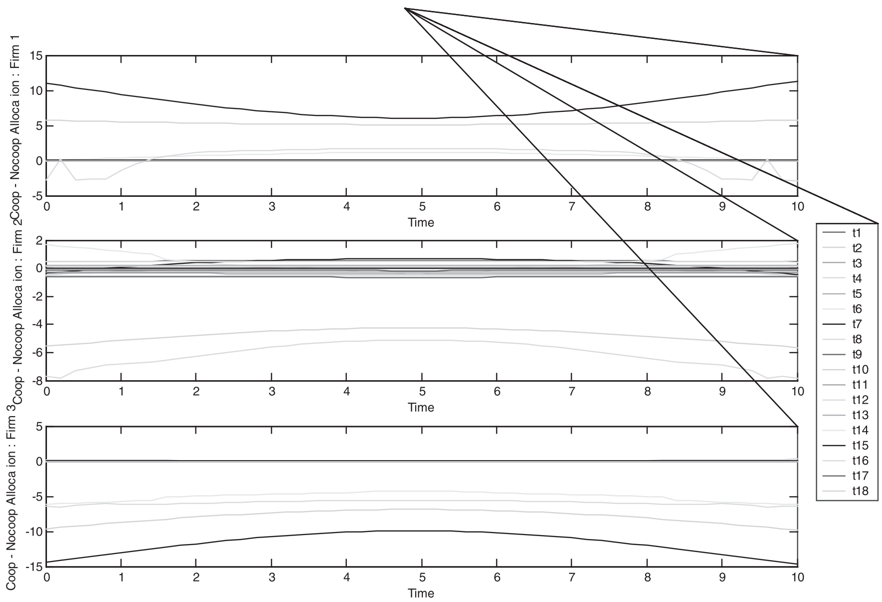


Figure 13: Plot of cooperative equilibrium rate of provision of services (grouped by firms) vs. time under markdown pricing scheme.

7 Conclusions and Future Work

We have shown that competitive service pricing and resource allocation, cornerstones of revenue management, may be articulated as differential variational inequalities. Additionally, we presented the numerical solution of an example service network pricing and resource allocation problem, motivated by the airline industry, using a fixed point algorithm. This example suggests that the differential variational inequality perspective for dynamic Cournot–Nash games is computationally tractable.

This work may be expanded upon to include more complicated allocation methods such as theft nesting. Other possible extension includes modeling scenarios where customers pay for service and are permitted to cancel their order before receiving the service. Such scenarios lead to overbooking policies as seen in the airline industry. In addition to the pricing and resource allocation decisions made in the model presented here, the firms must also determine a level of overbooking, in light of customer no-shows and cancellations, that maximizes the utilization of seats on a flight. We have studied the joint pricing-allocation-overbooking problem at some depth in [13].

References

[1] J. P. Aubin and A. Cellina, *Differential Inclusions: Set Valued Maps and Viability Theory*. Springer, New York, 1984.

[2] T. L. Friesz, M. Rigdon, and R. Mookherjee, Differential variational inequalities and shipper dynamic oligopolistic network competition, *to appear in Transportation Research B*, 2005.

[3] T. L. Friesz and R. Mookherjee, Differential variational inequalities with state-dependent time shifts, *Preprint*, 2005.

[4] J. I. McGill and G. J. van Ryzin, Revenue management: Research overview and prospects, *Transportation Science*, vol. 33, no. 2, pp. 233–256, 1999.

[5] G. Perakis and A. Sood, Competitive multi-period pricing for perishable products: A robut optimization approach, *to appear in Mathematical Programming B*, 2005.

[6] K. T. Talluri and G. J. van Ryzin, *The Theory and Practice of Revenue Management*. Kluwer Academic Publishers, New York, 2004.

[7] S. Karlin and C. R. Carr, *Prices of Optimal Inventory Policy*. In Studies in Applied Probability and Management Science, K. J. Arrow, S. Karlin and H. Scarf editors, Stanford University Press, Stanford, CA, 1962.

[8] F. Bernstein and A. Federgruen, Dynamic inventory and pricing models for competing retailers, *Naval Research Logistics*, vol. 51, pp. 258–274, 2004.

[9] S. P. Sethi and G. L. Thompson, *Optimal Control Theory: Applications to Management Science and Economics, Second edition*. Kluwer, Dordrecht, 2000.

[10] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*. Hemisphere Publishing Co., New York, 1975.

[11] M. Minoux, *Mathematical Programming Theory and Algorithms*. John Wiley, New York, 1986.

[12] R. Mookherjee and T. L. Friesz, Pricing and resource allocation in dynamic oligopolistic network competition with demand uncertainty, *Working paper, Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park*, 2005.

[13] R. Mookherjee and T. L. Friesz, Pricing, Allocation and overbooking in Dynamic Service Network competition when Demand is uncertain, under review, 2006.

PART VI
Numerical Methods and Algorithms in Dynamic Games

Numerical Methods for Stochastic Differential Games: The Ergodic Cost Criterion

Harold J. Kushner
Division of Applied Mathematics
Brown University
Providence RI 02912, USA
hjk@dam.brown.edu

Abstract

The Markov chain approximation method is a widely used, relatively easy to apply, and efficient family of methods for the bulk of stochastic control problems in continuous time, for reflected jump-diffusion type models. It is the most general method available, it has been shown to converge under broad conditions, via probabilistic methods, and there are good algorithms for solving the numerical problems, if the dimension is not too high (see the basic reference [16]). We consider a class of stochastic differential games with a reflected diffusion system model and ergodic cost criterion, where the controls for the two players are separated in the dynamics and cost function. The value of the game exists and the numerical method converges to this value as the discretization parameter goes to zero. The actual numerical method solves a stochastic game for a finite state Markov chain and ergodic cost criterion. The essential conditions are nondegeneracy of the diffusion and that a weak local consistency condition hold "almost everywhere" for the numerical approximations, just as for the control problem. The latter is close to a minimal condition. Such ergodic and "separated" game models occur in risk-sensitive and robust control.

1 Introduction

The Markov chain approximation method of [11,12,16] is a widely used method for the numerical solution of virtually all of the standard forms of stochastic control problems with reflected jump-diffusion models. It is robust and can be shown to converge under very broad conditions. Extensions to approximations for two-person differential games with discounted, finite time, stopping time, and pursuit-evasion games were given in [14] for reflected diffusion models where the controls for the two players are separated in the dynamics and cost rate functions, and the cost functions were of the discounted or stopping time types. This chapter outlines the basic ideas behind the methods of numerical approximation and the proofs of convergence for two-player stochastic dynamic games with the same systems model, but where the cost function is ergodic. It can also serve as a review of the Markov chain approximation method. The convergence results are adapted from [15], where more details on some of the proofs can be found. Such ergodic and "separated" models occur, for example, in risk-sensitive and robust control. See [8] for a formulation of risk-sensitive control in terms of a stochastic differential game for a diffusion that fits our form once "numerical boundaries" are added, provided that the noise covariance does not depend on the control. Our original motivation was the game formulation of risk-sensitive control problems for queues in heavy traffic where the state is confined to some convex polyhedron by boundary reflection [13]. In such problems, the robust control problem can be formulated as a game problem with a hyperrectangular state space. If the system state is not a priori confined to a bounded set, then for numerical purposes it is commonly necessary to bound the state space artificially by adding a reflecting boundary and then experimenting with the bounds to ensure that they have little effect in the regions of greatest concern. We restrict attention to the convex polyhedron to avoid minor details which can be distracting.

Results concerning convergence of numerical approximations for various forms of deterministic or stochastic continuous-state and time dynamic games have appeared. The upper value for a deterministic game (an ordinary differential equation (ODE) model) was treated by the Markov chain approximation method in [17,18]. Results for various deterministic problems are provided in [1–4,19,20]. The actual numerical methods which are used in the computations are of the Markov chain approximation type, although the proofs are sometimes based on subsequent partial differential equation (PDE) techniques. Indeed, the algorithms implied by viscosity solution methods are also of the Markov chain approximation type, as shown in [16, Chapter 16].

The proofs here are purely probabilistic and have the advantage of providing intuition concerning numerical approximations, as well as requiring relatively weak conditions. The essential conditions are *weak-sense* existence and uniqueness of the solution to the controlled equations, "almost everywhere" continuity of the dynamical and cost rate terms, and a natural "local consistency" condition. The consistency and continuity need hold only almost everywhere, allowing discontinuities in the drift and cost rate to be treated under appropriate conditions (see, e.g., [16, Theorem 10.5.2]). The numerical approximations can be represented as processes that are "close" to the original, which gives an intuitive content to the method. The actual proof of convergence of the numerical method is not difficult. However, as with numerical methods in general, convergence proofs depend on properties of the original model. In particular, one requires weak-sense uniqueness of the uncontrolled system, as well as model robustness in that small changes in the dynamics or control have only a small effect on the distributions. It is these considerations (that are of independent interest) that take most of the space, particularly since the properties of the underlying systems are not widely known. We need to know that there is a value to the game. But for this to be meaningful, we need to know that there is a policy that achieves it with a well-defined associated solution process.

The systems model, classes of controls, and the cost function are given in Section 2, together with the assumptions on the model. Three classes of controls are considered, ordinary nonanticipative controls, relaxed controls, and relaxed feedback controls, each of which has an important role to play. The dynamical model is the reflected stochastic differential equation (2.1) or (2.3). Owing to the boundary reflection, such models are examples of the *Skorohod problem* [6,13,16]. The conditions on the boundary of the state space A1 and A2 (Section 2), seem to cover the great majority of cases of current interest. In particular, the systems that arise from heavy traffic approximations to queueing and communications networks are covered. The conditions are obvious when the state space is a hyperrectangle and the reflection directions are the interior normals, and this condition is commonly used when the boundary conditions are used simply to bound the state space for numerical purposes.

Player 1 is minimizing and Player 2 maximizing. For the ergodic cost problem, it is usually desired that time-independent feedback controls be used. These controls are chosen at $t = 0$ and they can be either time-independent feedback or relaxed feedback controls. For the *min max* game Player 1 chooses its control first, and Player 2 chooses first for the *max min* game. The order will not matter since the upper and lower values will be the same. For technical reasons we also allow the player who chooses last to use a time-dependent control that is not necessarily feedback. The game has a value with this allowance and the value is associated with purely feedback controls, so the allowed use of time-dependent controls for the player who chooses last provides no advantage to him. We do not know whether it is possible for the player who chooses first to achieve a better value by using some non-feedback control strategy. But note that any approach that leads to a well-posed Isaacs equation will yield feedback controls. Since the player choosing first uses feedback controls, complications that arise due to the definitions of strategy in the time-dependent case do not occur. In this sense this chapter is simpler than [14]. However, the ergodic cost criterion creates its own set of difficulties.

The Markov chain approximation numerical method is discussed in Section 3 and the *local consistency condition* is stated. We obtain the approximating chain and cost function as in [16] for the pure control problem. The methods to be used are very different than those used in [14]. They are both based on the theory of weak convergence [5,7]. But the current case depends on the approximations to the ergodic cost control problem as developed in [13, Chapter 4], particularly on results concerning the "continuity" of the invariant measures and costs on the controls. Due to space limitations, we have simply stated some of the smoothness and approximation results, with citations to the proofs. We try to take advantage of the results in [13,16,15], wherever possible, while maintaining clarity in the flow of ideas.

The method of Girsanov transformations provides a very powerful tool for the treatment of stochastic control problems [9], and is widely used there, although it is less well known in the game theory literature. One starts with an uncontrolled system and the control is introduced by using the Girsanov measure transformation.

For our problem, with this technique one can show that basic properties of the uncontrolled system (say, concerning the nature of the transition function or invariant measure) carry over to the controlled system in a way that is "uniform" in the control. The general model is obtained in this way in Section 4, which also contains many background convergence and approximation results. The proof of convergence of the numerical method is given in Section 5, and it depends on the fact that the original game has a value. An outline of the proof of this last fact is given in Section 6.

2 The Model and Background

2.1 The Systems Model and Assumptions

The systems model with feedback controls is

$$dx = b(x, u(x))\,dt + \sigma(x)\,dw + dz, \quad x(t) \in G \subset \mathbb{R}^r, \quad u = (u_1, u_2), \quad (2.1)$$

where $\mathbb{R}^r$ denotes Euclidean r-space, $b(x, \alpha) = b_1(x, \alpha_1) + b_2(x, \alpha_2)$, and $w(\cdot)$ is a standard vector-valued Wiener process. The process is constrained to G by the reflection term $z(\cdot)$, which can change only at t where $x(t) \in \partial G$. The $x(\cdot), z(\cdot), u(\cdot)$ are always taken to be nonanticipative with respect to $w(\cdot)$. Player 1 (control u_1) is minimizing and Player 2 (control u_2) is maximizing. The following conditions are assumed to hold. Conditions A1 and A2 are standard for the Skorohod problem. They are unrestrictive in applications and an example is given below. The extension of $d(\cdot)$ in A3 is needed only because the approximating processes might be defined on a slightly larger set than G. See [13, Chapter 3] for more detail on reflected diffusions. Conditions A0 and A1 imply that $b(\cdot)$ is bounded.

A0. We have canonical control values $\alpha_i \in \mathcal{U}_i$ compact, $i = 1, 2$, $\sigma^{-1}(x)$ is bounded, $b(\cdot)$ is continuous, and $\sigma(\cdot)$ is Hölder continuous. Define $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2$.

A1. G is a convex polyhedron with a finite number of sides $\partial G_i, i = 1, \ldots,$ and an interior. Let ∂G_i^0 denote the relative interior of the ith face. Let n_i, $i = 1, \ldots,$ denote the interior normals. On ∂G_i^0, the (unit) reflection direction is d_i, and $\langle d_i, n_i \rangle > 0$. The possible reflection directions on the intersections of the ∂G_i are in the convex hull of the directions on the adjoining faces. No more than r constraints are active at any boundary point.

A2. Suppose that x lies in the intersection of $\{\partial G_{i_1}, \ldots, \partial G_{i_k}\}$, $k > 1$. Let $NC(x)$ denote the convex hull of the normals $\{n_{i_1}, \ldots, n_{i_k}\}$. Let $d(x)$ denote the set of reflection directions at x. Then there is a vector $v \in NC(x)$ such that $\gamma' v > 0$ for all $\gamma \in d(x)$ (see Figure 2.1 for an example). There is a neighborhood $N(\partial G)$ and an extension of $d(\cdot)$ to $\overline{N(\partial G)} - G$ that is *upper semicontinuous* in the following sense: For each $\epsilon > 0$, there is $\rho > 0$ that goes to zero as $\epsilon \to 0$ and such that if $x \in N(\partial G) - G$ and distance $(x, \partial G) \le \rho$, then $d(x)$ is in the convex hull of the directions $\{d(v);\, v \in \partial G, \text{distance } (x, v) \le \epsilon\}$.

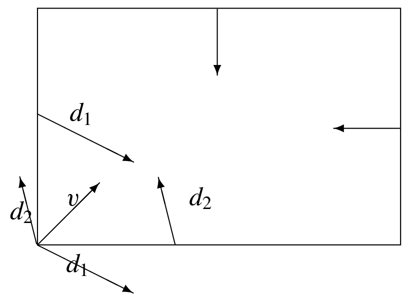


Figure 2.1a. Reflection directions for queue model of Figure 2.1b.

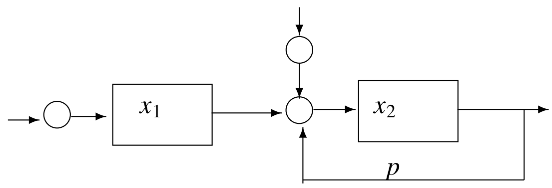


Figure 2.1b. A queue model with feedback.

A3. The uncontrolled model (where $b = 0$) has a unique weak-sense solution for each initial condition.

Condition A3 holds if $\sigma(\cdot)$ is Lipschitz continuous. There is an (a.e.) unique representation of the form $z(t) = \sum_i d_i y_i(t)$, where $y_i(\cdot)$ is the contribution to $z(\cdot)$ from the reflection on the ith face [13, Section 3.6 and Theorem 4.3.6]. The $y_i(\cdot)$ often have a physical meaning. In the example below the y_i corresponding to the outer boundaries correspond to lost messages due to buffer overflow and there would be a cost associated with that loss.

Figure 2.1 gives an example of reflection directions. The system is the heavy traffic limit [13] of the queue models of Figure 2.1b, where $x_1(t)$ (resp., $x_2(\cdot)$) is the scaled level of the first (resp., second) queue, and the scaled buffers are finite. The state space is a rectangle whose lower left corner is the origin, since $x_i(t) \geq 0$. The directions of reflection are determined by the probability $p \in [0, 1)$ of a customer being fed back, and A2 and A3 always hold.

2.2 Relaxed Feedback Controls $m_i(\cdot)$

In work on approximations, existence of optimal controls, and limit theorems for controls, it is usual to work with what are called relaxed controls. They are used for analytical (not practical) purposes, and they greatly facilitate the analysis. Although they are extensions of the ordinary controls, their use does not alter the min, max, minmax values, nor the form of the Bellman or Isaacs equations. A *relaxed feedback control* $m_i(x, \cdot)$, $i = 1, 2$, is a probability measure on the Borel sets of $\mathcal{U}_i$ for each $x \in G$, and $m_i(\cdot, A)$ is Borel measurable for each Borel set in $\mathcal{U}_i$.

Let $0 \in \mathcal{U}_i$, corresponding to no control ($b(x,0) = 0$). The associated controlled drift term for player i is

$$b_{i,m_i}(x) = \int_{\mathcal{U}_i} b_i(x,\alpha) m_i(x, d\alpha), \tag{2.2}$$

with controlled system

$$dx(t) = b_m(x(t))\, dt + \sigma(x(t))\, dw(t) + dz(t), \tag{2.3}$$

where $b_m(x) = b_{1,m_1}(x) + b_{2,m_2}(x)$. Define $m(x,\cdot) = m_1(x,\cdot)m_2(x,\cdot)$. Loosely speaking, a relaxed feedback control is a "continuous limit" of a randomized feedback control.

2.3 Relaxed Controls

General nonanticipative relaxed controls might also be used for the player that goes last. This allowance is made for technical reasons. But the game is still solved by purely feedback controls. An admissible *relaxed control* $r(\cdot)$ is a measure on $\mathcal{U} \times [0,\infty)$ such that $r(\mathcal{U} \times [0,t]) = t$ and is "adapted." Loosely and intuitively, $r(A \times [0,t]) =$ amount of time that the control takes values in $A \subset \mathcal{U}$ on $[0,t]$. Given a relaxed control $r(\cdot)$, there is a derivative $r_t(\cdot)$: $r(d\alpha dt) = r_t(d\alpha)\, dt$, where

$$r_t(A) = \lim_{\delta \to 0} \frac{r(A \times [t-\delta, t])}{\delta}.$$

For an ordinary control $u(\cdot)$, the relaxed control representation is $r_t(d\alpha) = I_{\{u(t) \in d\alpha\}}$. Define the relaxed controls $r_i(\cdot)$ for player i analogously.

2.4 The Weak Topology for Relaxed Controls

For relaxed controls $r^n(\cdot), r(\cdot)$, we have $r^n(\cdot) \to r(\cdot)$ if

$$\int_0^\infty \int_{\mathcal{U}} f(\alpha,s) r^n(d\alpha\, ds) \to \int_0^\infty \int_{\mathcal{U}} f(\alpha,s) r(d\alpha ds)$$

for each continuous $f(\cdot)$ with compact support. The set of relaxed controls is compact in the weak topology. The system can be written as follows, with $r_i(\cdot)$ being the relaxed control for player i:

$$\begin{aligned} dx(t) = {} & \int_{\mathcal{U}_1} b_1(x(t),\alpha_1) r_1(d\alpha_1 dt) + \int_{\mathcal{U}_2} b_2(x(t),\alpha_2) r_2(d\alpha_2 dt) \\ & + \sigma(x(t)) dw(t) + dz(t). \end{aligned}$$

2.5 The Cost Function

Define (continuous functions) $k(x,\alpha) = k_1(x,\alpha_1) + k_2(x,\alpha_2)$. For admissible relaxed controls $r_1(\cdot), r_2(\cdot)$ for players 1, 2, resp., define the relaxed control $r(\cdot)$

via its derivative as $r_t(d\alpha) = r_{1t}(d\alpha_1)r_{2t}(d\alpha_2)$. Let E_x^r denote the expectation under initial condition x and control r. Define (where $c_i \geq 0$)

$$k_r(x,t) = \int_{\mathcal{U}_1} k_1(x,\alpha_1)r_{1,t}(d\alpha_1) + \int_{\mathcal{U}_2} k_2(x,\alpha_2)r_{2,t}(d\alpha_2),$$
$$\gamma_T(x,r) = \frac{1}{T}E_x^r \int_0^T k_r(x(s),s)\,ds + \frac{1}{T}E_x^r c' y(T).$$

For a relaxed feedback control $m(\cdot)$, define $k_m(x) = \int_{\mathcal{U}} k(x,\alpha)m(x,d\alpha)$,

$$\gamma_T(x,m) = E_x^m \frac{1}{T}\int_0^T k_m(x(s))\,ds + \frac{1}{T}E_x^m c' y(T).$$

If both players use relaxed feedback controls, then the cost function is $\gamma(m) = \lim_T \gamma_T(x,m)$, and the limit will exist. If the player that goes last uses a relaxed (but not necessarily feedback) control, then the cost function is (for Player 1 choosing first; the use of $\liminf$ is just a convention) $\gamma(x,m_1,r_2) = \liminf_T \gamma_T(x,m_1,r_2)$, or (for Player 2 choosing first; the use of $\limsup$ is just a convention) $\gamma(x,r_1,m_2) = \limsup_T \gamma_T(x,r_1,m_2)$. Define

$$\bar{\gamma}^+ = \inf_{m_1}\sup_{m_2}\gamma(m_1,m_2) \geq \bar{\gamma}^- = \sup_{m_2}\inf_{m_1}\gamma(m_1,m_2).$$

3 Markov Chain Approximations

We will now give a quick overview of the Markov chain approximation method of [11,12,16], starting with some comments for the case where there is only one player. The method consists of two steps. Let $h > 0$ be an approximation parameter. The first is the determination of a finite-state controlled Markov chain ξ_n^h that has a continuous time interpolation that is an "approximation" of the process $x(\cdot)$. The second step solves the optimization problem for the chain and a cost function that approximates the one used for $x(\cdot)$. Under a natural "local consistency" condition, the minimal cost function for the chain converges to the minimal cost function for the original problem. In applications, the optimal control for the original problem is also approximated. The approximating chain and local consistency conditions are the same for the game problem. The reference [16] contains a comprehensive discussion of many automatic and simple methods for getting the transition probabilities of the chain. The approximations "stay close" to the physical model and can be adjusted to exploit local features.

The simplest state space for the chain (and the one that we will use for simplicity in the discussion) is based on the regular h-grid S_h in $\mathbb{R}^r$. It is only the subset of points in $G_h = S_h \cap G$ and their immediate neighbors that are of interest. On G_h the chain "approximates" the diffusion part of (2.1) or (2.3). If the chain tries to leave G_h, then it is returned immediately, consistently with the reflection direction where it exits. Let ∂G_h^+, an approximation to the reflecting boundary, denote the set of points not in G_h to which the chain might move in one step from some point in G_h. Define $G_h^+ = G_h \cup \partial G_h^+$. This two-step procedure on the boundary simplifies both coding and analysis.

Next we define the basic condition of *local consistency* for the part of a Markov chain ξ_n^h that is on G_h. Define $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$ and let $E_{x,n}^{h,\alpha}$ denote the expectation given the data to step n, the time of computation of ξ_n^h, with $\xi_n^h = x$ and control value α to be used in the next step. Suppose that there is $\Delta t^h(\cdot)$ (in our case this will be a constant of order $O(h^2)$ and it is easy to get; see the example below) such that

$$\begin{aligned} E_{x,n}^{h,\alpha}\Delta\xi_n^h &= b(x,\alpha)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)), \\ \operatorname{cov}_{x,n}^{h,\alpha}\big[\Delta\xi_n^h - E_{x,n}^{h,\alpha}\Delta\xi_n^h\big] &= \sigma(x)\sigma'(x)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)). \end{aligned} \tag{3.1}$$

In the sense of (3.1), the chain approximates the diffusion. The local consistency conditions on the reflecting boundary are obtained analogously, but we use $\Delta t^h = 0$ there (instantaneous reflection). Basically, for $x = \xi_n^h \in \partial G_h^+$, one sends the state back to a set of closest points in G_h such that the average direction is the local reflection direction, modulo an asymptotically (as $h \to 0$) vanishing error. An illustration is given in Figure 3.1, where G is the northeast sector bounded by the darker lines. In the figure, x goes to either x' or x'' with probabilities that ensure the mean reflection direction d_1. The interpolation intervals Δt_n^h are automatically given when the transition probabilities are calculated. But in the ergodic cost case they will be constant.

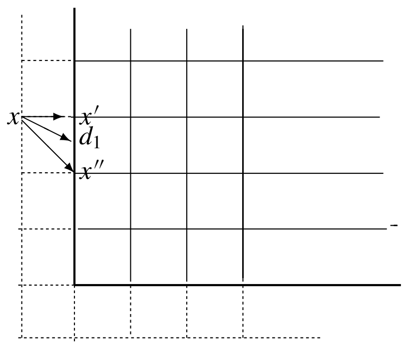


Figure 3.1. Illustration of reflection directions.

3.1 An Example of an Approximating Chain

The simplest case is one dimensional and where $h|b(\alpha,x)| \le \sigma^2(x)$. Then we can use the transition probabilities and interval, for $x \in G_h$,

$$p^h(x, x \pm h|\alpha) = \frac{\sigma^2(x) \pm hb(x,\alpha)}{2\sigma^2(x)}, \quad \Delta t^h(x,\alpha) = \frac{h^2}{\sigma^2(x)}, \;\; \Delta t_n^h = h^2/\sigma^2(\xi_n^h).$$

To get a chain with constant interpolation interval $\overline{\Delta}$ for $x \in G_h$, simply take the minimal interval and let the slack be taken up by transiting to x as well [16, Section 7.7], yielding

$$\bar{p}^h(x, y|\alpha) = p^h(x, y|\alpha)(1 - \bar{p}^h(x, x|\alpha)), \qquad \text{for } x \neq y,$$

$$\bar{p}^h(x, x|\alpha) = 1 - \frac{\overline{\Delta}^h}{\Delta t^h(x, \alpha)}, \quad \overline{\Delta}^h = h^2 / \max_{x \in G} \sigma^2(x).$$

Discretize the costs as follows. For $x \in \partial G_h^+$, write (modulo an error that goes to zero as $h \to 0$ due to the approximation of the d_i near the boundary according to condition A2) $E_x\left[\xi_1^h - x\right] I_{\left\{\xi_0^h = x \in \partial G_h^+\right\}} = \sum_i d_i \delta y_i^h$. The cost rate for the chain is $C(x, \alpha) = k(x, \alpha)\Delta t^h(x, \alpha) + c' dy^h$. The reflections are instantaneous and are not controlled. Under our condition that the controls are separated in $b(\cdot)$, if desired one can construct the chain so that the controls are "separated" in that the one-step transition probability has the form

$$p^h(x, y|\alpha) = p_1^h(x, y|\alpha_1) + p_2^h(x, y|\alpha_2). \tag{3.2}$$

It can be seen that the chain has the "local properties" (conditional mean change and covariance) of the diffusion process.

3.2 The Numerical Problem

Under the nondegeneracy conditions on $\sigma(\cdot)$, we can (and will) suppose that the chain is a single ergodic class under each control. If we assume (3.2) (which can always be arranged to hold) then the upper and lower values

$$\bar{\gamma}^{h,-} = \sup_{u_2} \inf_{u_1} \gamma^h(u_1, u_2), \quad \bar{\gamma}^{h,+} = \inf_{u_1} \sup_{u_2} \gamma^h(u_1, u_2)$$

exist and are equal (i.e., $\bar{\gamma}^{h,\pm} = \bar{\gamma}^h$) and the Isaacs equation

$$g^h(\cdot) = \sup_{\alpha_2 \in \mathcal{U}_2} \inf_{\alpha_1 \in \mathcal{U}_1} \left[\sum_y p^h(x, y|\alpha) g^h(y) + C^h(x, \alpha) - \gamma^h \right], \tag{3.3}$$

has a solution $\gamma^h, g^h(\cdot)$, where $\gamma^h = \bar{\gamma}^h$ is unique, and $g^h(x)$ is the "potential" or "relative value" function [10]. The convergence theorems do not require that $\bar{\gamma}^{h,+} = \bar{\gamma}^{h,-}$ or that (3.2) holds. Both the upper and lower values will still converge to $\bar{\gamma}$. If (3.2) does not hold, then the controls for the chain might have to be randomized or relaxed feedback controls used.

3.3 Continuous-Time Interpolations

The discrete-time chain ξ_n^h is used for the numerical computations. However, for the proofs of convergence a continuous-time interpolation $\psi^h(\cdot)$ approximating $x(\cdot)$ is used, and this will be a continuous-time Markov chain constructed as follows. Let ν_n^h, $n = 0, 1, \ldots$, be mutually independent and exponentially distributed random variables with unity mean, that are independent of the chain and controls. Recall that the interpolation interval Δt_n^h is either constant and positive for states in G^h (and will be denoted by $\overline{\Delta}^h$) or zero for states in ∂G_h^+. Let $u_n^h = (u_{1,n}^h, u_{2,n}^h)$ denote the control actions that are used at step n. Define $\Delta\tau_n^h = \nu_n^h \Delta t_n^h$ and $\tau_n^h = \sum_{i=0}^{n-1} \Delta\tau_i^h$. Define

$$\psi^h(t) = x(0) + \sum_{\tau_{i+1}^h \le t} \left[\xi_{i+1}^h - \xi_i^h\right], \tag{3.4}$$

$$z^h(t) = \sum_{\tau_{i+1}^h \le t} E_i^h\left[\xi_{i+1}^h - \xi_i^h\right] I_{\{\xi_i^h \in \partial G_h^+\}}, \tag{3.5}$$

where E_i^h denotes the expectation given the data to step i. Then the τ_n^h are the jump times.

Under any feedback control, the process $\psi^h(\cdot)$ is a continuous-time Markov chain. Define the continuous-time interpolations $u_i^h(\cdot)$ of the controls for player i by $u_i^h(t) = u_{i,n}^h$, $\tau_n^h \le t < \tau_{n+1}^h$, and let its relaxed control representation be denoted by $r_i^h(\cdot)$. We have the representation (modulo an asymptotically negligible error due to the approximation of the reflection direction close to but outside the boundary of G, as in A2)

$$z^h(t) = \sum_i d_i y_i^h(t),$$

where $y_i^h(\cdot)$ is the contribution from reflections on the ith face. Also (modulo an asymptotically negligible error that is due to the $o(\Delta t_n^h)$ terms in (3.1) and the use of the conditional expectation in (3.5)) there is a martingale $w^h(\cdot)$ (with respect to the filtration generated by the state and control processes, possibly augmented by an "independent" Wiener process) such that ([16, Sections 5.7.3 and 10.4.1])

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

$$\int_0^t b(\psi^h(s), u^h(s)) ds = \int_0^t \int_{\mathcal{U}} b(\psi^h(s), \alpha) r_s^h(d\alpha)\, ds.$$

Also, for any subsequence $h \to 0$, there is a further subsequence (also indexed by h for simplicity) such that [16] (the proof is the same as for the one player control case)[1]

$(w^h(\cdot), z^h(\cdot), \psi^h(\cdot), r^h(\cdot)) \Rightarrow (w(\cdot), z(\cdot), x(\cdot), r(\cdot))$, and

$$x(t) = x(0) + \int_0^t \int_{\mathcal{U}} b(x(s), \alpha) r_s(d\alpha)\, ds + \int_0^t \sigma(x(s))\, dw(s) + z(t),$$

[1]The symbol $\Rightarrow$ denotes weak convergence, the path space is that of Cadlag functions, and the Skorohod topology is used [7].

where $z(\cdot)$ is a reflection process and $(x(\cdot), z(\cdot), r(\cdot))$ are nonanticipative with respect to the standard Wiener process $w(\cdot)$. A main question is: If the $\{u_n^h\}$ solve the game problem for the chain, is the limit process optimal? We will return to this question in Section 5.

4 The Control Model via Girsanov Transformations

Let us be given a probability space with consistent measures $P_{x,t}$, $t < \infty$, filtration $\{\mathcal{F}_t\}$, $\mathcal{F} = \lim_t \mathcal{F}_t$, extension P_x of the $P_{x,t}$ to $\mathcal{F}$, and uncontrolled process $x(\cdot)$:

$$dx(s) = \sigma(x(s))\, dw(s) + dz(s), \quad x(0) = x \in G, \tag{4.1}$$

where $w(\cdot)$ is a standard $\mathcal{F}_t$-Wiener process, $z(\cdot)$ is the reflection process and $x(\cdot), z(\cdot)$ are $\mathcal{F}_t$-adapted. The control processes will be defined from the uncontrolled model (4.1) via the Girsanov measure transformation method.

Define the measure $P_{x,t}^m$ on $\mathcal{F}_t$ by

$$\zeta(t,m) = \int_0^t b_m'(x(s))\sigma^{-1}(x(s))\, dw(s) - \frac{1}{2}\int_0^t |\sigma^{-1}(x(s))b_m(x(s))|^2 ds,$$
$$R(t,m) = e^{\zeta(t,m)}, \qquad dP_{x,t}^m = R(t,m)\, dP_{x,t}.$$

$E_{x,t}R(t,m) = 1$ since $b(\cdot)$ is bounded, so that $P_{x,t}^m$ is a probability measure. Let P_x^m denote the extension of the $P_{x,t}^m$ to $\mathcal{F}$. Then the process $w_m(\cdot)$ defined by

$$dw_m(t) = dw(t) - [\sigma^{-1}(x(s))b_m(x(s))]dt$$

is an $\mathcal{F}_t$-standard Wiener process on $(\Omega, P_x^m, \mathcal{F})$ [9]. Rewrite the uncontrolled model (4.1) in terms of the new Wiener process

$$dx(t) = b_m(x(t))\, dt + \sigma(x(t))dw_m(t) + dz(t), \tag{4.2}$$

which yields the controlled model (2.3) under the new measure P_x^m. Write $P^m(x,t,\cdot)$ for the transition function under relaxed feedback control $m(\cdot)$. One can use relaxed (in lieu of relaxed feedback) controls in the construction. This method of defining the controlled model is good enough for our purposes. The probability space on which the processes are defined is not important and we work with an expected value cost criterion. Hence we can work with weak-sense solutions. The probability distribution of the process defined by the Girsanov transformation method is unique and equals that of (2.3), by the assumed weak-sense uniqueness for each control.

4.1 Illustration of Use of the Girsanov Transformation

To get a feeling for the way that the Girsanov transformation is used, details of some of the needed results will be presented. More details are given in [15].

Theorem 4.1. *Let* $m^n(y,\cdot) \Rightarrow m(y,\cdot)$ *for almost all* y. *Then for any* $0 < t_0 < t_1 < \infty$ *and bounded and measurable real-valued function* $f(\cdot)$, *with* $\sup_x |f(x)| \le 1$, *we have*

$$\int f(y)P^{m^n}(x,t,dy) \to \int f(y)P^m(x,t,dy) \tag{4.3}$$

uniformly in $f(\cdot)$, $x \in G$, $t \in [t_0,t_1]$. *For any* $t > 0$, $P^m(x,t,\cdot)$ *is absolutely continuous with respect to Lebesgue measure, uniformly in the relaxed feedback control* $m(\cdot)$ *and in* $(x,t) \in G\times[t_0,t_1]$. *For each relaxed feedback control* $m(\cdot)$, *the process* $x(\cdot)$ *defined in* (4.2) *is a strong Feller process. The stochastic differential equation has a unique weak-sense solution for each initial condition.*

Outline of part of proof. Concentrate on the uniformity in x. Expression (4.3) can be written as (P_x-measure used when writing E_x)

$$E_x f(x(t))R(t,m^n) - E_x f(x(t))R(t,m) \to 0. \tag{4.4}$$

For notational simplicity only, let $\sigma(x) = I$. Use the inequalities

$$|e^a - e^b| \le |a-b||e^a + e^b|,$$

$$E_x \left| \int_0^t b'_m(x(s))dw(s) - \int_0^t b'_{m^n}(x(s))dw(s)\right|^2$$
$$\le E_x \int_0^t |b_m(x(s)) - b_{m^n}(x(s))|^2 ds.$$

By the continuity and boundedness of $b(\cdot)$ and the weak convergence of the $m^n(y,\cdot)$,

$$b_{m^n}(y) = \int_{\mathcal{U}} b(y,\alpha)m^n(y,d\alpha) \to b_m(y) = \int_{\mathcal{U}} b(y,\alpha)m(y,d\alpha)$$

for almost all y. Define $\tilde{b}_n(y) = |b_m(y) - b_{m^n}(y)|^2$. By Egorov's theorem, for each $\epsilon > 0$ there is a measurable set A_ϵ with $l(A_\epsilon) \le \epsilon$ such that $\tilde{b}_n(y) \to 0$ uniformly in $y \notin A_\epsilon$. Furthermore, $P(x,t,\cdot)$ is absolutely continuous with respect to Lebesgue measure for each x and $t > 0$ (and uniformly in $(x,t) \in G\times[t_0,t_1]$ for any $0 < t_0 < t_1 < \infty$). These facts imply that $\int_0^t E_x\tilde{b}_n(x(s))ds \to 0$, uniformly in $x \in G$. The last expression implies the uniformity of convergence of (4.4). □

4.2 A Smoothed Control

Extend the definition of $m(y,\cdot)$ so that it is defined for y in R^r. For $\epsilon > 0$ and relaxed feedback control $m(\cdot)$, define the smoothed control

$$m^\epsilon(x,\cdot) = \frac{1}{\sqrt{2\pi\epsilon}} \int_{\mathbb{R}^r} e^{-|y-x|^2/2\epsilon} m(y,\cdot)dy, \quad x \in G.$$

The proof of the next theorem is straightforward.

Theorem 4.2. *$m^\epsilon(\cdot)$ is a relaxed feedback control and $m^\epsilon(x,\cdot) \Rightarrow m(x,\cdot)$ for almost all x, as $\epsilon \to 0$. The function $b_{m^\epsilon}(\cdot)$ is continuous for each $\epsilon > 0$, and $b_{m^\epsilon}(\cdot)$ converges almost everywhere to $b_m(\cdot)$.*

We also have the following important results for the invariant measures $\mu_m(\cdot)$, which exist since $x(\cdot)$ is a Feller process on a compact state space for each relaxed feedback control $m(\cdot)$.

Theorem 4.3. *$\mu_m(\cdot)$ is continuous in the control in that if $m^n(x,\cdot) \Rightarrow m(x,\cdot)$ for almost all x, then for each Borel set A,*

$$\mu_{m^n}(A) \to \mu_m(A). \tag{4.5}$$

Proof. By extracting a weakly convergent subsequence of $\{\mu_{m^n}(\cdot)\}$ and relabeling the indices, we can suppose that there is $\mu(\cdot)$ such that $\mu_{m^n}(\cdot) \Rightarrow \mu(\cdot)$. Let $f(\cdot)$ be bounded and continuous. Then the definition of invariant measure implies that, for $t > 0$,

$$\int \mu_{m^n}(dx) \int f(y) P^{m^n}(x,t,dy) = \int f(x)\mu_{m^n}(dx) \to \int f(x)\mu(dx). \tag{4.6}$$

It follows from Theorem 4.1 that (only measurability of $f(\cdot)$ is needed for this)

$$\int f(y) P^{m^n}(x,t,dy) \to \int f(y) P^m(x,t,dy),$$

uniformly in $x \in G$, and the right-hand side is continuous in $x \in G$ since $x(\cdot)$ is a strong Feller process (see [15] for the strong Feller property) under the relaxed feedback control $m(\cdot)$. Thus, (4.6) implies that

$$\begin{aligned}\int f(x)\mu(dx) &= \lim_n \int \mu_{m^n}(dx) \int f(y) P^m(x,t,dy) \\ &= \int \mu(dx) \int f(y) P^m(x,t,dy). \end{aligned} \tag{4.7}$$

The expression (4.7), the arbitrariness of $f(\cdot)$, and the fact that the invariant measure is unique imply that $\mu(\cdot) = \mu_m(\cdot)$. Now, for a Borel set A, let $f(x) = I_A(x)$. Then the two right-hand terms in (4.7) are still equal, and they equal the limit of the left-hand term in (4.6), which proves (4.5). □

Theorem 4.4. *If $m^n(x,\cdot) \Rightarrow m(x,\cdot)$ for almost all x, then $\gamma(m^n) \to \gamma(m)$.*

4.3 Additional Comments

Fix $m_1(\cdot)$ and maximize over $m_2(\cdot)$. Let $m_2^n(\cdot)$ be a maximizing sequence. Consider $m^n(x, d\alpha)\,dx = m_1(x, d\alpha_1)m_2^n(x, d\alpha_2)\,dx$ and take a weakly convergent subsequence. Write the limit as $m_1(x, d\alpha_1)\tilde{m}_2(x, d\alpha_2)\,dx$ (see the remarks after

Theorem 4.5 for more detail on this product form representation), and write $\tilde{m}_2 = \overline{m}_2(m_1)$. Then $\overline{m}_2(m_1)$ is maximizing in that

$$\sup_{m_2} \gamma(m_1, m_2) = \gamma(m_1, \overline{m}_2(m_1)).$$

Now, by minimizing over m_1 we obtain an optimal relaxed feedback policy for Player 1 (if it chooses first), called $\overline{m}_1(\cdot)$. The analogous result holds in the other direction.

The value of the cost for the maximizing control problem is not increased by using relaxed controls, as stated in the following theorem. The analog for the minimization problem with $m_2(\cdot)$ fixed also holds.

Theorem 4.5. *[13, Theorem 4.6.1] Fix $m_1(\cdot)$ and let $\overline{m}_2(m_1)$ be an optimal relaxed feedback control and $r_2(\cdot)$ an admissible relaxed control. Then for each x, $\gamma(x, m_1, r_2) \le \gamma(m_1, \overline{m}_2(m_1))$.*

4.4 Comment on the Product Form Representation of the Limit

Write the limit as $m(x, d\alpha_1, d\alpha_2)\, dx$ and define $m_2(x, d\alpha_2) = m(x, U_1, d\alpha_2)$. Let $f_i(\cdot)$ be continuous. By the convergence assumption,

$$\int_{\mathcal{U}_2} f_2(\alpha_2) m_2^n(x, d\alpha_2) \to \int_{\mathcal{U}_2} f_2(\alpha_2) m_2(x, d\alpha_2)$$

for almost all x. Thus, for almost all x,

$$\left[\int_{\mathcal{U}_1} f_1(\alpha_1) m_1(x, d\alpha_1)\right]\left[\int_{\mathcal{U}_2} f_2(\alpha_2) m_2^n(x, d\alpha_2)\right]$$
$$\to \int_{\mathcal{U}_1}\int_{\mathcal{U}_2} f_1(\alpha_1) f_2(\alpha_2) m_1(x, d\alpha_1) m_2(x, d\alpha_2).$$

Hence

$$\int_G \int_{\mathcal{U}_1}\int_{\mathcal{U}_2} f_1(\alpha_1) f_2(\alpha_2) f_3(x) m_1(x, d\alpha_1) m_2^n(x, d\alpha_2)\, dx$$
$$\to \int_G \int_{\mathcal{U}_1}\int_{\mathcal{U}_2} f_1(\alpha_1) f_2(\alpha_2) f_3(x) m_1(x, d\alpha_1) m_2(x, d\alpha_2)\, dx,$$

which implies that the limit measure has the product form $m_1(x, d\alpha_1) m_2(x, d\alpha_2)\, dx$.

5 Convergence of the Numerical Procedure

For the model (2.3), let Player 1 choose first, with $m_1^\epsilon(\cdot)$ being a smooth ϵ-optimal control. This can be constructed by smoothing the optimal control and appealing to Theorems 4.2–4.4. Now, apply $m_1^\epsilon(\cdot)$ to the finite state chain ξ_n^h (equivalently, to $\psi^h(\cdot)$). This can be done as follows. For $\xi_n^h = x \in G^h$, Player 1 uses $m_1^\epsilon(x, \cdot)$. If $m_1^\epsilon(\cdot)$ is a classical feedback control, the application is clear. If it is a relaxed feedback control, then select the control for the chain by randomizing according to the measure $m_1^\epsilon(x, \cdot)$. It can be shown that the errors due to the randomization have vanishing effect as $h \to 0$. Let $\tilde{u}_2^h(\cdot)$ be the consequent optimal policy for Player 2. This is allowed to depend on m_1^ϵ and the past history of control actions and states, but not on the current realization of the randomization associated with m_1^ϵ, if any. In fact, the optimal policy for Player 2 can be taken to be feedback and not randomized. Then

$$\bar{\gamma}^{+,h} \leq \sup_{u_2^h} \gamma^h(m_1^\epsilon, u_2^h) = \gamma^h\big(m_1^\epsilon, \tilde{u}_2^h\big).$$

There is an admissible relaxed control $\tilde{r}_2$ (see the comments in the next paragraph) such that

$$\limsup_h \bar{\gamma}^{+,h} \leq \gamma\big(m_1^\epsilon, \tilde{r}_2\big) \leq \sup_{r_2} \gamma\big(m_1^\epsilon, r_2\big) \leq \bar{\gamma}^+ + \epsilon. \tag{5.1}$$

There is a similar argument on the other side:

$$\begin{aligned}\bar{\gamma}^{-,h} &\geq \inf_{u_1^h} \gamma^h\big(u_1^h, m_2^\epsilon\big) = \gamma^h\big(\tilde{u}_1^h, m_2^\epsilon\big)\\ &\to \gamma\big(\tilde{r}_1, m_2^\epsilon\big) \geq \inf_{r_1} \gamma\big(r_1, m_2^\epsilon\big) \geq \bar{\gamma}^- - \epsilon.\end{aligned}$$

Hence, since ϵ is arbitrarily small and m_i is optimal for player i if it selects first,

$$\bar{\gamma}^- \leq \liminf_h \bar{\gamma}^{-,h} \leq \limsup_h \bar{\gamma}^{+,h} \leq \bar{\gamma}^+. \tag{5.2}$$

Thus, convergence holds if (the proof of which is outlined in Section 6)

$$\bar{\gamma}^+ = \bar{\gamma}^-.$$

So we must show that there is the asserted relaxed control $\tilde{r}_1$, and a value to the original game when the first player's policies are restricted to relaxed feedback. The first issue is addressed in the next section and the second in the following paragraph.

5.1 Existence of the Relaxed Controls $\tilde{r}_i(\cdot)$ in (5.1)

Consider the game for the approximating chain. Let Player 1 use $m_1^\epsilon(\cdot)$, and let the feedback control $\tilde{u}_2^h(\cdot)$ be optimal for Player 2. Let ξ_n^h denote the chain under these controls, with $\psi^h(\cdot)$ the continuous-time interpolation. For the purpose of computing the cost we can suppose that it is stationary, without loss of generality.

Thus, $\psi^h(0)$ has the stationary measure. Let $\tilde{r}_2^h(\cdot)$ denote the relaxed control representation of $\tilde{u}_2^h(\psi^h(\cdot))$, and $r^h(\cdot)$ the relaxed control representation of the pair of controls for the continuous-time chain. Then the measure of the increments

$$\psi^h(t+\cdot), \tilde{r}_2^h(t+\cdot) - \tilde{r}_2^h(t), z^h(t+\cdot) - z^h(t), y^h(t+\cdot) - y^h(t)$$

does not depend on t, and neither does the weak sense limit. [That is, the limit process $x(\cdot)$ is stationary.] That this $\tilde{r}_2(\cdot)$ can be used in (5.1) follows from the following computation. By the weak convergence and stationarity of the processes (E^{m_1,u_2} is the expectation of functions of the stationary chain under controls m_1, u_2),

$$\begin{aligned}
\gamma^h(m_1^\epsilon, \tilde{u}_2^h) &= E^{m_1^\epsilon, \tilde{u}_2^h} \int_0^1 [k(\psi^h(s), \tilde{u}^h(\psi^h(s))\, ds + c' dy^h(s)] \\
&= E^{m_1^\epsilon, \tilde{u}_2^h} \int_0^1 \left[\int_{\mathcal{U}} k(\psi^h(s), \alpha) \tilde{r}_s^h(d\alpha)\, ds + c' dy^h(s) \right] \\
&= E^{m_1^\epsilon, \tilde{u}_2^h} \int_0^1 \left[\int_{\mathcal{U}_1} k_1(\psi^h(s), \alpha_1) m_1^\epsilon(\psi^h(s), d\alpha_1)\, ds \right. \\
&\qquad \left. + \int_{\mathcal{U}_2} [k_2(\psi^h(s), \alpha_2) \tilde{r}_{2,s}^h(d\alpha_2)\, ds + c' dy^h(s) \right] \\
&\to E \int_0^1 \left[\int_{\mathcal{U}_1} k_1(x(s), \alpha_1) m_1^\epsilon(x(s), d\alpha_1)\, ds \right. \\
&\qquad \left. + \int_{\mathcal{U}_2} [k_2(x(s), \alpha_2) \tilde{r}_{2,s}(d\alpha_2)\, ds + c' dy(s) \right] \\
&= \lim_T \frac{1}{T} E \int_0^T \left[\int_{\mathcal{U}_1} k_1(x(s), \alpha_1) m_1^\epsilon(x(s), d\alpha_1)\, ds \right. \\
&\qquad \left. + \int_{\mathcal{U}_2} [k_2(x(s), \alpha_2) \tilde{r}_{2,s}(d\alpha_2)\, ds + c' dy(s) \right] \\
&= \gamma\big(m_1^\epsilon, \tilde{r}_2\big).
\end{aligned}$$

6 Existence of the Value: Outline of the Argument

The procedure uses an approximation to the process (2.3). We start with a formal discussion to introduce the general idea and the problems to be contended with. For approximation parameter h, let the optimal controls for the approximating Markov chain be feedback with relaxed feedback control representation $\overline{m}_1^h, \overline{m}_2^h$, and suppose that (which holds for the Markov chain in Section 3, under (3.2))

$$\bar{\gamma}^{+,h} = \bar{\gamma}^{-,h} = \bar{\gamma}^h = \gamma^h(\overline{m}_1^h, \overline{m}_2^h). \tag{6.1}$$

Recall that

$$\bar{\gamma}^{-} \leq \liminf_{h} \bar{\gamma}^{-,h} \leq \limsup_{h} \bar{\gamma}^{+,h} \leq \bar{\gamma}^{+}. \tag{6.2}$$

Suppose that there are relaxed feedback controls $\tilde{m}_i$ such that

$$\overline{m}_1^h(x, d\alpha_1)\overline{m}_2^h(x, d\alpha_2)\, dx \Rightarrow \tilde{m}_1(x, d\alpha_1)\tilde{m}_2(x, d\alpha_2)\, dx. \tag{6.3}$$

Suppose that, if for a sequence $m^h(\cdot)$ there is $m(\cdot)$ such that

$$m_1^h(x, d\alpha_1)m_2^h(x, d\alpha_2)\, dx \Rightarrow m_1(x, d\alpha_1)m_2(x, d\alpha_2)\, dx,$$

then

$$\gamma^h(m^h) \to \gamma(m_1, m_2). \tag{6.4}$$

Then (6.2) and (6.3) imply that $\bar{\gamma}^{-} \leq \gamma(\tilde{m}_1, \tilde{m}_2) \leq \bar{\gamma}^{+}$.

Suppose that $\tilde{m}_1$ is not optimal for Player 1 if it chooses first in that $\sup_{m_2} \gamma(\tilde{m}_1, m_2) > \bar{\gamma}^{+}$. Then there is an $\hat{m}_2$ such that $\gamma(\tilde{m}_1, \hat{m}_2) > \bar{\gamma}^{+}$. Now, smooth $\hat{m}_2$ as described before Theorem 4.2 (call the smoothed form $\hat{m}_2^{\epsilon}$), and adapt it for use on the chain (or equivalently the process $\psi^h(\cdot)$). Then, by (6.3), $\overline{m}_1^h(x, d\alpha_1)\hat{m}_2^{\epsilon}(x, d\alpha_2)dx \Rightarrow \tilde{m}_1(x, d\alpha_1)\hat{m}_2^{\epsilon}(x, d\alpha_2)dx$, and by (6.4) $\gamma^h(\overline{m}_1^h, \hat{m}_2^{\epsilon}) \to \gamma(\tilde{m}_1, \hat{m}_2^{\epsilon}) > \bar{\gamma}^{+}$. This implies that $\gamma^h(\bar{m}_1^h, \hat{m}_2^{\epsilon}) > \bar{\gamma}^{h,+}$ for small h, which contradicts the optimality of $\bar{m}_1^h, \bar{m}_2^h$. Hence,

$$\sup_{m_2} \gamma(\tilde{m}_1, m_2) = \bar{\gamma}^{+} = \gamma(\tilde{m}_1, \tilde{m}_2).$$

Do the same on the other side. Conclude that if there is a value to the numerical game and the controls and costs converge, as above, then there is a value to the original game, when the controls of the player that goes first are restricted to relaxed feedback. Thus we need to show (6.1), (6.3), (6.4). Unfortunately, this cannot be done for the finite-state Markov chain numerical procedure. The main difficulty is the existence of the $\tilde{m}_i$ in the weak-sense limit in (6.3), and a primary reason for the problem is that ξ_n^h is concentrated on a grid and does not have a density that is absolutely continuous with respect to Lebesgue measure.

An alternative approximation. To carry out the outlined argument, we can use any alternative approximation which converges in the sense that (6.1)–(6.4) hold. Only a brief outline of the several steps will be given. This alternative approximation is not to be used for numerical purposes, only to show that the value exists, and it will be constructed by discretizing time, and not space. Let $\Delta > 0$ be the time discretization interval. We need to construct processes that have transition densities that are mutually absolutely continuous with respect to Lebesgue measure, uniformly in (x, Δ, $[t_0, t_1]$, control) for any $0 < t_1 < t_2 < \infty$.

Choose the controls that are to be used in the appropriate order (depending on whether we are computing inf sup or sup inf) at $t = 0$, the start of the procedure.

Then in each interval, only one player will exercise its control, as follows. At each time $k\Delta$, $k = 0, 1, \ldots$, flip a fair coin. With probability 1/2, Player 1 will act during that interval. Otherwise, Player 2 will act. The controls will depend on the state at the start of the interval and be constant during the interval, with the appropriate associated drifts. The process will have the form

$$dx = b^{\Delta}(x, u(x^{\Delta}))\,dt + \sigma(x)\,dw + dz, \tag{6.5}$$

where $x^{\Delta}(t) = x(l\Delta)$, $t \in [l\Delta, l\Delta+\Delta)$. As was done for the chain below (3.5) (or for the original process (2.3)), we can write (modulo an error that goes to zero as $h \to 0$ due to the approximation of the d_i near the boundary according to A2) $z(t) = \sum_i d_i y_i(t)$. The value of $b^{\Delta}(\cdot)$ is determined by the randomization procedure. In particular, at $t \in [l\Delta, l\Delta + \Delta)$, $b^{\Delta}(x(t), u(x^{\Delta}(t))) = 2b_i(x(t), u_i(x^{\Delta}(t)))$, for $i = 1$ or $i = 2$ according to the random choice made at $l\Delta$.

If the control is relaxed feedback, then in lieu of (6.5) write

$$dx = b^{\Delta}(x, m(x^{\Delta}))\,dt + \sigma(x)\,dw + dz, \tag{6.6}$$

where at $t \in [l\Delta, l\Delta+\Delta)$, $b^{\Delta}(x(t), m(x^{\Delta}(t))) = 2\int_{\mathcal{U}_i} b_i(x(t), \alpha_i)m_i(x(l\Delta), d\alpha_i)$ for $i = 1$ or $i = 2$ according to the random choice made at $l\Delta$.

Let $E^{\Delta,i,\alpha_i}_{x(l\Delta)}$ denote the expectation of functionals on $[l\Delta, l\Delta + \Delta)$ when player i acts on that interval and uses control action α_i. The conditional mean increment in the total cost on $[l\Delta, l\Delta + \Delta)$ is, for $u_i^{\Delta}(x(l\Delta)) = \alpha_i$, $i = 1, 2$,

$$\begin{aligned} &C^{\Delta}(x(l\Delta), \alpha) \\ &\quad = \frac{1}{2}\sum_{i=1,2} E^{\Delta,i,\alpha_i}_{x(l\Delta)}\left[\int_{l\Delta}^{l\Delta+\Delta} 2k_i(x(s), \alpha_i)ds + c'(y(l\Delta + \Delta) - y(l\Delta))\right], \end{aligned}$$

which is continuous in $x(l\Delta)$, α.

The randomization separates the dynamics: The formal Isaacs equation for the value of the discrete-time problem is

$$\begin{aligned} \bar{\gamma}^{\Delta} + \bar{g}^{\Delta}(x) = \inf_{\alpha_1}\sup_{\alpha_2}\Bigg[&\frac{1}{2}\int \bar{g}^{\Delta}(x + y)p_1^{\Delta}(x, dy|\alpha_1) \\ &+\frac{1}{2}\int \bar{g}^{\Delta}(x + y)p_2^{\Delta}(x, dy|\alpha_2) + C^{\Delta}(x, \alpha)\Bigg], \end{aligned} \tag{6.7}$$

where p_i^{Δ} is the one-step transition function when player i acts on the interval, $\bar{g}^{\Delta}(\cdot)$ is the relative value or potential function, and $\bar{\gamma}^{\Delta}$ is the value. Note that $\sup\inf = \inf\sup$ in (6.7).

The uncontrolled system is

$$dx = \sigma(x)\,dw + dz, \tag{6.8}$$

and (6.5) or (6.6) is obtained from this by a Girsanov transformation, as follows. Define the measure $P_{x,t}^{\Delta,m}$ on $\mathcal{F}_t$ by

$$\begin{aligned}\zeta^\Delta(t,m) &= \int_0^t b^\Delta(x(s),m(x^\Delta(s)))\sigma^{-1}(x(s))\,dw(s)\\ &\quad -\frac{1}{2}\int_0^t |\sigma^{-1}(x(s))b^\Delta(x(s),m(x^\Delta(s)))|^2,\\ R^\Delta(t,m)) &= e^{\zeta^\Delta(t,m)},\quad dP_{x,t}^{\Delta,m} = R^\Delta(t,m)dP_{x,t}.\end{aligned}$$

Let $P_x^{\Delta,m}$ denote the extension of $P_{x,t}^{\Delta,m}$ to $\mathcal{F}$. Then the process $w(\cdot)$ defined by

$$dw_m^\Delta(t) = dw(t) - [\sigma^{-1}(x(t))b^\Delta(x(s),m(x^\Delta(s)))]dt$$

is an $\mathcal{F}_t$-standard Wiener process on $(\Omega, P_x^{\Delta,m}, \mathcal{F})$. Now, rewrite the model (6.8) in terms of the new Wiener process as

$$dx(t) = b^\Delta(x(t),m(x^\Delta(t)))\,dt + \sigma(x(t))\,dw_m^\Delta(t) + dz(t),$$

which yields the model (6.6).

The basic properties of (2.3) carry over. In particular, the process defined by (6.6) has the strong Feller property when sampled at $l\Delta$, $l > 0$, and $P^\Delta(x, dy, l\Delta|m)$, the transition probability at $t = l\Delta$, is mutually absolutely continuous with respect to Lebesgue measure, uniformly in $m(\cdot)$, x, $[t_0, t_1]$ for any $0 < t_0 < t_1 < \infty$ that are multiples of Δ. Hence the Doeblin condition is satisfied and for each $m(\cdot)$ there is a unique invariant measure $\mu_m^\Delta(\cdot)$ that is uniformly (in $m(\cdot)$) absolutely continuous with respect to Lebesgue measure. There are controls $\overline{m}_i^{\Delta,+}(\cdot)$, $i = 1, 2$, which are optimal if Player 1 chooses its control first (i.e., for the upper value), and $\overline{m}_i^{\Delta,-}(\cdot)$, $i = 1, 2$, which are optimal if Player 2 chooses its control first (i.e., for the lower value).

Given $m(\cdot)$, define the mean value

$$\gamma^\Delta(m) = \int_G \mu_m^\Delta(dx)\left[\int_{\mathcal{U}} C^\Delta(x,\alpha)m(x,d\alpha)\right],$$

and the potential function

$$g^\Delta(x,m) = \sum_{l=0}^{\infty}\big[E_x^{\Delta,m}C^\Delta(x(l\Delta), m(x(l\Delta)) - \gamma^\Delta(m))\big].$$

The summands converge to zero exponentially, uniformly in $(x, m(\cdot))$. Then, a direct evaluation yields the relationship

$$\bar\gamma^{\Delta,\pm} + g^{\Delta,\pm}(x) = E_x^{\Delta,\overline{m}^{\Delta,\pm}}\big[g^{\Delta,\pm}(x(\Delta)) + C^\Delta(x,\overline{m}^{\Delta,\pm}(x))\big], \tag{6.9}$$

where $\bar\gamma^{\Delta,\pm}$, $g^{\Delta,\pm}(x)$ correspond to $\overline{m}^{\Delta,\pm}(\cdot)$.

Next, using (6.9) and the definitions of $\bar{\gamma}^{\Delta,\pm}$, we show that the Isaacs equation (6.7) holds for almost all x, and that $\bar{\gamma}^{+,\Delta} = \bar{\gamma}^{-,\Delta} = \bar{\gamma}^{\Delta}$. Finally, given a sequence $m^{\Delta}(\cdot)$, there is always a subsequence (indexed also by $\Delta \to 0$ for notational simplicity) and relaxed feedback controls $\tilde{m}_i$ such that

$$m_1^{\Delta}(x, d\alpha_1)m_2^{\Delta}(x, d\alpha_2)dx \Rightarrow m_1(x, d\alpha_1)m_2(x, d\alpha_2)dx.$$

We then show that this implies that

$$\gamma^{\Delta}(m^{\Delta}) \to \gamma(m).$$

The details for the assertions in the last sentence are similar to what one does for the original continuous-time single player system. Additional details on these points are given in Section 4 and [15, Chapter 4].

Acknowledgments

The work of this author was supported by contracts DAAD-19-02-1-0425 from the Army Research Office and National Science Foundation Grant ECS 0097447.

References

[1] M. Bardi, S. Bottacini, and M. Falcone. Convergence of discrete schemes for discontinuous value functions of pursuit-evasion games. In G.I. Olsder, editor, *New Trends in Dynamic Games and Applications*. Birkhäuser, Boston, 1995.

[2] M. Bardi, M. Falcone, and P. Soravia. Fully discrete schemes for the value function of pursuit-evasion games. In T. Basar and A. Haurie, editors, *Advances in Stochastic Games and Applications*. Birkhäuser, Boston, 1994.

[3] M. Bardi, M. Falcone, and P. Soravia. Numerical methods for pursuit-evasion games via viscosity solutions. In M. Bardi, T.E.S. Raghavan, and T. Parthasarathy, editors, *Stochastic and Differential Games: Theory and Numerical Methods*. Birkhäuser, Boston, 1998.

[4] M. Bardi and P. Soravia. Approximation of differential games of pursuit-evasion by discrete time games. In R.P. Hämäläinen and H.K. Ehtamo, editors, *Differential Games: Developments in Modeling and Computation*. Springer-Verlag, Berlin and New York, 1991.

[5] P. Billingsley. *Convergence of Probability Measures; second edition*. Wiley, New York, 1999.

[6] P. Dupuis and H. Ishii. On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stochastics Stochastic Rep.*, 35:31–62, 1991.

[7] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.

[8] W.H. Fleming and W. McEneaney. Risk-sensitive control on an infinite time horizon. *SIAM J. Control Optim.*, 33:1881–1815, 1995.

[9] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, 1988.

[10] H. J. Kushner. The Gauss-Seidel numerical procedure for Markov stochastic games. *IEEE Trans in Automatic Control*, 49:1779–1782, 2004.

[11] H.J. Kushner. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. Academic Press, New York, 1977.

[12] H.J. Kushner. Numerical methods for stochastic control problems in continuous time. *SIAM J. Control Optim.*, 28:999–1048, 1990.

[13] H.J. Kushner. *Heavy Traffic Analysis of Controlled Queueing and Communication Networks*. Springer-Verlag, Berlin and New York, 2001.

[14] H.J. Kushner. Numerical methods for stochastic differential games. *SIAM J. Control Optim.*, 41:457–486, 2002.

[15] H.J. Kushner. Numerical approximations for stochastic differential games: The ergodic case. *SIAM J. Control Optim.*, 42:1911–1933, 2004.

[16] H.J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, Berlin and New York, 1992. Second edition, 2001.

[17] O. Pourtallier and M. Tidball. Approximation of the value function for a class of differential games with target. Research report 2942, INRIA, France, 1996.

[18] O. Pourtallier and B. Tolwinski. Discretization of Isaac's equation. Report: INRIA, France, 1992.

[19] M. Tidball. Undiscounted zero-sum differential games with stopping times. In G.J. Olsder, editor, *New Trends in Dynamic Games and Applications*. Birkhäuser, Boston, 1995.

[20] M. Tidball and R.L.V. González. Zero-sum differential games with stopping times: Some results and about its numerical resolution. In T. Basar and A. Haurie, editors, *Advances in Dynamic Games and Applications*. Birkhäuser, Boston, 1994.

Gradient Transformation Trajectory Following Algorithms for Determining Stationary Min-Max Saddle Points

Walter J. Grantham
School of Mechanical and Materials Engineering
Washington State University
Pullman, WA 99164-2920, USA
grantham@mme.wsu.edu

Abstract

For finding a stationary min-max point of a scalar-valued function, we develop and investigate a family of gradient transformation differential equation algorithms. This family includes, as special cases: Min-Max Ascent, Newton's method, and a Gradient Enhanced Min-Max (GEMM) algorithm that we develop. We apply these methods to a sharp-spined "Stingray" saddle function, in which Min-Max Ascent is globally asymptotically stable but stiff, and Newton's method is not stiff, but does not yield global asymptotic stability. However, GEMM is both globally asymptotically stable and not stiff. Using the Stingray function we study the stiffness of the gradient transformation family in terms of Lyapunov exponent time histories. Starting from points where Min-Max Ascent, Newton's method, and the GEMM method do work, we show that Min-Max Ascent is very stiff. However, Newton's method is not stiff and is approximately 60 to 440 times as fast as Min-Max Ascent. In contrast, the GEMM method is globally convergent, is not stiff, and is approximately 3 times faster than Newton's method and approximately 175 to 1000 times faster than Min-Max Ascent.

Key words. Min-max saddle points, stationary points, trajectory following differential equations, stiff systems, Lyapunov exponents.

1 Introduction

We consider the problem of finding a stationary point $\mathbf{x}^* \in R^s$ for a scalar-valued function $\phi(\mathbf{x})$. We are particularly concerned with finding a min-max saddle point. In this chapter we develop differential equation algorithms of the form $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, where $(\dot{\ })$ denotes $\mathrm{d}()/\mathrm{d}t$. We choose the function $\mathbf{f}(\cdot)$ with the objective of having solutions $\mathbf{x}(t) \to \mathbf{x}^*$ as $t \to \infty$. Such "trajectory following" algorithms have received considerable attention in recent years. In [15] steepest descent differential equations are used to design controllers for nonlinear systems. In [5] optimal control differential equations are used to design new discrete minimization algorithms. In [1] and [12] differential equation algorithms are investigated for min-max optimization problems. In [6] and [7] differential equations for Newton's method are used to find all of the stationary points of a function. In [8] a Gradient Enhanced Newton (GEN) algorithm is developed for finding a stationary proper minimum point.

In this chapter we develop and investigate a family of "gradient transformation" differential equation algorithms for finding a stationary min-max point. This family includes, as special cases: Min-Max Ascent [1,12], Newton's method, and a Gradient Enhanced Min-Max (GEMM) algorithm that we develop. We apply these algorithms to a sharp-spined "Stingray" saddle function, a min-max analogue to Rosenbrock's "banana" function, used for testing minimization algorithms [8]. We are particularly concerned with eliminating the stiff differential equations that can result from Min-Max Ascent. We study the stiffness of the gradient transformation family in terms of Lyapunov exponent time histories. Applied to the Stingray function, we show that Newton's method does not yield global asymptotic stability. We also show that Newton's method (from an initial point where it does work) is not stiff and is approximately 60 to 440 times as fast as Min-Max Ascent. But we show that the GEMM method is also not stiff, is globally convergent, and is approximately 3 times as fast as Newton's method and approximately 175 to 1000 times as fast as Min-Max Ascent.

In Section 2 we define min-max points and present necessary conditions for a min-max. In Section 3 we present the "Stingray" saddle function and examine some of its properties. In Section 4 we define Lyapunov exponents, for investigating stiffness. In Section 5 we develop the family of gradient transformation algorithms. We apply Min-Max Ascent and Newton's method to the Stingray saddle function. In Section 6 we develop and analyze the GEMM method and apply it to the Stingray saddle function. In Section 7 we present numerical results comparing Min-Max Ascent, Newton's method, and the GEMM method. In Section 8 we present the conclusions.

2 Min-Max Saddle Point

Let $\mathbf{M}^{\mathsf{T}}$ denote the transpose of a matrix $\mathbf{M}$. For $\mathbf{x}^{\mathsf{T}} = [\mathbf{u}^{\mathsf{T}}, \mathbf{v}^{\mathsf{T}}]$, with $\mathbf{u} \in \mathcal{U} \subseteq R^n$, $\mathbf{v} \in \mathcal{V} \subseteq R^m$, and $\mathbf{x} \in R^s$, $s = n + m$, we are concerned with finding a point $\mathbf{x}^* = (\mathbf{u}^*, \mathbf{v}^*)$ to yield a min-max for a C^2 scalar-valued function $\phi(\mathbf{x}) = \phi(\mathbf{u}, \mathbf{v})$, such that $\mathbf{u}^*$ minimizes ϕ and $\mathbf{v}^*$ maximizes ϕ. That is,

$$\phi(\mathbf{u}^*, \mathbf{v}) \le \phi(\mathbf{u}^*, \mathbf{v}^*) \le \phi(\mathbf{u}, \mathbf{v}^*)$$

for all $\mathbf{u} \in \mathcal{U}$ and $\mathbf{v} \in \mathcal{V}$. Denote the gradient of ϕ by

$$\mathbf{g} = \left[\frac{\partial \phi}{\partial \mathbf{x}}\right]^{\mathsf{T}} = \begin{bmatrix} \mathbf{g}_u \\ \mathbf{g}_v \end{bmatrix} = \begin{bmatrix} \left[\frac{\partial \phi}{\partial \mathbf{u}}\right]^{\mathsf{T}} \\ \left[\frac{\partial \phi}{\partial \mathbf{v}}\right]^{\mathsf{T}} \end{bmatrix}$$

and the Hessian of ϕ by

$$\mathbf{G} = \frac{\partial^2 \phi}{\partial \mathbf{x}^2} = \begin{bmatrix} \mathbf{G}_{uu} & \mathbf{G}_{uv} \\ \mathbf{G}_{uv}^{\mathsf{T}} & \mathbf{G}_{vv} \end{bmatrix},$$

where $\mathbf{g}_u \in R^n$, $\mathbf{g}_v \in R^m$, $\mathbf{G}_{uu} = \partial^2\phi/\partial \mathbf{u}^2 \in R^{n\times n}$, $\mathbf{G}_{vv} = \partial^2\phi/\partial \mathbf{v}^2 \in R^{m\times m}$, and $\mathbf{G}_{uv} = \partial^2\phi/\partial \mathbf{u}\partial \mathbf{v} \in R^{n\times m}$.

We are particularly concerned with finding a **proper stationary min-max point** $\mathbf{x}^*$, at which:

(1) $\phi(\mathbf{u}^*, \mathbf{v}) < \phi(\mathbf{u}^*, \mathbf{v}^*) < \phi(\mathbf{u}, \mathbf{v}^*)$ for all $\mathbf{u} \in \mathcal{U} - \{\mathbf{u}^*\}$ and $\mathbf{v} \in \mathcal{V} - \{\mathbf{v}^*\}$,
(2) $\mathbf{g}^* = \mathbf{g}(\mathbf{x}^*) = \mathbf{0}$,
(3) $\mathbf{G}_{uu}^* = \mathbf{G}_{uu}(\mathbf{x}^*) \geq 0$ (positive semidefinite),
(4) $\mathbf{G}_{vv}^* = \mathbf{G}_{vv}(\mathbf{x}) \leq 0$ (negative semidefinite),
(5) $|\mathbf{G}^*| = |\mathbf{G}(\mathbf{x}^*)| < 0$.

In addition we assume that $\mathbf{g}(\mathbf{x}) \neq \mathbf{0}$ for $\mathbf{x} \neq \mathbf{x}^*$ and that $\|\mathbf{g}(\mathbf{x})\| \to \infty$ as $\|\mathbf{x} - \mathbf{x}^*\| \to \infty$, where $\|\cdot\|$ denotes the Euclidian norm.

For $\mathbf{u} \in \mathcal{U}$ and $\mathbf{v} \in \mathcal{V}$ denote the rational reaction set for the minimizing player $\mathbf{u}$ by $\mathcal{R}_u = \{(\mathbf{u}, \mathbf{v}) : \mathbf{v} \in \mathcal{V}$ and $\phi(\mathbf{u}, \mathbf{v}) \leq \phi(\bar{\mathbf{u}}, \mathbf{v})$ for all $\bar{\mathbf{u}} \in \mathcal{U}\}$ and let $\mathcal{R}_v = \{(\mathbf{u}, \mathbf{v}) : \mathbf{u} \in \mathcal{U}$ and $\phi(\mathbf{u}, \bar{\mathbf{v}}) \leq \phi(\mathbf{u}, \mathbf{v})$ for all $\bar{\mathbf{v}} \in \mathcal{V}\}$ denote the rational reaction set for the maximizing player $\mathbf{v}$. On $\mathcal{R}_u$ with $\mathbf{u} \in \overset{o}{\mathcal{U}}$ (interior of $\mathcal{U}$) it is necessary [13, p. 149] that $\mathbf{0} = \mathbf{g}_u(\mathbf{u}, \mathbf{v}) = [\partial\phi(\mathbf{u}, \mathbf{v})/\partial \mathbf{u}]^{\mathsf{T}}$ and $\mathbf{G}_{uu}(\mathbf{u}, \mathbf{v}) = \partial^2\phi(\mathbf{u}, \mathbf{v})/\partial \mathbf{u}^2 \geq 0$. On $\mathcal{R}_v$ with $\mathbf{v} \in \overset{o}{\mathcal{V}}$ it is necessary that $\mathbf{0} = \mathbf{g}_v(\mathbf{u}, \mathbf{v}) = [\partial\phi(\mathbf{u}, \mathbf{v})/\partial \mathbf{v}]^{\mathsf{T}}$ and $\mathbf{G}_{vv}(\mathbf{u}, \mathbf{v}) = \partial^2\phi(\mathbf{u}, \mathbf{v})/\partial \mathbf{v}^2 \leq 0$.

3 Stingray Saddle Function

For $a > 0$ and $c > 0$ we consider the "Stingray" saddle function

$$\phi = \frac{a}{2}u^2 + \frac{c}{2}(u-1)v^2,$$

with gradient

$$\mathbf{g} = \begin{bmatrix} g_u \\ g_v \end{bmatrix} = \begin{bmatrix} au + \frac{c}{2}v^2 \\ c(u-1)v \end{bmatrix}$$

and Hessian

$$\mathbf{G} = \begin{bmatrix} G_{uu} & G_{uv} \\ G_{uv} & G_{vv} \end{bmatrix} = \begin{bmatrix} a & cv \\ cv & c(u-1) \end{bmatrix}.$$

The function has a unique proper min-max point at $\mathbf{x}^* = (u^*, v^*) = (0, 0)$, with $\mathbf{g} \neq \mathbf{0}$ for $\mathbf{x} \neq \mathbf{0}$ and $\|\mathbf{g}\| \to \infty$ as $\|\mathbf{x}\| \to \infty$. Note that $|\mathbf{G}| = ac(u-1) - c^2v^2 = 0$ on $u = 1 + \frac{c}{a}v^2$. Also note that $G_{uu} = a > 0$ for all (u, v), but $G_{vv} = c(u-1) < 0$ only for $u < 1$. The Stingray function $\phi(u, v)$ is convex in u for each v, but is concave in v only for $u < 1$. For $u > 1$ the function is convex in v.

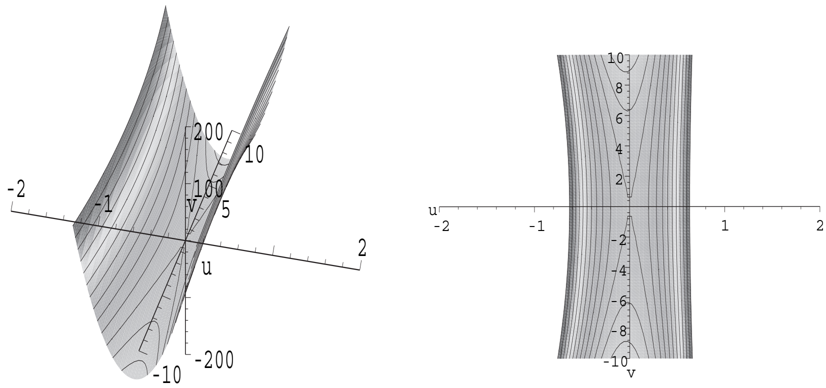


Figure 1: Banana saddle ($a = 1000$, $c = 1$).

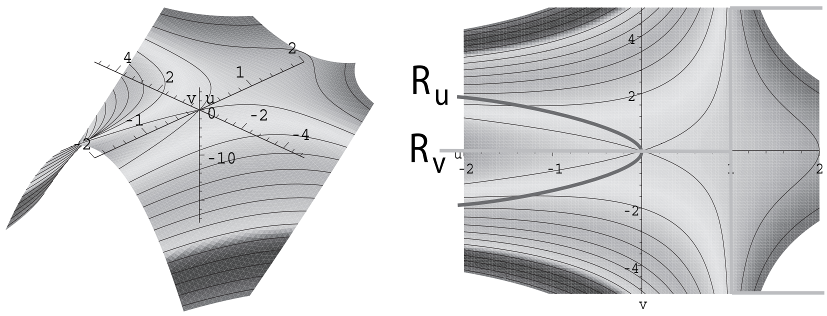


Figure 2: Stingray saddle ($a = 1$, $c = 1$).

Figures 1 and 2 show three-dimensional and contour plots for various values of a and c. For $a = 1000$ and $c = 1$ the function is similar to Rosenbrock's "banana" function, having a steep-walled canyon with a parabolic valley, except that the stationary point is a saddle point instead of a minimum point. For $a = 1$ and increasing values of c the function looks like a stingray flapping its wings. Unless otherwise specified, we will consider the case where $a = 1$ and $c = 100$. For these parameter values the Stingray function has a sharp local $\max_v$ ridge on $v = 0$, $u < 1$, and a local $\min_u$ valley on $u = -\frac{c}{2a}v^2$.

As illustrated in Figure 2, for $v_{\min} \le v \le v_{\max}$ with $v_{\min} < 0 < v_{\max}$, the $\min_u \max_v \phi$ rational reaction sets are $\mathcal{R}_u = \{(u, v) : u = -\frac{c}{2a}v^2\}$ and $\mathcal{R}_v = \{(u, v) : v = v^\circ(u)\}$, where

$$v^\circ(u) = \begin{cases} 0 & \text{if } u < 1 \\ \in [v_{\min}, v_{\max}] & \text{if } u = 1 \\ v_{\max} & \text{if } u > 1 \text{ and } v_{\max} \ge |v_{\min}| \\ v_{\min} & \text{if } u > 1 \text{ and } v_{\max} \le |v_{\min}| \,. \end{cases}$$

In particular, while the minimizing player u seeks $g_u = 0$, the maximizing player v only seeks $g_v = 0$ for $u < 1$. For $u > 1$ the maximizing player seeks either the upper or lower bound on v. Nevertheless, the intersection $\mathcal{R}_u \cap \mathcal{R}_v$ of the reaction sets is the min-max point $u^* = v^* = 0$, where both $g_u = 0$ and $g_v = 0$.

4 Lyapunov Exponents

Stiff systems are systems which have two or more widely separated time scales, usually specified in terms of eigenvalues. For nonlinear systems we will use Lyapunov exponents [14, p. 205]. Lyapunov exponents are generalizations of eigenvalues and characteristic (Floquet) multipliers that provide information about the (average) rates at which neighboring trajectories converge or diverge in a nonlinear system. Let $\mathbf{x}(t)$ and $\tilde{\mathbf{x}}(t)$ be solutions to

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{1}$$

starting from neighboring initial conditions, and let $\rho(t) = \|\tilde{\mathbf{x}}(t) - \mathbf{x}(t)\|$ be the distance between the trajectory $\mathbf{x}(t)$ and the perturbed trajectory $\tilde{\mathbf{x}}(t)$ at time t. If $\rho(0)$ is arbitrarily small and $\rho(t) \to \rho(0)e^{\sigma t}$ as $t \to \infty$ then σ is called a Lyapunov exponent for the reference trajectory $\mathbf{x}(t)$. The distance between the trajectory points $\mathbf{x}(t)$ and $\tilde{\mathbf{x}}(t)$ grows, shrinks, or remains constant for $\sigma > 0$, $\sigma < 0$, or $\sigma = 0$, respectively. In an s-dimensional state space there are s real Lyapunov exponents, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_s$, corresponding to exponential growth rates in s orthogonal directions. For a given trajectory $\mathbf{x}(t)$ the Lyapunov exponents are unique, but are functions of the initial state. Arbitrarily close initial states (e.g., on and to either side of a separatrix) may yield trajectories with different Lyapunov exponents, corresponding to different behaviors as $t \to \infty$.

If $\mathbf{f}(\cdot)$ is continuous and continuously differentiable the Lyapunov exponents can be computed [16] in terms of the state perturbation equations

$$\dot{\boldsymbol{\eta}} = \mathbf{A}(t)\boldsymbol{\eta}, \quad \mathbf{A}(t) = \frac{\partial \mathbf{f}[\mathbf{x}(t)]}{\partial \mathbf{x}}, \tag{2}$$

where, with $\boldsymbol{\eta}(t) \in R^s$, $\mathbf{A}(t) \in R^{s\times s}$ is evaluated along a trajectory $\mathbf{x}(t)$ and, for small ϵ, $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) + \epsilon\boldsymbol{\eta}(t) + \mathbf{O}(\epsilon^2)$ is an initially neighboring trajectory. If $\mathbf{f}(\cdot)$ is discontinuous across some "switching surface" in the state space certain "jump conditions" must be imposed to accurately compute the Lyapunov exponents [9].

At $t > 0$ we define the instantaneous Lyapunov exponents as

$$\sigma_i(t) = \frac{1}{t} \ln\left[\frac{\|\boldsymbol{\eta}_i(t)\|}{\|\boldsymbol{\eta}_i(0)\|}\right], \tag{3}$$

with the Lyapunov exponents $\sigma_i = \lim_{t\to\infty}\{\sigma_i(t)\}$. For the special case of an equilibrium $\mathbf{x}(t) = \textit{constant}$, so that $\mathbf{A}$ is constant, the Lyapunov exponents σ_i, $i = 1, \ldots, s$, are the real parts of the eigenvalues μ_i of $\mathbf{A}$. The same result holds for trajectories that asymptotically approach an equilibrium.

5 Gradient Transformation Systems

We will consider the class of systems

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) = -\mathbf{P}(\mathbf{x})\mathbf{g}, \tag{4}$$

where the $s \times s$ matrix $\mathbf{P}(\mathbf{x})$ is to be chosen. If $|\mathbf{P}(\mathbf{x})| \neq 0$ for all $\mathbf{x}$ then the only equilibrium $\mathbf{x}^*$ is where $\mathbf{g}(\mathbf{x}) = \mathbf{0}$. We wish to choose $\mathbf{P}(\mathbf{x})$ to control the eigenvalues of $\mathbf{P}(\mathbf{x}^*)$, to control the stiffness of (4), and to make $\mathbf{x}^*$ (globally) asymptotically stable.

5.1 State Perturbation Equations

Let $\mathbf{p}_i^\mathsf{T}(\mathbf{x})$, $i = 1, \ldots, s$, denote the ith row of $\mathbf{P}(\mathbf{x})$. Then along a trajectory $\mathbf{x}(t)$ the state perturbation equations are given by (2) with $\mathbf{A}(\mathbf{x}) = \partial \mathbf{f}(\mathbf{x})/\partial \mathbf{x}$ given by

$$\mathbf{A}(\mathbf{x}) = -\mathbf{P}(\mathbf{x})\mathbf{G}(\mathbf{x}) - \begin{bmatrix} \mathbf{g}^\mathsf{T}(\mathbf{x}) \frac{\partial \mathbf{p}_1(\mathbf{x})}{\partial \mathbf{x}} \\ \vdots \\ \mathbf{g}^\mathsf{T}(\mathbf{x}) \frac{\partial \mathbf{p}_s(\mathbf{x})}{\partial \mathbf{x}} \end{bmatrix}. \tag{5}$$

At a stationary point $\mathbf{x}^*$ of $\phi(\mathbf{x})$, $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$ and $\mathbf{A}(\mathbf{x}^*) = -\mathbf{P}(\mathbf{x}^*)\mathbf{G}(\mathbf{x}^*)$. This result also holds for $\mathbf{P}$ constant.

The separation of the eigenvalues μ_i of $\mathbf{A}(\mathbf{x}^*)$ indicates the stiffness of the system, at least near equilibrium. Note that this stiffness cannot be altered merely by multiplying the state equations (1) by a scalar $\beta(\mathbf{x})$, such as $1/(1+\|\mathbf{f}(\mathbf{x})\|)$. Specifically, $\dot{\mathbf{x}} = \beta \mathbf{f}$ would have a state perturbation matrix $\bar{\mathbf{A}} = \beta \mathbf{A} + \mathbf{f}[\partial \beta/\partial \mathbf{x}] \to \beta(\mathbf{x}^*)\mathbf{A}(\mathbf{x}^*)$ as $\mathbf{f} \to \mathbf{0}$, which has eigenvalues $\bar{\mu}_i = \beta(\mathbf{x}^*)\mu_i$. Thus widely separated μ_i yield widely separated $\bar{\mu}_i$, with $\beta(\mathbf{x}^*)$ as a scale factor.

5.2 Steepest Descent

Taking $\mathbf{P}(\mathbf{x}) = \mathbf{I}$ yields the **Steepest Descent** algorithm

$$\dot{\mathbf{x}} = -\mathbf{g}(\mathbf{x}). \tag{6}$$

When applied to *minimize* a function $\phi(\mathbf{x})$, such as Rosenbrock's function, Steepest Descent yields global asymptotic stability of the equilibrium at $\mathbf{x}^*$. However, the dynamic system may be stiff. In [8] we investigate gradient transformation systems applied to minimization problems. Global asymptotic stability is established by considering $\dot{\phi} = \mathbf{g}^\mathsf{T}\dot{\mathbf{x}} = -\mathbf{g}^\mathsf{T}\mathbf{g} < 0$ for $\mathbf{g} \neq \mathbf{0}$. If $V(\mathbf{x}) = \phi(\mathbf{x}) - \phi(\mathbf{x}^*) > 0$ and $\nabla V = [\partial V/\partial \mathbf{x}]^\mathsf{T} = \mathbf{g}(\mathbf{x}) \neq \mathbf{0}$ for $\mathbf{x} \neq \mathbf{x}^*$ then $V(\mathbf{x})$ is a Lyapunov function [14, p. 217] and $\mathbf{x}^*$ is asymptotically stable. In addition, if $\|\mathbf{g}(\mathbf{x})\| \to \infty$ as $\|\mathbf{x} - \mathbf{x}^*\| \to \infty$ then $\mathbf{x}^*$ is globally asymptotically stable.

When applied to finding a min-max saddle point, $\phi(\mathbf{x})$ would not be a candidate for a Lyapunov function, since the stationary point is a saddle point, not a proper minimal point [$\phi(\mathbf{x}) = \textit{constant}$ contours do not surround $\mathbf{x}^*$]. This makes the search for a suitable Lyapunov function very difficult. Nevertheless, the concept of steepest descent (or ascent) can be applied to the saddle point problem.

5.3 Min-Max Ascent

Since $\mathbf{u}$ seeks $\min_u \phi(\mathbf{u}, \mathbf{v})$ and $\mathbf{v}$ seeks $\max_v \phi(\mathbf{u}, \mathbf{v})$, the first min-max algorithm investigated by researchers [1] was steepest descent on $\mathbf{u}$ and steepest ascent on $\mathbf{v}$.

Let $\mathbf{I}_u$ and $\mathbf{I}_v$ denote the $n \times n$ and $m \times m$ identity matrices, respectively. Taking $\mathbf{P}(\mathbf{x}) = \text{diag}[\mathbf{I}_u, -\mathbf{I}_v]$ yields the **Min-Max Ascent** algorithm

$$\begin{aligned} \dot{\mathbf{u}} &= -\mathbf{g}_u \\ \dot{\mathbf{v}} &= \quad \mathbf{g}_v \end{aligned} \tag{7}$$

with the state perturbation equations

$$\begin{bmatrix} \dot{\boldsymbol{\eta}}_u \\ \dot{\boldsymbol{\eta}}_v \end{bmatrix} = \begin{bmatrix} -\mathbf{G}_{uu} & -\mathbf{G}_{uv} \\ \mathbf{G}_{uv}^{\mathsf{T}} & \mathbf{G}_{vv} \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_u \\ \boldsymbol{\eta}_v \end{bmatrix}.$$

For the Stingray saddle function $\phi = \frac{a}{2}u^2 + \frac{c}{2}(u-1)v^2$ the Min-Max Ascent system is given by

$$\begin{aligned} \dot{u} &= -g_u = -au - \frac{c}{2}v^2 \\ \dot{v} &= g_v = c(u-1)v, \end{aligned}$$

with the state perturbation equations

$$\begin{bmatrix} \dot{\boldsymbol{\eta}}_u \\ \dot{\boldsymbol{\eta}}_v \end{bmatrix} = \begin{bmatrix} -a & -cv \\ cv & c(u-1) \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_u \\ \boldsymbol{\eta}_v \end{bmatrix}.$$

At the stationary point the state perturbation matrix $\mathbf{A}(\mathbf{x}^*) = \text{diag}[-a, -c]$ has eigenvalues $\{-a, -c\}$. For $a = 1$ and $c = 100$ Min-Max Ascent yields a very stiff system.

Figure 3 shows Min-Max Ascent trajectories for the case where $a = 1$ and $c = 100$. For numerical integration we use a fixed step size ($\Delta t = 10^{-5}$, because of stiffness) standard $4th$-order Runge–Kutta method. Trajectories for $u < 1$ rapidly approach the $v = 0$ ($g_v = 0$) surface (the sharp local maximum ridge of the Stingray) and then slowly move along the ridge toward the saddle point at the origin. This is caused by the stiffness of the system. Notice the tendency, in the region $u > 1$, for trajectories to diverge from the $g_v = 0$ surface rather than converge to it. This is caused by G_{vv} not being negative definite everywhere.

In [12] a sufficiency condition for Min-Max Ascent is presented. It does not hold for the Stingray system, but a variation can be used to establish global asymptotic stability. The positive definite function $V = \frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{x}$ has Lyapunov derivative $\dot{V} = \mathbf{x}^{\mathsf{T}}\dot{\mathbf{x}} = -au^2 + \frac{c}{2}uv^2 - cv^2 = -\mathbf{x}^{\mathsf{T}}\mathbf{Q}\mathbf{x}$, with

$$\mathbf{Q} = \begin{bmatrix} a & -\frac{c}{4}v \\ -\frac{c}{4}v & c \end{bmatrix}$$

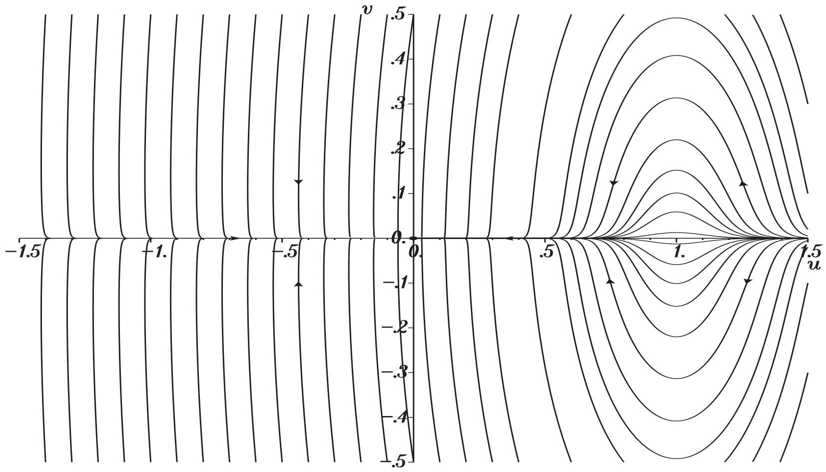


Figure 3: Min-Max Ascent ($a = 1$, $c = 100$).

and $|\mathbf{Q}| = ac - c^2v^2/16$. Thus $\dot{V} \le 0$ only in the region $v^2 < 16a/c$, where $\mathbf{Q}$ is positive definite. Thus, V is a local Lyapunov function, establishing local asymptotic stability of the equilibrium at $\mathbf{x}^* = (u^*, v^*) = (0, 0)$.

To establish global asymptotic stability we use a three-stage Lyapunov approach. The function $V_1 = u$ yields $\dot{V}_1 = \dot{u} = -au - \frac{c}{2}v^2 < 0$ for all (u, v) such that $u > 0$. In particular, every trajectory starting in the region $V_1(u, v) = u > 0$ eventually enters the positively invariant region $u < 1$ and remains there. In that region consider $V_2 = \frac{1}{2}v^2$. Then $\dot{V}_2 = v\dot{v} = c(u-1)v^2 < 0$ for $u < 1$ and $v \neq 0$. Hence for arbitrary $\epsilon > 0$ every trajectory eventually enters the positively invariant region $\mathcal{D}_\epsilon = \{(u, v) : u \le 1, |v| \le \epsilon\}$. In $\mathcal{D}_\epsilon$ consider $V_3 = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{x} = \frac{1}{2}(u^2 + v^2)$, with $\dot{V}_3 = -\mathbf{x}^\mathsf{T}\mathbf{Q}\mathbf{x} < 0$ for $(u, v) \in \mathcal{D}_\epsilon - \{\mathbf{0}\}$ for $\epsilon < 4\sqrt{\frac{a}{c}}$. We conclude that the Min-Max Ascent equilibrium at $\mathbf{x}^* = (u^*, v^*) = (0, 0)$ is globally asymptotically stable.

5.4 Newton's Method

Newton's method, in which $\mathrm{d}\mathbf{g}/\mathrm{d}t = -\mathbf{g}$ [hence, $\mathbf{g}(t) = \mathbf{g}(0)\mathrm{e}^{-t} \to \mathbf{0}$ as $t \to \infty$], corresponds to $\mathbf{P}(\mathbf{x}) = \mathbf{G}^{-1}(\mathbf{x})$. Applied to the Stingray saddle function, Newton's method is given by

$$\begin{aligned} \dot{\mathbf{x}} &= -\mathbf{G}^{-1}\mathbf{g} \\ &= -\frac{c}{|\mathbf{G}|}\begin{bmatrix} (u-1)\left(au + \frac{1}{2}cv^2\right) - cv^2(u-1) \\ -v\left(au + \frac{1}{2}cv^2\right) + a(u-1)v \end{bmatrix}, \end{aligned} \tag{8}$$

where $|\mathbf{G}| = ac(u-1) - c^2v^2$. Figure 4 shows trajectories for Newton's method applied to the Stingray saddle function ($a = 1, c = 100$) using a $4th$-order Runge–Kutta method ($\Delta t = 10^{-3}$).

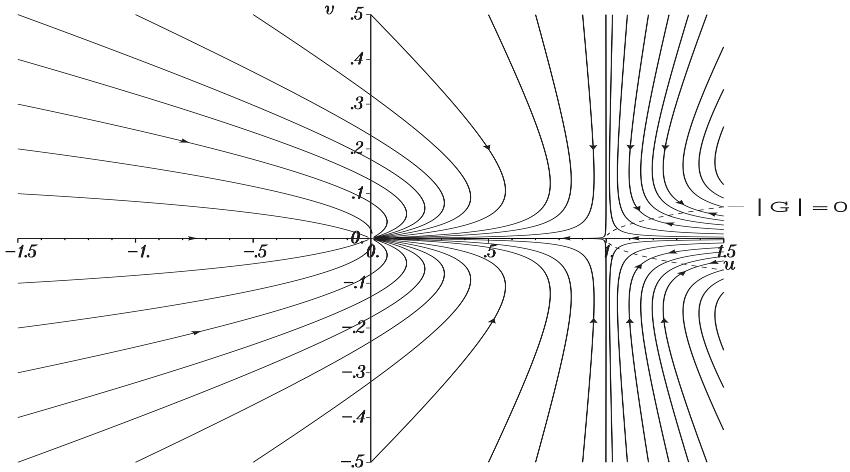


Figure 4: Newton's method ($a = 1$, $c = 100$).

At $\mathbf{x}^* = (u^*, v^*) = (0, 0)$ the state perturbation equations yield $\mathbf{A}(\mathbf{x}^*) = \text{diag}[-1, -1]$ with eigenvalues $\{-1, -1\}$. This is clearly not a stiff system near $\mathbf{x}^*$. Trajectories move at a much better speed than in Min-Max Ascent, as indicated by the step size. However, Newton's method is not globally asymptotically stable to $\mathbf{x}^*$. Note that solutions to (8) only exist for $|\mathbf{G}| \neq 0$ and that $|\mathbf{G}| = 0$ on $v^2 = (u-1)a/c$. The domain of attraction to $\mathbf{x}^*$ is only the region $u < 1$, that is, the region where $\mathbf{G}_{vv} < 0$.

5.5 Hamiltonian Systems

As with Newton's method, Min-Max Ascent may not always yield global asymptotic stability. Min-Max Ascent may in fact produce Hamiltonian systems, of the form (for $m = n$) $\dot{\mathbf{u}} = [\partial H/\partial \mathbf{v}]^\mathsf{T}$, $\dot{\mathbf{v}} = -[\partial H/\partial \mathbf{u}]^\mathsf{T}$ for some Hamiltonian $H(\mathbf{u}, \mathbf{v})$. Then $\dot{H} = [\partial H/\partial \mathbf{u}]\dot{\mathbf{u}} + [\partial H/\partial \mathbf{v}]\dot{\mathbf{v}} \equiv 0$. Hence $H = \textit{constant}$. Such systems cannot have asymptotically stable equilibria. For more details see [2].

5.5.1 Example (Hyperbolic Saddle)

Consider the function [12] $\phi = uv$, which has a (nonproper) $\min_u \max_v$ saddle at $\mathbf{x}^* = (u^*, v^*) = (0, 0)$. The gradient is $\mathbf{g} = \begin{bmatrix} v & u \end{bmatrix}^\mathsf{T}$ and the Hessian is

$$\mathbf{G} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The Min-Max Ascent system $\dot{u} = -\partial\phi/\partial u = -v$, $\dot{v} = \partial\phi/\partial v = u$ has trajectories that are circles centered at the origin. The difficulty is that this Min-Max

Ascent system is a Hamiltonian system $\dot{u} = \partial H/\partial v$, $\dot{v} = -\partial H/\partial u$, with Hamiltonian $H(u, v) = -(u^2 + v^2)/2$.

Despite the lack of even local asymptotic stability for the Min-Max Ascent system in this example, there does exist a gradient transformation system that yields global asymptotic stability. Consider the Newton system $\dot{\mathbf{x}} = -\mathbf{G}^{-1}\mathbf{g}$ given by $[\dot{u} \quad \dot{v}]^\mathsf{T} = [-u \quad -v]^\mathsf{T}$, which yields $[u \quad v]^\mathsf{T} = [u(0) \quad v(0)]^\mathsf{T}\mathrm{e}^{-t}$. Thus $\mathbf{x}(t) \to \mathbf{x}^* = [0 \quad 0]^\mathsf{T}$ as $t \to \infty$. Since $|\mathbf{G}(\mathbf{x})| \neq 0$ anywhere, the domain of attraction is all of R^2.

5.6 Gradient Enhanced Newton Minimization

Consider, for a moment, Newton's method applied to the problem of finding a proper *minimum* point for a function $\phi(\mathbf{x})$. For the case where $\mathbf{G}(\mathbf{x})$ is not positive definite everywhere, the Levenberg–Marquardt modification to Newton's method [4, pp. 145–149] is given by $(\alpha\mathbf{I} + \mathbf{G})\dot{\mathbf{x}} = -\mathbf{g}$, where $\alpha \geq 0$ and $\mathbf{I}$ denotes the $s \times s$ identity matrix. If $\mathbf{F} = \alpha\mathbf{I} + \mathbf{G}$ is positive definite then $\dot{\mathbf{x}} = -\mathbf{P}(\mathbf{x})\mathbf{g}$, with $\mathbf{P}(\mathbf{x}) = \mathbf{F}^{-1} = (\alpha\mathbf{I} + \mathbf{G})^{-1}$. Then $\dot{\phi} = \mathbf{g}^\mathsf{T}\dot{\mathbf{x}} = -\mathbf{g}^\mathsf{T}\mathbf{P}\mathbf{g} < 0$ for $\mathbf{g} \neq \mathbf{0}$ establishes (global) asymptotic stability.

Let μ_i and $\boldsymbol{\xi}_i$, $i = 1, \ldots, s$, denote the eigenvalues and eigenvectors of $\mathbf{G}$, respectively. For symmetric $\mathbf{G}$ the eigenvalues are all real, but may not all be positive. The matrix $\mathbf{F} = \alpha\mathbf{I} + \mathbf{G}$ has eigenvalues $\lambda_i = \mu_i + \alpha$ and eigenvectors $\boldsymbol{\xi}_i$, since $\mathbf{F}\boldsymbol{\xi}_i = (\mu_i + \alpha)\boldsymbol{\xi}_i$. Thus, at a point $\mathbf{x}$, if α is sufficiently large, all of the eigenvalues of $\mathbf{F}$ will be positive. As $\alpha \to 0$ the method approaches Newton's method applied to $\phi(\mathbf{x})$, and as $\alpha \to \infty$ the method approaches Steepest Descent applied to $\phi(\mathbf{x})/\alpha$.

The Levenberg–Marquardt minimization method generally will not work with constant α. If $|\mathbf{G}(\mathbf{x})|$ changes sign somewhere, then for constant α the determinant $|\mathbf{F}| = |\alpha\mathbf{I} + \mathbf{G}|$ will also generally change sign, although at a different place than $|\mathbf{G}(\mathbf{x})|$.

In [8] we develop a **Gradient Enhanced Newton** (GEN) minimization method, in which $\alpha = \gamma\|\mathbf{g}\| = \gamma\sqrt{\mathbf{g}^\mathsf{T}\mathbf{g}}$ with constant $\gamma \geq 0$, yielding

$$\dot{\mathbf{x}} = -\mathbf{P}(\mathbf{x})\mathbf{g} = -[\gamma\|\mathbf{g}\|\mathbf{I} + \mathbf{G}]^{-1}\mathbf{g}. \tag{9}$$

The ideas behind this minimization method are: 1) at points where $\|\mathbf{g}\| \neq 0$ we can make $\mathbf{F}$ be positive definite for sufficiently large $\gamma \geq 0$; 2) for small γ or near places where $\mathbf{g} = \mathbf{0}$ the method behaves like Newton's method; 3) for large $\|\mathbf{g}\|$ the speed $\|\dot{\mathbf{x}}\| \approx 1/\gamma$. In [8] it is shown that, for sufficiently large $\gamma \geq 0$, GEN is globally asymptotically stable for functions that have a single proper stationary minimum point and satisfy a Lyapunov growth condition. In addition, when applied to Rosenbrock's "banana" function, GEN is uniformly nonstiff and approximately 25 times faster than Newton's method and approximately 2500 times faster than Steepest Descent.

6 Gradient Enhanced Min-Max

The Levenberg–Marquardt modification of Newton's method cannot be used for min-max problems, but a variation of it can. Consider the Hessian

$$\mathbf{G}(\mathbf{x}) = \begin{bmatrix} \mathbf{G}_{uu} & \mathbf{G}_{uv} \\ \mathbf{G}_{uv}^{\mathsf{T}} & \mathbf{G}_{vv} \end{bmatrix},$$

which is positive definite at a proper minimum point. But at a proper min-max point $\mathbf{x}^*$, $\mathbf{G}_{uu}^* = \mathbf{G}_{uu}(\mathbf{x}^*) \geq 0$, $\mathbf{G}_{vv}^* = \mathbf{G}_{vv}(\mathbf{x}^*) \leq 0$, and $|\mathbf{G}^*| = |\mathbf{G}(\mathbf{x}^*)| < 0$. Thus, the eigenvalues of $\mathbf{G}_{uu}^*$ are ≥ 0, the eigenvalues of $\mathbf{G}_{vv}^*$ are ≤ 0, and the product of the eigenvalues of $\mathbf{G}^*$ is negative. When $|\mathbf{G}(\mathbf{x})|$ passes through zero, so does one or more of its eigenvalues. The Levenberg–Marquardt matrix $\mathbf{F} = \alpha\mathbf{I} + \mathbf{G}$ could be used to make all of its eigenvalues be positive (or all of them negative, for $\alpha < 0$) at any given point $\hat{\mathbf{x}}$. But if $\alpha = \alpha(\mathbf{x}) \geq 0$ (or ≤ 0), with $\alpha(\mathbf{x}^*) = 0$ and $|\mathbf{G}^*| = |\mathbf{G}(\mathbf{x}^*)| < 0$, then somewhere between $\mathbf{x}^*$ and $\hat{\mathbf{x}}$ we would have $|\mathbf{F}(\mathbf{x})| = 0$, as one of the positive eigenvalues goes negative or one of the negative eigenvalues goes positive. What we need to do, to ensure that the replacement matrix $\mathbf{F}(\mathbf{x})$ for $\mathbf{G}(\mathbf{x})$ is nonsingular, is to keep the positive eigenvalues positive and the negative eigenvalues negative, yielding $|\mathbf{F}(\mathbf{x})| < 0$.

Consider

$$\dot{\mathbf{x}} = -\mathbf{P}\mathbf{g}, \tag{10}$$

with

$$\mathbf{P} = \mathbf{F}^{-1} = \begin{bmatrix} \alpha_u \mathbf{I}_u + \mathbf{G}_{uu} & \mathbf{G}_{uv} \\ \mathbf{G}_{uv}^{\mathsf{T}} & -\alpha_v \mathbf{I}_v + \mathbf{G}_{vv} \end{bmatrix}^{-1}. \tag{11}$$

For $\alpha_u = \alpha_v = \alpha \to \infty$ the method approaches Min-Max Ascent applied to ϕ/α. For $\alpha \to 0$ the method approaches Newton's method applied to ϕ. The **Gradient Enhanced Min-Max** (GEMM) method is given by (10)–(11) with $\alpha_u = \gamma_u \|\mathbf{g}\|$ and $\alpha_v = \gamma_v \|\mathbf{g}\|$ for constants $\gamma_u \geq 0$ and $\gamma_v \geq 0$. That is, $\mathbf{P} = \mathbf{F}^{-1}$, with

$$\mathbf{F} = \begin{bmatrix} \gamma_u \|\mathbf{g}\| \mathbf{I}_u + \mathbf{G}_{uu} & \mathbf{G}_{uv} \\ \mathbf{G}_{uv}^{\mathsf{T}} & -\gamma_v \|\mathbf{g}\| \mathbf{I}_v + \mathbf{G}_{vv} \end{bmatrix}. \tag{12}$$

As for establishing global asymptotic stability of the resulting equilibrium, we note that using $W(\mathbf{x}) = \mathbf{g}^{\mathsf{T}}\mathbf{g}$ as a descent function [14, p. 276] would not work, since $\dot{W} = \mathbf{g}^{\mathsf{T}}\dot{\mathbf{g}} + \dot{\mathbf{g}}^{\mathsf{T}}\mathbf{g} = \mathbf{g}^{\mathsf{T}}\mathbf{G}\dot{\mathbf{x}} + \dot{\mathbf{x}}^{\mathsf{T}}\mathbf{G}\mathbf{g} = -\mathbf{g}^{\mathsf{T}}\mathbf{Q}\mathbf{g}$, with $\mathbf{Q} = \mathbf{G}\mathbf{P} + \mathbf{P}^{\mathsf{T}}\mathbf{G}$ not being positive definite if $|\mathbf{G}|$ changes sign (see Lyapunov's lemma, [14, p. 223]).

Also note that replacing the $\min_u \max_v \phi$ problem with Newton's method (or the Levenberg–Marquardt modification) applied to the least squares problem [4, pp. 146–148] of minimizing $W(\mathbf{x}) = \mathbf{g}^{\mathsf{T}}\mathbf{g}$, via $\dot{\mathbf{x}} = -\mathbf{H}^{-1}(\mathbf{x})\nabla W$, where $\nabla W = [\partial W/\partial \mathbf{x}]^{\mathsf{T}}$ and $\mathbf{H}(\mathbf{x}) = \partial^2 W/\partial \mathbf{x}^2$, would involve third derivatives of $\phi(\mathbf{x})$.

6.1 Nonsingularity of the Partitioned Matrix F

We will show that for sufficiently large constants $\gamma_u \geq 0$ and $\gamma_v \geq 0$ the matrix $\mathbf{F}$ in (12) is nonsingular for all $\mathbf{x}$. Hence the only equilibrium for (10)–(12) is at $\mathbf{x}^*$. A Lyapunov approach can be used to investigate whether the unique equilibrium at $\mathbf{x}^*$ is (globally) asymptotically stable.

To prove that $\mathbf{F}$ is nonsingular, we have the following results.

Lemma 6.1. *For $\mathbf{x} \in R^s$ let $\mathbf{M}(\mathbf{x})$ be an $n \times n$ matrix whose elements are functions of class C^q, $q \geq 0$, in a neighborhood of $\hat{\mathbf{x}} \in R^s$, with distinct eigenvalues at $\hat{\mathbf{x}}$. Then the eigenvalues $\lambda_j(\mathbf{x})$, $j = 1, \ldots, n$, of $\mathbf{M}(\mathbf{x})$ are of class C^q in a neighborhood of $\hat{\mathbf{x}}$.*

Proof. Consider the characteristic equation $0 = \psi(\lambda, \mathbf{x}) = |\lambda\mathbf{I} - \mathbf{M}(\mathbf{x})| = \lambda^n + p_{n-1}\lambda^{n-1} + \cdots + p_1\lambda + p_0$, where $\mathbf{I}$ denotes the $n \times n$ identity matrix. The coefficients $p_j(\mathbf{x})$ are C^q since they can be determined from Newton's identities [14, p. 227] in terms of the trace$(\mathbf{M}^j)$, $j = 1, \ldots, n$, of powers of $\mathbf{M}(\mathbf{x})$, which only involves products and sums of the elements of $\mathbf{M}(\mathbf{x})$. Then the lemma follows from the implicit function theorem [13, p. 21], with Jacobian $\mathrm{d}\psi(\lambda_j, \hat{\mathbf{x}})/\mathrm{d}\lambda \neq 0$ for the case where the eigenvalues λ_j, $j = 1, \ldots, n$, are distinct. □

For repeated eigenvalues, the elements of $\mathbf{M}(\mathbf{x})$ can be perturbed by an arbitrarily small amount $\epsilon > 0$ to yield distinct eigenvalues [11, p. 89]. For a more detailed analysis of the case of repeated eigenvalues, see [10, p. 134]. Henceforth, we will consider only the case of distinct eigenvalues.

Theorem 6.1. *For $\mathbf{x} \in R^s$ let $\mathbf{M}(\mathbf{x}) \in R^{n\times n}$ be a continuous symmetric matrix with $\mathbf{M}(\mathbf{x}^*) \geq 0$ (≤ 0) and let $\phi(\mathbf{x})$ be a scalar-valued function of class C^q, $q \geq 1$. Let $\mathbf{g}(\mathbf{x}) = [\partial\phi/\partial\mathbf{x}]^\mathsf{T}$. If $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$, with $\mathbf{g}(\mathbf{x}) \neq \mathbf{0}$ for $\mathbf{x} \neq \mathbf{x}^*$ and $\|\mathbf{g}(\mathbf{x})\| \to \infty$ as $\|\mathbf{x} - \mathbf{x}^*\| \to \infty$, then for $\gamma \geq 0$ (≤ 0) with $|\gamma|$ sufficiently large, the $n \times n$ matrix $\mathbf{N}(\mathbf{x}) = \gamma\|\mathbf{g}(\mathbf{x})\|\mathbf{I} + \mathbf{M}(\mathbf{x})$ is positive definite (negative definite) for all $\mathbf{x} \neq \mathbf{x}^*$.*

Proof. We consider the positive semidefinite case for $\mathbf{M}(\mathbf{x}^*)$. The proof for the negative semidefinite case is analogous. At $\mathbf{x}$ let $\mu(\mathbf{x})$ denote the smallest (possibly negative) eigenvalue of $\mathbf{M}(\mathbf{x})$, with corresponding unit eigenvector $\boldsymbol{\xi}(\mathbf{x})$. For $\gamma \geq 0$ let $\lambda(\mathbf{x}) = \mu(\mathbf{x}) + \gamma\|\mathbf{g}(\mathbf{x})\|$ denote the corresponding smallest eigenvalue of $\mathbf{N}(\mathbf{x})$, with corresponding unit eigenvector $\boldsymbol{\xi}(\mathbf{x})$, where $\lambda(\mathbf{x}) = \boldsymbol{\xi}^\mathsf{T}\mathbf{N}\boldsymbol{\xi} = \lambda\boldsymbol{\xi}^\mathsf{T}\boldsymbol{\xi} = \mu(\mathbf{x}) + \gamma\|\mathbf{g}(\mathbf{x})\|$. Let $\mathcal{B}_r = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^*\| \leq r\}$. From Lemma 6.1 $\mu(\mathbf{x})$ is continuous on R^s, with $\mu(\mathbf{x}^*) \geq 0$ and all the other eigenvalues of $\mathbf{M}(\mathbf{x}^*)$ positive. For arbitrarily small $\epsilon > 0$ let $\bar{\mathbf{x}}$ be a minimal point for $\mu(\mathbf{x})$ on $\mathcal{B}_\epsilon$. If $\bar{\mathbf{x}} = \mathbf{x}^*$ choose any $\bar{\gamma} > 0$. If $\bar{\mathbf{x}} \neq \mathbf{x}^*$ choose $\bar{\gamma} > \max\{0, -\mu(\bar{\mathbf{x}})/\|\mathbf{g}(\bar{\mathbf{x}})\|\}$. Then for $\gamma > \bar{\gamma}$, $\mu(\mathbf{x}) > 0\ \forall\ \mathbf{x} \in \mathcal{B}_\epsilon - \{\mathbf{x}^*\}$. For any $r \geq \epsilon$ let $\mathcal{X}_r = \{\mathbf{x} : \epsilon \leq \|\mathbf{x} - \mathbf{x}^*\| \leq r\}$, with $\|\mathbf{g}(\mathbf{x})\| > 0\ \forall\ \mathbf{x} \in \mathcal{X}_r$. From the theorem of Weierstrass $\mu(\mathbf{x})/\|\mathbf{g}(\mathbf{x})\|$ takes on a minimum value at some point $\hat{\mathbf{x}} \in \mathcal{X}_r$. Let $\hat{\gamma}(r) = \max\{0, -\mu(\hat{\mathbf{x}})/\|\mathbf{g}(\hat{\mathbf{x}})\|\} \geq 0$. Then for $\gamma > \hat{\gamma}(r)$ we have $\lambda(\mathbf{x})/\|\mathbf{g}(\mathbf{x})\| = \gamma + \mu(\mathbf{x})/\|\mathbf{g}(\mathbf{x})\| \geq \gamma + \mu(\hat{\mathbf{x}})/\|\mathbf{g}(\hat{\mathbf{x}})\| \geq \gamma - \hat{\gamma}(r) > 0$

$\forall\, \mathbf{x} \in \mathcal{X}_r$. The conditions on $\mathbf{g}(\mathbf{x})$ ensure that $\|\mathbf{g}(\mathbf{x})\| \not\to 0$ as $\|\mathbf{x}-\mathbf{x}^*\| \to \infty$. Thus $\hat{\gamma} = \lim_{r\to\infty}\{\hat{\gamma}(r)\}$ exists. Then $\lambda(\mathbf{x}) > 0\ \forall\, \mathbf{x} \neq \mathbf{x}^*$ provided $\gamma > \max(\bar{\gamma}, \hat{\gamma})$. □

Lemma 6.2. *Let* $\mathbf{A} \in R^{n\times n}$ *be symmetric and* $\mathbf{B} \in R^{n\times m}$. *If* $\mathbf{A}$ *is positive definite* ($\mathbf{A} > 0$) *then* $\mathbf{B}^\mathsf{T}\mathbf{A}\mathbf{B}$ *is at least positive semidefinite* ($\mathbf{B}^\mathsf{T}\mathbf{A}\mathbf{B} \geq 0$).

Proof. For $\mathbf{z} \in R^m$ and $\mathbf{y} \in R^n$, let $\mathbf{y} = \mathbf{B}\mathbf{z}$. Then $\psi = \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y} > 0$ for $\mathbf{y} \neq \mathbf{0}$. Hence $\psi = \mathbf{z}^\mathsf{T}\mathbf{B}^\mathsf{T}\mathbf{A}\mathbf{B}\mathbf{z} \geq 0$ for $\mathbf{z} \neq \mathbf{0}$. □

Lemma 6.3. *Let* $\mathbf{A} \in R^{n\times n}$ *be symmetric and* $\mathbf{B} \in R^{n\times n}$. *If* $\mathbf{A} > 0$ (< 0) *and* $\mathbf{B} \geq 0$ (≤ 0) *then* $\mathbf{A}+\mathbf{B} > 0$ (< 0).

Proof. For $\mathbf{y} \in R^n$ we have $\psi = \mathbf{y}^\mathsf{T}(\mathbf{A}+\mathbf{B})\mathbf{y} = \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y} + \mathbf{y}^\mathsf{T}\mathbf{B}\mathbf{y} > 0$ (< 0) for all $\mathbf{y} \neq \mathbf{0}$. □

Theorem 6.2. *For* $\mathbf{A} \in R^{n\times n}$ *symmetric,* $\mathbf{B} \in R^{n\times m}$, *and* $\mathbf{D} \in R^{m\times m}$ *symmetric, the matrix*

$$\mathbf{F} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\mathsf{T} & \mathbf{D} \end{bmatrix}$$

is nonsingular, with $|\mathbf{F}| < 0$, *if* $\mathbf{A} > 0$ *and* $\mathbf{D} < 0$ (*or if* $\mathbf{A} < 0$ *and* $\mathbf{D} > 0$).

Proof. Premultiplying the first block row by $\mathbf{B}^\mathsf{T}\mathbf{A}^{-1}$ and subtracting from the second block row yields

$$|\mathbf{F}| = \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{B}^\mathsf{T}\mathbf{A}^{-1}\mathbf{B} \end{vmatrix} = |\mathbf{A}||\mathbf{D} - \mathbf{B}^\mathsf{T}\mathbf{A}^{-1}\mathbf{B}|.$$

For $\mathbf{A} > 0$, we have [3, p. 128] $|\mathbf{A}| = \lambda_1 \cdots \lambda_n > 0$, where λ_j, $j = 1, \ldots, n$, are the eigenvalues of $\mathbf{A}$. From Lemma 6.2 $\mathbf{B}^\mathsf{T}\mathbf{A}^{-1}\mathbf{B} \geq 0$. Thus from Lemma 6.3 with $\mathbf{D} < 0$ we have $\mathbf{D} - \mathbf{B}^\mathsf{T}\mathbf{A}^{-1}\mathbf{B} < 0$. Hence $|\mathbf{D} - \mathbf{B}^\mathsf{T}\mathbf{A}^{-1}\mathbf{B}| = \mu_1 \cdots \mu_m < 0$, where μ_k, $k = 1, \ldots, m$, are the eigenvalues of $\mathbf{D} - \mathbf{B}^\mathsf{T}\mathbf{A}^{-1}\mathbf{B}$. Thus, $|\mathbf{F}| = \lambda_1 \cdots \lambda_n \mu_1 \cdots \mu_m < 0$. □

Theorem 6.3. *If* $\mathbf{x}^* = (\mathbf{u}^*, \mathbf{v}^*)$ *is a proper stationary min-max point for a scalar-valued* C^2 *function* $\phi(\mathbf{u}, \mathbf{v})$, *with* $\mathbf{g}(\mathbf{x}) = [\partial\phi(\mathbf{x})/\partial\mathbf{x}]^\mathsf{T} \neq \mathbf{0}$ *for* $\mathbf{x} \neq \mathbf{x}^*$ *and* $\|\mathbf{g}(\mathbf{x})\| \to \infty$ *as* $\|\mathbf{x} - \mathbf{x}^*\| \to \infty$, *and with*

$$\mathbf{G}(\mathbf{x}) = \frac{\partial^2\phi(\mathbf{x})}{\partial\mathbf{x}^2} = \begin{bmatrix} \mathbf{G}_{uu} & \mathbf{G}_{uv} \\ \mathbf{G}_{uv}^\mathsf{T} & \mathbf{G}_{vv} \end{bmatrix},$$

then for sufficiently large $\gamma_u \geq 0$ *and* $\gamma_v \geq 0$ *the matrix*

$$\mathbf{F} = \begin{bmatrix} \gamma_u\|\mathbf{g}\|\mathbf{I}_u + \mathbf{G}_{uu} & \mathbf{G}_{uv} \\ \mathbf{G}_{uv}^\mathsf{T} & -\gamma_v\|\mathbf{g}\|\mathbf{I}_v + \mathbf{G}_{vv} \end{bmatrix}$$

is nonsingular, with $|\mathbf{F}| < 0$, *for all* $\mathbf{x} = (\mathbf{u}, \mathbf{v})$.

Proof. At $\mathbf{x}^*$ we have $|\mathbf{F}^*| = |\mathbf{G}^*| < 0$. From Theorem 6.1, for $\mathbf{x} \neq \mathbf{x}^*$, $\gamma_u \|\mathbf{g}\| \mathbf{I}_u + \mathbf{G}_{uu} > 0$ for sufficiently large $\gamma_u \geq 0$ and $-\gamma_v \|\mathbf{g}\| \mathbf{I}_v + \mathbf{G}_{vv} < 0$ for sufficiently large $\gamma_v \geq 0$. Then $|\mathbf{F}| < 0$ follows from Theorem 6.2 with $\mathbf{A} = \gamma_u \|\mathbf{g}\| \mathbf{I}_u + \mathbf{G}_{uu}$, $\mathbf{B} = \mathbf{G}_{uv}$, and $\mathbf{D} = -\gamma_v \|\mathbf{g}\| \mathbf{I}_v + \mathbf{G}_{vv}$. $\square$

6.2 Stingray Saddle

For a and $c > 0$ in the Stingray saddle function $\phi = \frac{a}{2}u^2 + \frac{c}{2}(u-1)v^2$ consider

$$\mathbf{F} = \begin{bmatrix} \alpha_u \mathbf{I}_u + \mathbf{G}_{uu} & \mathbf{G}_{uv} \\ \mathbf{G}_{uv}^\mathsf{T} & -\alpha_v \mathbf{I}_v + \mathbf{G}_{vv} \end{bmatrix} = \begin{bmatrix} \alpha_u + a & cv \\ cv & -\alpha_v + c(u-1) \end{bmatrix}.$$

The determinant $|\mathbf{F}| = (\alpha_u + a)[-\alpha_v + c(u-1)] - c^2v^2$ is zero on the parabola $c^2v^2 = (\alpha_u + a)[-\alpha_v + c(u-1)]$ provided $-\alpha_v + c(u-1) \geq 0$. Since $\alpha_u + a > 0$ for all $\alpha_u \geq 0$ with $a > 0$ and $c > 0$, a necessary and sufficient condition for $|\mathbf{F}(u, v)| < 0$ for all u, v is that $-\alpha_v \mathbf{I}_v + \mathbf{G}_{vv} = -\alpha_v + c(u-1) < 0$ for all u. We can ensure that $|\mathbf{F}(u, v)| < 0$ for all u, v by taking $\alpha_v = \gamma_v \|\mathbf{g}\| = \gamma_v \sqrt{(au + \frac{c}{2}v^2)^2 + c^2(u-1)^2v^2}$ with sufficiently large $\gamma_v > 0$. To see this, note that

$$|\mathbf{F}| = -\gamma_v(\alpha_u + a)\sqrt{\left(au + \frac{c}{2}v^2\right)^2 + c^2(u-1)^2v^2} + (\alpha_u + a)c(u-1) - c^2v^2.$$

The $\max_v |\mathbf{F}|$ occurs on $v = 0$, with $|\mathbf{F}|_{v=0} = (\alpha_u + a)[-\gamma_v a|u| + c(u-1)]$. For $u \leq 0$ we have $|\mathbf{F}|_{v=0} < 0$. For $u > 0$ we have $0 = |\mathbf{F}|_{v=0} = (\alpha_u + a)[(c - \gamma_v a)u - c]$ at $u = 1/(1 - \gamma_v \frac{a}{c})$, which yields $u < 0$ (a contradiction) for $\gamma_v > c/a$. Hence $|\mathbf{F}(u, v)| < 0$ for all u, v if we take $\alpha_u = \gamma_u \|\mathbf{g}\|$ and $\alpha_v = \gamma_v \|\mathbf{g}\|$, with $\gamma_u \geq 0$ and $\gamma_v > c/a$. We also have

$$\begin{aligned} -\alpha_v + c(u-1) &= -\gamma_v \sqrt{\left(au + \frac{c}{2}v^2\right)^2 + c^2(u-1)^2v^2} + c(u-1) \\ &< -\gamma_v a|u| + c(u-1) \\ &< 0 \text{ for all } u, v \text{ if } \gamma_v > c/a. \end{aligned} \tag{13}$$

Applied to the Stingray function, the GEMM algorithm is given by

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = -\mathbf{F}^{-1}\mathbf{g} = -\frac{1}{|\mathbf{F}|} \begin{bmatrix} [-\gamma_v \|\mathbf{g}\| + c(u-1)]\left(au + \frac{1}{2}cv^2\right) - c^2v^2(u-1) \\ -cv\left(au + \frac{1}{2}cv^2\right) + (\gamma_u \|\mathbf{g}\| + a)c(u-1)v \end{bmatrix},$$

where $0 > |\mathbf{F}| = (\gamma_u \|\mathbf{g}\| + a)[-\gamma_v \|\mathbf{g}\| + c(u-1)] - c^2v^2$ for all u, v.

For $a = 1$, $c = 100$, and $\gamma_v = 101$, Figures 5 and 6 show trajectories for the GEMM algorithm for $\gamma_u = 0$ and 10, respectively.

To establish global asymptotic stability we use a three-stage Lyapunov approach. For $V_1 = u$ we have

$$\dot{V}_1 = \dot{u} = -\frac{1}{|\mathbf{F}|}\left\{[-\gamma_v \|\mathbf{g}\| + c(u-1)]\left(au + \frac{1}{2}cv^2\right) - c^2v^2(u-1)\right\}$$

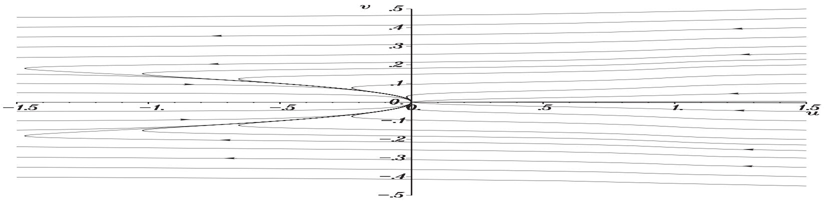


Figure 5: GEMM trajectories ($\gamma_u = 0$, $\gamma_v = 101$).

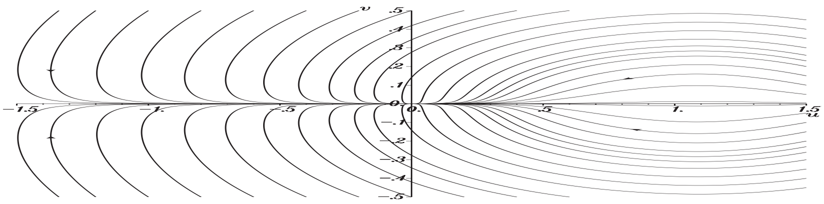


Figure 6: GEMM trajectories ($\gamma_u = 10$, $\gamma_v = 101$).

and from (13) $-\gamma_v\|\mathbf{g}\| + c(u-1) < 0$. Thus $\dot{u} < 0$ for all u, v with $u \geq 1$. Hence every trajectory starting in the region $V_1(u, v) = u \geq 1$ eventually enters the positively invariant region $u < 1$ and remains in that region. For $u \leq 1$ consider the function $V_2 = \frac{1}{2}v^2$, with

$$\begin{aligned}\dot{V}_2 = v\dot{v} &= -\frac{1}{|\mathbf{F}|}cv^2\left\{\gamma_u\|\mathbf{g}\|(u-1) - a - \frac{1}{2}cv^2\right\}\\ &< 0 \quad \text{for } u \leq 1 \text{ and } v \neq 0.\end{aligned}$$

Thus for arbitrary $\epsilon > 0$ every trajectory eventually enters the positively invariant region $\mathcal{D}_\epsilon = \{(u, v) : u \leq 1, |v| \leq \epsilon\}$. In $\mathcal{D}_\epsilon$ consider

$$V_3 = \frac{1}{2}(u^2 + v^2),$$

with

$$\begin{aligned}\dot{V}_3 = &-\frac{1}{|\mathbf{F}|}u\left\{[-\gamma_v\|\mathbf{g}\| + c(u-1)]\left(au + \frac{1}{2}cv^2\right) - c^2v^2(u-1)\right\}\\ &-\frac{1}{|\mathbf{F}|}cv^2\left\{\gamma_u\|\mathbf{g}\|(u-1) - a - \frac{1}{2}cv^2\right\}\\ \approx &-\frac{1}{|\mathbf{F}|}[-\gamma_v\|\mathbf{g}\| + c(u-1)]au^2\\ < &\ 0 \quad \text{for } (u, v) \in \mathcal{D}_\epsilon - \{\mathbf{0}\}.\end{aligned}$$

We conclude that the GEMM equilibrium at $\mathbf{x}^* = (u^*, v^*) = (0, 0)$ is globally asymptotically stable.

7 Performance Comparisons

For comparison of Min-Max Ascent, Newton's method, and the GEMM method, we consider the trajectories starting from $(u, v) = (-1.5, 0.5)$ for the Stingray saddle function. We use a fixed time step standard $4th$-order Runge–Kutta method with the time step Δt chosen to control the approximate initial displacement $\|\dot{\mathbf{x}}(0)\|\Delta t$. The trajectories are terminated when $\|\mathbf{g}\| < 10^{-3}$. We consider two cases: Table 1 shows results for the "Banana saddle" ($a = 1000$, $c = 1$, $\gamma_u = \gamma_v = 1$, stiffness ≈ 1000), and Table 2 shows results for the "Stingray saddle" ($a = 1$, $c = 100$, $\gamma_u = 1$, $\gamma_v = 101$, stiffness ≈ 100). The results indicate that Newton's method is about 60 to 440 times faster than Min-Max Ascent, and that the GEMM method is about 2 to 3 times faster than Newton's method and about 175 to 1000 times faster than Min-Max Ascent.

These results are consistent with the results in [8] for the GEN minimization method. When applied to Rosenbrock's function, GEN is approximately 25 time faster than Newton's method and approximately 2500 times faster than Steepest Descent.

We use the "Stingray saddle" case to study Lyapunov exponents. Figure 7 shows the time history of $u(t)$ for the trajectories. Figure 8 shows the instantaneous Lyapunov exponents (3) for the Newton and Gradient Enhanced Min-Max methods, and the first Lyapunov exponent for Min-Max Ascent. Figure 9 shows the second Lyapunov exponent for Min-Max Ascent.

Initially, Newton's method is only slightly stiff, with $\Delta\sigma(t) = \sigma_1(t) - \sigma_2(t)$ and $\max \Delta\sigma(t) \approx 0.6$, but soon becomes nonstiff, with $\sigma_1 = \sigma_2 = -1$ as $t \to \infty$. The GEMM method is uniformly nonstiff, with $\max \Delta\sigma(t) \approx 0.03$ and $\sigma_1 = \sigma_2 = -1$ as $t \to \infty$. Min-Max Ascent is uniformly stiff, with $\Delta\sigma(t) \approx 100$ for all t and $\sigma_1 = -1$ and $\sigma_2 = -100$ as $t \to \infty$.

Table 1: Banana saddle results ($a = 1000$, $c = 1$).

Method	Δt	$\|\dot{\mathbf{x}}(0)\|\Delta t$	**Final** t	**# Steps**	**Ratio**
Min-Max Ascent	10^{-6}	1.499×10^{-3}	6.211	6, 210, 995	980.2
Newton	10^{-3}	1.513×10^{-3}	14.221	14, 221	2.24
GEMM	2.5×10^{-3}	1.499×10^{-3}	15.84	6, 336	1

Table 2: Stingray saddle results ($a = 1$, $c = 100$).

Method	Δt	$\|\dot{\mathbf{x}}(0)\|\Delta t$	**Final** t	**# Steps**	**Ratio**
Min-Max Ascent	10^{-5}	1.256×10^{-3}	7.330	732, 973	175.5
Newton	10^{-3}	1.296×10^{-3}	11.74	11, 740	2.81
GEMM	1.5×10^{-2}	1.254×10^{-3}	62.625	4, 175	1

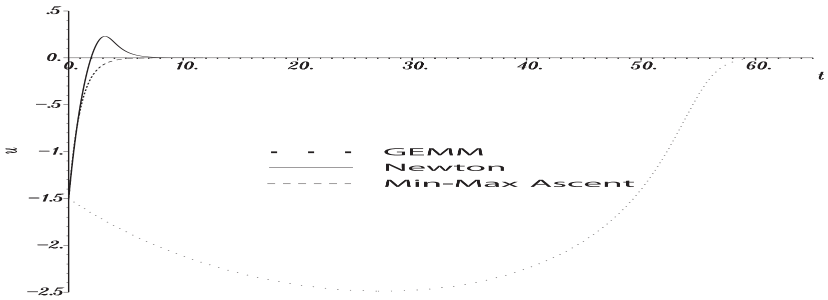


Figure 7: Time response ($a = 1$, $c = 100$).

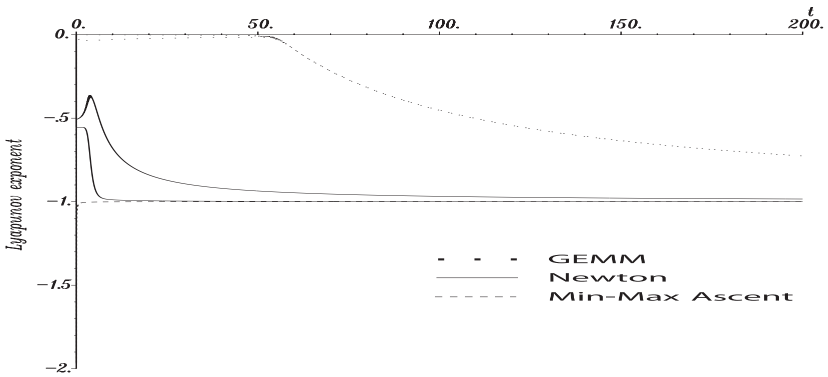


Figure 8: Lyapunov exponents ($a = 1$, $c = 100$).

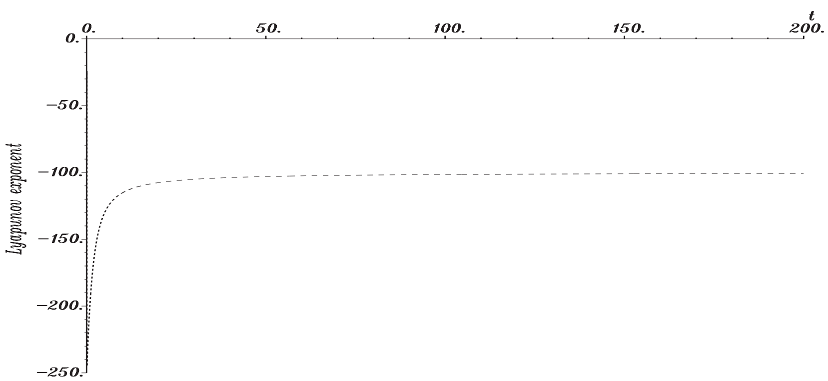


Figure 9: Second Lyapunov exponent for Min-Max Ascent ($a = 1$, $c = 100$).

8 Conclusions

The GEMM method provides global asymptotic stability to the saddle point for functions such as the Stingray saddle function, which have a single proper stationary min-max point and satisfy a Lyapunov growth condition. For the Stingray function Newton's method is not stiff but does not provide global asymptotic stability. Min-Max Ascent, applied to the Stingray function, provides global asymptotic stability but is very stiff. When applied to the Stingray function, the GEMM method is very fast and is not stiff, whereas Min-Max Ascent is very slow and very stiff. The GEMM method is approximately 3 times faster than Newton's method and approximately 175 to 1000 times faster than Min-Max Ascent.

References

[1] Arrow, K. J., and Hurwicz, L., Gradient Method for Concave Programming, I: Local Results, in *Studies in Linear and Non-Linear Programming*, Arrow, K. J., Hurwicz, L., and Uzawa, H., eds., Stanford Univ. Press, Stanford, CA, 1958, pp. 117–126.

[2] Bloch, A. M., Brockett, R. W., and Ratiu, T. S., On the Geometry of Saddle Point Algorithms, *Proc. 31st I.E.E.E. Conf. on Decision and Control*, Tucson, AZ, December 1992, pp. 1482–1487.

[3] Brogan, W. L., *Modern Control Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

[4] Chong, E. K. P., and Zak, S. H., *An Introduction to Optimization*, 2nd Ed., Wiley, New York, 2001.

[5] Goh, B. S., Algorithms for Unconstrained Optimization Problems via Control Theory, *J. of Optimization Theory and Applications*, Vol. 92, No. 3, March 1997, pp. 581–604.

[6] Gomulka, J., Remarks on Branin's Method for Solving Nonlinear Equations, in *Towards Global Optimization*, Dixon, L. C. W. and Szegö, G. P., eds., North-Holland, Amsterdam, 1972, pp. 96–106.

[7] Gomulka, J., Two Implementations of Branin's Method: Numerical Experience, in *Towards Global Optimization 2*, Dixon, L. C. W. and Szegö, G. P., eds., North-Holland, Amsterdam, 1978, pp. 151–164.

[8] Grantham, W. J., Trajectory Following Optimization by Gradient Transformation Differential Equations, *Proc. 42nd I.E.E.E. Conf. on Decision and Control*, Maui, HI, December 9–12, 2003, pp. 5496–5501.

[9] Grantham, W. J., and Lee, B. S., A Chaotic Limit Cycle Paradox, *Dynamics and Control*, Vol. 3, No. 2, April 1993, pp. 159–173.

[10] Kato, T., *A Short Introduction to Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1982.

[11] Luenberger, D. G., *An Introduction to Dynamic Systems*, Wiley, New York, 1979.

[12] Vincent, T. L., Goh, B. S., and Teo, K. L., Trajectory-Following Algorithms for Min-Max Optimization Problems, *J. of Optimization Theory and Applications*, Vol. 75, No. 3, December 1992, pp. 501–519.

[13] Vincent, T. L., and Grantham, W. J., *Optimality in Parametric Systems*, Wiley, New York, 1981.

[14] Vincent, T. L., and Grantham, W. J., *Nonlinear and Optimal Control Systems*, Wiley, New York, 1991.

[15] Vincent, T. L., and Grantham, W. J., Trajectory Following Methods in Control System Design, *J. of Global Optimization*, Vol. 23, 2002, pp. 267–282.

[16] Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A., Determining Lyapunov Exponents from a Time Series, *Physica 16D*, 1985, pp. 285–317.

Singular Perturbation Trajectory Following Algorithms for Min-Max Differential Games

Dale B. McDonald
School of Mechanical and Materials Engineering
Washington State University
Pullman, WA 99164-2920, USA
dbmcdonald@gocougs.wsu.edu

Walter J. Grantham
School of Mechanical and Materials Engineering
Washington State University
Pullman, WA 99164-2920, USA
grantham@mme.wsu.edu

Abstract

This chapter examines trajectory following algorithms for differential games subject to simple bounds on player strategy variables. These algorithms are trajectory following in the sense that closed-loop player strategies are generated directly by the solutions to ordinary differential equations. Player strategy differential equations are based upon Lyapunov optimizing control techniques and represent a balance between the current penetration rate for an appropriate descent function and the current cost accumulation rate. This numerical strategy eliminates the need to solve 1) a min-max optimization problem at each point along the state trajectory and 2) nonlinear two-point boundary-value problems. Furthermore, we address "stiff" systems of differential equations that arise during the design process and seriously degrade algorithmic performance. We use standard singular perturbation methodology to produce a numerically tractable algorithm. This results in the Efficient Cost Descent (ECD) algorithm which possesses desirable characteristics unique to the trajectory following method. Equally important as a specification of a new trajectory following algorithm is the observation and resolution of several issues regarding the design and implementation of a trajectory following algorithm in a differential game setting.

Key words. Differential game, trajectory following, stiffness, singular perturbation.

1 Introduction

This chapter considers two-player zero-sum differential games where trajectory following optimization and Lyapunov optimizing control methods dictate one or both player strategies. Player strategies are implemented that represent a trade-off between the current rate of cost accumulation and the penetration rate for that player's Lyapunov descent function. That is, we specify closed-loop feedback player strategies directly, by solving ordinary differential equations that approximately optimize a desired merit function. These strategies eliminate the need to solve a min-max optimization problem at each point of the state trajectory, as well as the need to solve two-point boundary-value problems usually associated with differential games. Since this trajectory following strategy is approximate or suboptimal, we introduce a gain on a player's differential equation strategy update to influence the manner in which the game is played. This gain amounts to a singular perturbation on the control variable and affects the performance of the algorithm. By construction, this creates a multi-time scale system that makes numerical calculations difficult and interferes with the "on-line" nature of our controller. Singular perturbation techniques are used to deal with these "stiff" equations. What results is a new trajectory following algorithm for the solution of such games that is generated completely on-line and is not stiff. This algorithm is therefore easy to implement numerically. This paper also considers several issues regarding the design and implementation of a trajectory following algorithm in a differential game setting and resolves practical issues that arise during such a process.

The trajectory following method has proven to be quite effective, with application to control systems, optimization, and game theory. In [6] Steepest Descent differential equations are used to design controllers for nonlinear systems. In [3] "Gradient Enhanced Newton" differential equations are used to provide global asymptotic stability to the minimum point of a nonlinear function. In [7] trajectory following algorithms are developed to solve nonlinear min-max optimization problems.

In Section 2 we define the general type of differential game we will consider. In Section 3 we state necessary conditions for solving such a game. In Section 4 we discuss trajectory following and singular perturbation techniques that will allow us to avoid solving the min-max optimization problem of Section 3 at each point along a state trajectory. In Section 5 we consider three trajectory following strategies applied to a linear-quadratic game. The results of this example help us derive the Efficient Cost Descent (ECD) algorithm in Section 6 and formally specify it in Section 7. In Section 8 we present the conclusions.

2 Zero-Sum Differential Games

In general, we consider a differential game [5, pp. 513–514] having a cost functional

$$J[u(\cdot), v(\cdot)] = \int_0^{t_f} f_0(\mathbf{x}, u, v)\,\mathrm{d}t \tag{1}$$

which is associated with transporting the state $\mathbf{x}(t) \in R^{n_x}$ of a system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u, v) \tag{2}$$

from some initial point

$$\mathbf{x}(0) = \mathbf{x}_i \tag{3}$$

to any final state $\mathbf{x}(t_f)$ in a specified terminal set $\mathcal{X} \subset R^{n_x}$,

$$\mathbf{x}(t_f) \in \mathcal{X} = \{\mathbf{x} \in R^{n_x} : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}, \tag{4}$$

where $\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}) \cdots g_{n_g}(\mathbf{x})]^{\mathsf{T}}$, $()^{\mathsf{T}}$ denotes transpose, $\dot{()} = \mathrm{d}()/\mathrm{d}t$, and t denotes time. The final time t_f is generally left unspecified and is defined as the first time that the state reaches the target. The player strategy variables u and v, respectively, are constrained to be elements of the constraint sets $\mathcal{U} = \{u \in R^{n_u} : \mathbf{h}(u) \geq \mathbf{0}\}$ and $\mathcal{V} = \{v \in R^{n_v} : \hat{\mathbf{h}}(v) \geq \mathbf{0}\}$, where $\mathbf{h}(u) = [h_1(u) \cdots h_{n_h}(u)]^{\mathsf{T}}$ and $\hat{\mathbf{h}}(v) = [\hat{h}_1(v) \cdots \hat{h}_{n_{\hat{h}}}(v)]^{\mathsf{T}}$. Player 1 chooses the strategy $u(\mathbf{x})$ and desires to minimize (1). Player 2 chooses the strategy $v(\mathbf{x})$ and desires to maximize (1). We assume that $f_0(\mathbf{x}, u, v)$ and each component of the vector functions $\mathbf{f}(\mathbf{x}, u, v)$, $\mathbf{g}(\mathbf{x})$, $\mathbf{h}(u)$, and $\hat{\mathbf{h}}(v)$ are continuous and continuously differentiable with respect to their arguments. Furthermore, we assume that all points in $\mathcal{X}, \mathcal{U}$, and $\mathcal{V}$ are regular points [5, p. 125]. In particular, we assume that the gradient vectors $\partial g_i(\mathbf{x})/\partial \mathbf{x}$ are linearly independent at each point in $\mathcal{X}$. The gradient vectors $\partial h_i(u)/\partial u$ are also assumed to be linearly independent at any active $\mathcal{U}$ constraint $h_i(u) = 0$. Similarly, the gradient vectors $\partial \hat{h}_i(v)/\partial v$ are assumed to be linearly independent at any active $\mathcal{V}$ constraint $\hat{h}_i(v) = 0$ [5, p. 514].

3 Min-Max Optimization

Solution of a min-max differential game requires that we confront an underlying, possibly nonlinear, min-max optimization problem at each point $\mathbf{x}$ along a state trajectory $\mathbf{x}(t)$. In general terms we must perform

$$\min_{u \in \mathcal{U}} \max_{v \in \mathcal{V}} H(\mathbf{x}, u, v, \boldsymbol{\lambda}, \lambda_0) \tag{5}$$

subject to

$$\begin{aligned} \mathbf{h}(u) &\geq \mathbf{0} \\ \hat{\mathbf{h}}(v) &\geq \mathbf{0}. \end{aligned} \tag{6}$$

The differential game H function is given by $H(\mathbf{x}, u, v, \boldsymbol{\lambda}, \lambda_0) = \lambda_0 f_0(\mathbf{x}, u, v) + \boldsymbol{\lambda}^{\mathsf{T}} \mathbf{f}(\mathbf{x}, u, v)$, where f_0 is the accumulated cost, $\mathbf{f}$ contains the system dynamics, λ_0 is a constant, and $\boldsymbol{\lambda}$ is the adjoint vector. The vector $\mathbf{x}$ is the state, and u and v are control vectors with $\mathcal{U} \subseteq R^{n_u}$ and $\mathcal{V} \subseteq R^{n_v}$ specified constraint sets consisting of n_h-dimensional and $n_{\hat{h}}$-dimensional vectors of inequality constraint functions $\mathbf{h}(u) \geq \mathbf{0}$ and $\hat{\mathbf{h}}(v) \geq \mathbf{0}$, respectively. Our assumptions on $f_0(\cdot)$ and $\mathbf{f}(\cdot)$ imply that $H(\mathbf{x}, u, v, \boldsymbol{\lambda}, \lambda_0)$ is continuous and continuously differentiable in u and v.

The necessary conditions for a game theoretic saddle point solution can be summarized as follows [5, pp. 521–522].

Let u^* and v^* be regular points of $\mathcal{U}$ and $\mathcal{V}$, respectively. If H takes on a $\min_u \max_v$ *at $u = u^*$ and $v = v^*$, then Lagrange multipliers $\boldsymbol{\gamma} = [\gamma_1 \cdots \gamma_{n_h}]^\mathsf{T}$ and $\hat{\boldsymbol{\gamma}} = [\hat{\gamma}_1 \cdots \hat{\gamma}_{n_{\hat{h}}}]^\mathsf{T}$ exist such that*

$$\begin{aligned} \frac{\partial L}{\partial u} &= \mathbf{0}^\mathsf{T} \\ \frac{\partial L}{\partial v} &= \mathbf{0}^\mathsf{T} \\ \mathbf{h}(u^*) &\geq \mathbf{0} \\ \hat{\mathbf{h}}(v^*) &\geq \mathbf{0} \\ \boldsymbol{\gamma}^\mathsf{T}\mathbf{h}(u^*) &= 0 \\ \hat{\boldsymbol{\gamma}}^\mathsf{T}\hat{\mathbf{h}}(v^*) &= 0 \\ \boldsymbol{\gamma} &\geq \mathbf{0} \\ \hat{\boldsymbol{\gamma}} &\leq \mathbf{0}, \end{aligned} \tag{7}$$

where the Lagrangian function is $L(\mathbf{x}, u, v, \boldsymbol{\lambda}, \lambda_0, \boldsymbol{\gamma}, \hat{\boldsymbol{\gamma}}) = H(\mathbf{x}, u, v, \boldsymbol{\lambda}, \lambda_0) - \boldsymbol{\gamma}^\mathsf{T}\mathbf{h}(u) - \hat{\boldsymbol{\gamma}}^\mathsf{T}\hat{\mathbf{h}}(v)$ and the partial derivatives are evaluated at $u = u^*$ and $v = v^*$.

For state-independent control constraints the above results can be re-stated in terms of a differential game min-max principle [5, pp. 522–523].

Min-Max Principle: Given the constraint sets $\mathcal{U}$ and $\mathcal{V}$, if $u^*(\mathbf{x}) \in \mathcal{U}$ and $v^*(\mathbf{x}) \in \mathcal{V}$ are min-max controls for the differential game (1)–(4), then there must exist a continuous and piecewise differentiable vector function $\boldsymbol{\lambda}(t) = [\lambda_1 \cdots \lambda_{n_x}]^\mathsf{T}$, a constant $\boldsymbol{\lambda}_0 \geq 0$ ($= 0$ or $= 1$), with $(\lambda_0, \boldsymbol{\lambda}) \neq \mathbf{0}$, and a constant multiplier vector $\boldsymbol{\rho} = [\rho_1 \cdots \rho_{n_x}]^\mathsf{T}$ such that $\boldsymbol{\lambda}(t)$ satisfies the adjoint equations

$$\dot{\boldsymbol{\lambda}}^\mathsf{T} = -\frac{\partial H}{\partial \mathbf{x}}, \tag{8}$$

where, at the terminal set $\mathbf{g}[\mathbf{x}(t_f)] = \mathbf{0}$, $\boldsymbol{\lambda}(t_f)$ satisfies the transversality conditions

$$\boldsymbol{\lambda}^\mathsf{T}(t_f) = \boldsymbol{\rho}^\mathsf{T}\frac{\partial \mathbf{g}[\mathbf{x}(t_f)]}{\partial \mathbf{x}} \tag{9}$$

and such that H takes on a global $\min_u \max_v$ value with respect to u and v at each point $\mathbf{x}$ along the trajectory generated by $u^*(\mathbf{x})$ and $v^*(\mathbf{x})$. The $\min_u \max_v$ value of H at every such point is zero.

In general, solution of such a differential game from $\mathbf{x}(0) = \mathbf{x}_i$ generates open-loop controls $u(t)$ and $v(t)$ along a trajectory $\mathbf{x}(t)$, and involves the solution of a two-point boundary-value problem, with initial conditions (3) specified on $\mathbf{x}(t)$ at $t = 0$ and terminal conditions (9) specified on $\boldsymbol{\lambda}(t)$ at $t = t_f$. To generate closed-loop strategies $u(\mathbf{x})$ and $v(\mathbf{x})$ would require solving the two-point boundary-value problems from each initial state $\mathbf{x}$.

4 Methodology

Implementation of a trajectory following algorithm, through the use of singular perturbation techniques, eliminates the need to solve the (possibly) nonlinear min-max optimization problem (5) of Section 3 at each point along the state trajectory, as well as the need to solve two-point boundary-value problems. Such an algorithm is nonstiff and easy to implement numerically in an on-line manner.

4.1 Trajectory Following Optimization

The forthcoming derivation will focus upon the minimizing player u with player 2, controlling v, representing an uncertain input. We develop trajectory following algorithms for $u(\mathbf{x})$, where the strategy variables u and v are subject to simple inequality constraints, by forming differential equations taken from a function $\dot{W}_0$. Qualitatively, this function represents a trade-off between the current rate of cost accumulation and the penetration rate for that player's Lyapunov descent function W [5, p. 276]. This combination of function minimization and Lyapunov stability techniques will provide structure for our u player trajectory following algorithms.

Consider the accumulated cost [5, pp. 295–297] given by

$$x_0 = \int_0^t f_0(\mathbf{x}, u, v)\,\mathrm{d}t, \tag{10}$$

which accumulates at a rate

$$\dot{x}_0 = f_0(\mathbf{x}, u, v), \tag{11}$$

with the initial condition $x_0(0) = 0$. Now let $W(\mathbf{x})$ denote a descent function, such as distance to the target $\mathcal{X}$ (for example, $W = \mathbf{x}^\mathsf{T}\mathbf{x}$ with $\mathcal{X}$ specified as the origin, or a small spherical radius of the origin), and define an augmented descent function $W_0 = x_0 + W(\mathbf{x})$. The time derivative of W_0 is

$$\dot{W}_0 = \frac{\partial W_0}{\partial x_0}\dot{x}_0 + \frac{\partial W_0}{\partial \mathbf{x}}\dot{\mathbf{x}} = f_0(\mathbf{x}, u, v) + \frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}, u, v) \tag{12}$$

which balances the rate f_0 of accumulation of cost with the current penetration rate of the descent function W along trajectories generated by the state equations (2).

We first consider steepest descent and "total derivative" updates based on (12), subject to the control constraints (6). From that analysis we derive our Efficient Cost Descent (ECD) strategy update. Simple bounds on player strategy variables will be handled as in [6]. These algorithms will have the general form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, u, v) \\ \dot{u} &= \hat{\mathbf{f}}(\mathbf{x}, u, v) \end{aligned} \tag{13}$$

with $u_{\min} \le u \le u_{\max}$, $v_{\min} \le v \le v_{\max}$, where we will derive $\hat{\mathbf{f}}(\mathbf{x}, u, v)$ from (12). Use of the trajectory following method in a differential game setting (1)–(4) has essentially “elevated” the strategy u to state variable status in (13).

4.2 Singular Perturbations

It may occur that a gain, or equivalently, a singular perturbation introduced into (13) creates a more robust strategy update $\dot{u}$. Defining a small, positive, singular perturbation parameter ϵ, we now have

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, u, v) \\ \epsilon \dot{u} &= \hat{\mathbf{f}}(\mathbf{x}, u, v). \end{aligned} \tag{14}$$

The system (14) possesses a two-time-scale property with n_x slow modes and n_u fast modes [1]. Equations (14) represent a stiff system of differential equations, which can make numerical integration difficult. Singular perturbation solution methodology will be used to relieve the numerical stiffness and produce a fast, efficient trajectory following algorithm. Stiff differential equation solvers can handle (14), but we will address this issue up front and give the analyst low-level control over the process.

Singular perturbation theory provides a means by which the system (14) may be converted into slow and fast subsystems [1]. Each subsystem may be examined on a single time scale to determine information about the original system. In this chapter we will extract such information from the slow and fast subsystems to develop a trajectory following update $\dot{u}$ that can be efficiently handled with a nonstiff numerical integrator for differential equations.

4.2.1 Slow Subsystem

The slow subsystem of (14) is found by assuming that the fast modes are infinitely fast, essentially letting $\epsilon \to 0$ [1]. The slow subsystem is then

$$\begin{aligned} \dot{\mathbf{x}}_s &= \mathbf{f}(\mathbf{x}_s, u_s, v_s) \\ \mathbf{0} &= \hat{\mathbf{f}}(\mathbf{x}_s, u_s, v_s), \end{aligned} \tag{15}$$

where the subscript “s” denotes a slow variable.

4.2.2 Fast Subsystem

The fast subsystem of (14) is found by assuming that the slow variables are constant during the initial transient phase. In general, the fast subsystem is given by

$$\epsilon \dot{u}_f = \hat{\mathbf{f}}(\mathbf{x}_f, u_f, v_f), \tag{16}$$

where the subscript “f” denotes a fast variable.

5 Zero-Sum Linear-Quadratic Game

Prior to the derivation of the ECD algorithm we consider the following linear quadratic differential game.

Example 5.1. Consider a two-player linear-quadratic differential game where the state, governed by

$$\begin{aligned}\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u, v) &= [\mathbf{A}\mathbf{x} + \mathbf{B}u + \mathbf{E}v] \\ &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u + \begin{bmatrix} 0 \\ 1 \end{bmatrix}v, \end{aligned} \tag{17}$$

evolves under the control of two inputs $u \in R^1$ and $v \in R^1$, with $x \in R^2$. The u and v players seek to minimize and maximize, respectively, a quadratic cost functional

$$\begin{aligned} J[u(\cdot), v(\cdot)] &= \frac{1}{2}\int_0^{t_f} (\mathbf{x}^{\mathsf{T}}\mathbf{Q}\mathbf{x} + u^{\mathsf{T}}\mathbf{R}u - \rho^2 v^{\mathsf{T}} v)\,\mathrm{d}t \\ &= \frac{1}{2}\int_0^{t_f} (x_1^2 + x_2^2 + u^2 - \rho^2 v^2)\,\mathrm{d}t, \end{aligned} \tag{18}$$

subject to $|u| \le u_{\max} = 2$ and $|v| \le v_{\max} = 1$. Let the descent function to the target (the origin; more precisely, a small radius of the origin) be given by

$$W(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathsf{T}}\begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}\mathbf{x} = \frac{1}{2}(x_1^2 + x_1x_2 + x_2^2), \tag{19}$$

and we now have

$$\dot{W}_0 = \frac{1}{2}\big(x_1^2 + x_2^2 + u^2 - \rho^2 v^2\big) + \frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}, u, v). \tag{20}$$

We let the parameter $\rho^2 = 1.1$ and refer the reader to [5, pp. 530–531] for further information. This system corresponds to a linearized inverted pendulum, where a gravitational term has been replaced by the uncertain input v [8]. The state variable x_1 is the angular displacement of the pendulum from the vertical, x_2 is the angular velocity, and the control u is the applied nongravitational torque. Player 1, controlling u, desires to minimize (18) while driving the system to the target. The target is given by equality constraints on the state as in (4), which for this example are

$$g(\mathbf{x}) = x_1^2 + x_2^2 - r^2 = 0, \tag{21}$$

where r is a specified small radius from the origin (we take $r = 10^{-4}$). Player 2, controlling the uncertain input v, desires to maximize (18) while driving the system away from the target (21).

5.1 Minimum Cost Descent Algorithm

We define the Minimum Cost Descent (MCD) trajectory following algorithm by setting the $\dot{u}$ strategy update for player 1 equal to the negative of the gradient of $\dot{W}_0$ with respect to u. That is,

$$\dot{u} = -\left[\frac{\partial \dot{W}_0}{\partial u}\right]^{\mathsf{T}}. \tag{22}$$

For Example 5.1 this yields

$$\dot{u} = -\frac{\partial \dot{W}_0}{\partial u} = -\left(\frac{1}{2}x_1 + x_2 + u\right), \tag{23}$$

which is desirable since $\dot{u} \to 0$ only as $\partial \dot{W}_0/\partial u \to 0$. We use a "saturation" version of (23); if u is at its upper or lower bound and $\dot{u}$ from (23) would cause u to exceed its bounds, then instead we set $\dot{u} = 0$.

Assuming that the minimizing strategy variable has not saturated and ignoring the uncertain input for the moment, we may examine the eigenvalues of the augmented linear system (13). We have

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{u} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{1}{2} & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} \tag{24}$$

with eigenvalues $\mu_1 = -0.1761 + \mathrm{i}0.8607$, $\mu_2 = -0.1761 - \mathrm{i}0.8607$, and $\mu_3 = -0.6478$.

5.2 Total Descent Algorithm

Consider a trajectory following update for the minimizing strategy variable by minimizing $\dot{W}_0$ with respect to u. From the necessary conditions for an interior minimum we have $\mathbf{0}^{\mathsf{T}} = \partial \dot{W}_0/\partial u$. One technique [4, p. 47] from singular perturbation theory would be to set the time derivative of this condition to zero. This suggests that we update u according to

$$\mathbf{0} = \frac{\partial}{\partial u}\left[\frac{\partial \dot{W}_0}{\partial u}\right]^{\mathsf{T}} \dot{u} + \frac{\partial}{\partial \mathbf{x}}\left[\frac{\partial \dot{W}_0}{\partial u}\right]^{\mathsf{T}} \dot{\mathbf{x}}. \tag{25}$$

That is, we set

$$\dot{u} = -\left[\frac{\partial^2 \dot{W}_0}{\partial u^2}\right]^{-1} \frac{\partial^2 \dot{W}_0}{\partial u \partial \mathbf{x}} \dot{\mathbf{x}}, \tag{26}$$

resulting in the Total Descent (TD) algorithm. Implementing this strategy for Example 5.1 results in

$$\dot{u} = -\left(\frac{1}{2}x_2 + u\right) \tag{27}$$

and once again, we use a saturation version of (27).

Assuming that the minimizing strategy variable has not saturated and ignoring the uncertain input, we may examine the eigenvalues of the augmented linear system (13). We have

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{u} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} \tag{28}$$

with eigenvalues $\mu_1 = -\frac{1}{2} + \mathrm{i}\frac{1}{2}$, $\mu_2 = -\frac{1}{2} - \mathrm{i}\frac{1}{2}$, and $\mu_3 = 0$.

5.2.1 MCD and TD Example 5.1 Results

Let the initial conditions for Example 5.1 be given by $x_1(0) = 1$, $x_2(0) = 1$, $u(0) = -1.5$ and for the moment let $v = 0$. Trajectories evolving from these initial conditions for the MCD and TD trajectory following algorithms are shown in Figure 1. The terminal costs (18) for the MCD and TD algorithms were 5.37 and 3.31, respectively, and each algorithm drove the state to the target. The TD algorithm has driven the state to the target and done so at a lower cost than the MCD algorithm.

Now let the initial conditions be given by $x_1(0) = 1$, $x_2(0) = 1$, and $u(0) = -2$ where again we let $v(t) = 0$. Trajectories evolving from these initial conditions for the MCD and TD trajectory following algorithms are shown in Figure 2. The MCD algorithm drove the system to the target with a terminal cost of 6.62; the TD algorithm failed to drive the system to the target. We expected this latter result as one eigenvalue of (28) was zero; previous convergence to the target was a product of the choice of initial conditions. Given $x_1(0) = 1$ and $x_2(0) = 1$, $\min[\dot{W}_0(1, 1, u)] \Rightarrow u^*(0) = -1.5$.

The results of Example 5.1, along with Figures 1 and 2, illustrate the behavior of the MCD and TD algorithms. In particular, we note that the TD algorithm could

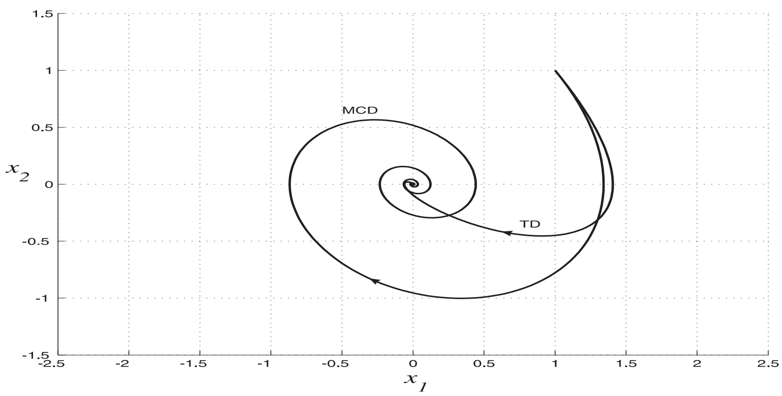


Figure 1: TD and MCD state trajectories for Example 5.1 [$u(0) = -1.5$].

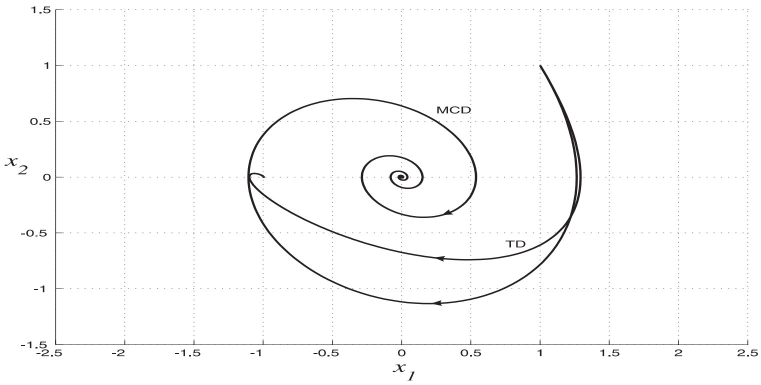


Figure 2: TD and MCD state trajectories for Example 5.1 [$u(0) = -2$].

only drive the system to the target for very specific initial conditions, *even in the absence of the disturbance*! We now turn to the v-reachable set [5, p. 504] to provide a qualitative means by which we may compare the effectiveness of any player 1 strategy $u(\mathbf{x})$ in the presence of the disturbance v.

Definition 5.1. A given point in state space is v-**reachable** from the origin if there exists an admissible control law for v such that, under a given admissible control law for $u(\mathbf{x})$, the system can be driven from the origin to the point in finite time.

Definition 5.2. The v-**reachable set** is the set of all v-reachable points under the specified control law $u(\mathbf{x})$.

The v-reachable set provides a qualitative measure of the capabilities of the disturbance to drive the system to points other than the origin (target) that is particularly well suited for two-dimensional problems. Under a specified control law $u(\mathbf{x})$, the v-reachable set is determined by finding the reachable set [5, p. 315] from the origin for the v player. That is, this process reduces to an abnormal control problem for the v player. Necessary conditions for trajectories lying in the boundary of the v-reachable set are given as the reachability maximum principle [5, p. 342]. The v-reachable set is found by integrating the system state equations forward in time, starting near the origin, subject to the input v obtained from the reachability maximum principle [5, p. 507]. For Example 5.1, this input is $v = v_{\max} \operatorname{sgn}(x_2)$.

The v-reachable set for the MCD algorithm, which is not shown, is unbounded. The "unknown" disturbance is able to drive the system from the origin to any point in state space. Modification of (23) is needed to counter the effect of the disturbance, resulting in a bounded v-reachable set. We did not include the TD algorithm in this analysis, since the "state variable" u has a state constraint $|u| \leq 2$. This would make the determination of the boundary of the v-reachable set very difficult [2]. In fact, setting $v = v_{\max} \operatorname{sgn}(x_2)$ and integrating (28) forward in time results in a trajectory that approaches infinity as time approaches infinity. So the υ-reachable set for the TD algorithm is also unbounded.

5.3 Singular Perturbation Minimum Cost Descent

A singular perturbation will be incorporated into the MCD algorithm to counter the effect of the disturbance input. We then have the Singular Perturbation Minimum Cost Descent (SPMCD) algorithm

$$\epsilon \dot{u} = -\frac{\partial \dot{W}_0}{\partial u} = -\left(\frac{1}{2}x_1 + x_2 + u\right), \tag{29}$$

with saturation. Assuming that the minimizing strategy variable has not saturated and ignoring the uncertain input, we may examine the eigenvalues of the augmented linear system (13). We have

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \epsilon \dot{u} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{1}{2} & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix} \tag{30}$$

with eigenvalues $\mu_1 = -0.5 + \mathrm{i}0.5$, $\mu_2 = -0.5 - \mathrm{i}0.5$, and $\mu_3 = -9999$ for $\epsilon = 0.0001$.

Let the initial conditions for Example 5.1 be given by $x_1(0) = 1$, $x_2(0) = 1$, $u(0) = -2.0$, with $\upsilon = 0$. Figure 3 displays the state trajectories for all three trajectory following algorithms we have considered to this point. As before, the MCD algorithm drove the state to the target while the TD algorithm did not. The SPMCD algorithm drove the state to the target with a terminal cost of 3.311.

We examine the υ-reachable set for SPMCD; Figure 4 displays the resulting limit cycle boundary of the υ-reachable set. The SPMCD algorithm yields a bounded

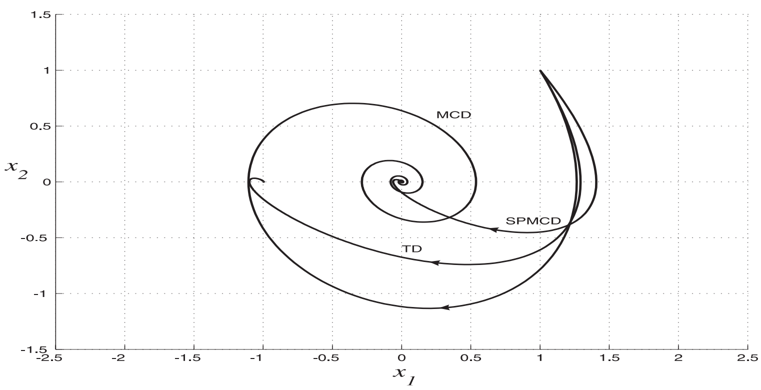


Figure 3: State trajectories for Example 5.1 [$u(0) = -2$].

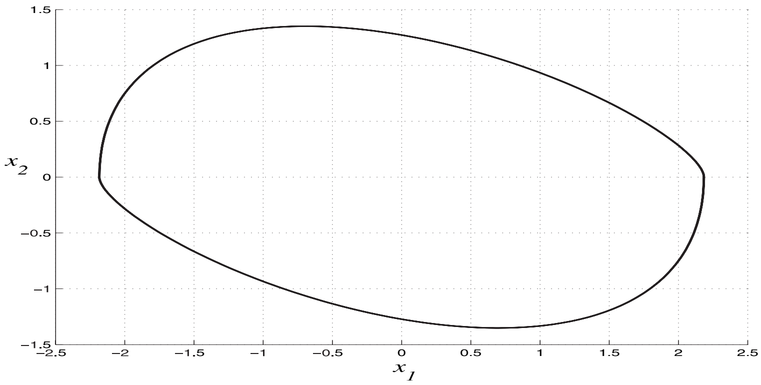


Figure 4: The υ-reachable set for the SPMCD algorithm, Example 5.1.

υ-reachable set; the algorithm is much more robust in regard to the unknown disturbance than the MCD or TD algorithms.

The results of Example 5.1 illustrate the strengths and weaknesses of each of the three algorithms MCD, TD, and SPMCD. The MCD algorithm drove the system to the target but only if the input disturbance was ignored. The TD algorithm only drove the system to the target for very specific initial conditions; when it did work its terminal cost was quite low. Both the MCD and TD algorithms operated on a single time scale (nonstiff) as shown by their eigenvalues. In the absence of a disturbance the SPMCD converged to the target and its terminal cost was the smallest. The eigenvalues of (30) clearly show that the SPMCD algorithm yields a stiff system of differential equations.

In the presence of a bounded disturbance υ, the υ-reachable set for Example 5.1 is all of R^2 for the MCD algorithm, and is at least unbounded for the TD algorithm, but is a bounded set for the SPMCD algorithm. We now appeal to singular perturbation analysis techniques to derive a trajectory following algorithm for player 1 of the differential game (1)–(4) that drives the system to the target in the absence of the disturbance input. This algorithm should be generally nonstiff and therefore easily implemented numerically. Finally, we desire the υ-reachable set for any algorithm we develop to be no larger than that of the SPMCD algorithm.

6 SPMCD Time Scale Analysis

In Example 5.1, in the absence of a disturbance υ, the SPMCD algorithm drove the system to the target and the terminal cost was lower than that of the MCD and TD algorithms. With a disturbance the υ-reachable set was unbounded for the MCD and TD algorithms, but was bounded for the SPMCD algorithm, suggesting that the primary shortcoming of the SPMCD method is stiffness, which makes numerical calculations difficult. The ECD algorithm is derived from the time scale decomposition of the SPMCD algorithm.

Singular perturbation theory [1] tells us that the fast modes of a given system of differential equations are important only during a short initial period; at the conclusion of this period the slow modes determine the behavior of the system. This suggests that the effect of the perturbation ϵ in (30) need only be felt during a short initial time interval. Limiting the use of the perturbation would immediately relieve the stiffness of the system of differential equations. Neglecting ϵ completely degrades the overall performance of the algorithm, as seen by comparing the terminal costs and υ-reachable sets of MCD and SPMCD in Example 5.1.

6.1 Slow Subsystem

We derive the slow subsystem of (30) by assuming the fast modes are infinitely fast. That is, we let $\epsilon = 0$. We have

$$\begin{bmatrix} \dot{x}_{1\mathrm{s}} \\ \dot{x}_{2\mathrm{s}} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{1}{2} & -1 & -1 \end{bmatrix} \begin{bmatrix} x_{1\mathrm{s}} \\ x_{2\mathrm{s}} \\ u_{\mathrm{s}} \end{bmatrix}, \tag{31}$$

where a subscript “s” represents a slow variable. In (30) we have set $\epsilon \dot{u} = -\partial \dot{W}_0/\partial u$; as a product of that choice we see that the slow system contains the original dynamics and the necessary condition for $\min_u \dot{W}_0$ for this problem. This suggests that the bulk of the information needed to drive this system along the SPMCD trajectory is contained in the slow system, on a single time scale.

6.2 Fast Subsystem

During the derivation of the fast subsystem, we assume that the slow variables are constant. In the case of Example 5.1 we then have $\dot{x}_{1\mathrm{s}} = \dot{x}_{2\mathrm{s}} = \dot{u}_{\mathrm{s}} = 0$. With x_1 and x_2 given as slow variables, the third equation in (30) may be rewritten as $\epsilon \dot{u} = -\frac{1}{2}x_{1\mathrm{s}} - x_{2\mathrm{s}} - u$. We may substitute u_{s} from (30) into (31) and noting that $\dot{u}_{\mathrm{s}} = 0$ we have $\epsilon(\dot{u} - \dot{u}_{\mathrm{s}}) = u_{\mathrm{s}} - u$. If we let $u_{\mathrm{f}} = u - u_{\mathrm{s}}$ we have $\epsilon \dot{u}_{\mathrm{f}} = -u_{\mathrm{f}}$, which has a solution $u_{\mathrm{f}} = u_{\mathrm{f}}(0)\mathrm{e}^{-t/\epsilon}$. We observe that in a trajectory following setting, where we have introduced a singular perturbation on the player strategy, the fast system for Example 5.1 depends only upon the strategy variable itself. For this particular type of example, where the strategy differential equation is linear in u, the fast system variable u_{f} is given by a rapidly decaying exponential function. This suggests that the SPMCD algorithm (30) may be modified with an exponential function to reduce the effect of ϵ yet still produce the desired results.

6.3 Algorithm Derivation

The ECD algorithm is based on singular perturbation techniques applied to the SPMCD algorithm. It consists of three components taken from the full system, the slow subsystem, and the fast subsystem. Taken together, these components create a state trajectory that approximates that of the SPMCD algorithm, but is generally nonstiff and therefore much more efficient.

6.3.1 Slow Subsystem Component

Consider general, possibly nonlinear, algebraic equations found by setting $\epsilon = 0$ in an augmented system of differential equations (14). Ignoring v for the moment, the resulting equations are of the form $\hat{\mathbf{g}}(\mathbf{x}, u) = \mathbf{0}$ [$\hat{g}(\mathbf{x}, u) = \partial \dot{W}_0/\partial u = -\hat{f}(\mathbf{x}, u)$ for SPMCD in Example 5.1], which is on the same time scale as the state variables. An update strategy for u may be found by differentiating the constraint $\hat{\mathbf{g}}(\mathbf{x}, u)$ with respect to time [4, p. 47],

$$\mathbf{0} = \frac{\partial \hat{\mathbf{g}}(\mathbf{x}, u)}{\partial u}\dot{u} + \frac{\partial \hat{\mathbf{g}}(\mathbf{x}, u)}{\partial \mathbf{x}}\dot{\mathbf{x}}, \tag{32}$$

yielding

$$\dot{u} = -\left[\frac{\partial \hat{\mathbf{g}}(\mathbf{x}, u)}{\partial u}\right]^{-1} \frac{\partial \hat{\mathbf{g}}(\mathbf{x}, u)}{\partial \mathbf{x}}\dot{\mathbf{x}}. \tag{33}$$

This update (33) cannot stand alone; if $\partial \hat{\mathbf{g}}(\mathbf{x}, u)/\partial u$ is singular the $\dot{u}$ algorithm, from (32), could cause convergence to an equilibrium point that is not contained in the target set, which is undesirable. The SPMCD strategy update algorithm may be stiff in Example 5.1, but it only produces an equilibrium point when $\hat{\mathbf{g}}(\mathbf{x}, u) = \mathbf{0}$.

6.3.2 Fast System Component

The solution to a possibly nonlinear system of differential equations $\epsilon \dot{u}_{\mathrm{f}} = -\hat{\mathbf{g}}(u_{\mathrm{f}})$ will not in general reduce to a simple exponential function. However, the results of Section 6.2 along with singular perturbation theory show that the fast subsystem is only important during a short initial period. For this reason we will design the ECD algorithm to contain a term of the form

$$\dot{u} = -\frac{1}{\epsilon}\mathrm{e}^{-\alpha t}\hat{\mathbf{g}}, \tag{34}$$

which is not on the same time scale as the state variables. This second time scale quickly decays with the exponential term.

6.3.3 Full System Component

If $\partial \hat{\mathbf{g}}(\mathbf{x}, u)/\partial u$ is not singular along a trajectory generated by the augmented state equations, the components (33) and (34) attempt to drive the system to the target with only a very short initial period of stiffness. If $\partial \hat{\mathbf{g}}(\mathbf{x}, u)/\partial u$ is singular along that trajectory we need a third component to our algorithm to ensure that progress toward the target is made. To this end we let the ECD algorithm contain the following term:

$$\dot{u} = -\beta \|\hat{\mathbf{g}}\|\hat{\mathbf{g}}. \tag{35}$$

7 Efficient Cost Descent

The initial form of the ECD algorithm is defined by its three components (33)–(35) yielding

$$\dot{u} = -\left[\frac{1}{\epsilon}\mathrm{e}^{-\alpha t} + \beta\|\hat{\mathbf{g}}\|\right]\hat{\mathbf{g}} - \left[\frac{\partial\hat{\mathbf{g}}}{\partial u}\right]^{-1}\frac{\partial\hat{\mathbf{g}}}{\partial\mathbf{x}}\dot{\mathbf{x}}, \tag{36}$$

with saturation. The ECD algorithm is generally nonstiff and therefore efficient. This behavior is inherited from the slow subsystem decomposition of SPMCD representing the far right term in (36). At locations where $\|\hat{\mathbf{g}}\|$ is small, the algorithm behaves like the TD algorithm. The ECD algorithm does not require the solution of a (possibly) nonlinear optimization problem, even at the initial point of the trajectory. The fast subsystem information contained in the exponential term of (36) drives the state near a point where the necessary conditions $\partial\dot{W}_0/\partial u = \mathbf{0}^{\mathsf{T}}$ are satisfied despite a poor choice for $u(0)$. The nonnegative constants α and β may be chosen to further control the speed of the algorithm and to avoid points where $\partial\hat{\mathbf{g}}(\mathbf{x}, u)/\partial u$ is singular. For Example 5.1 we note that $\partial\hat{\mathbf{g}}(\mathbf{x}, u)/\partial u = \partial^2\dot{W}_0/\partial u^2 \equiv 1$ is nonsingular.

7.1 Performance Comparison: ECD and SPMCD

Figure 5 illustrates the trajectory generated by the ECD algorithm for Example 5.1; the SPMCD trajectory was not included as it quickly converged to the ECD trajectory. The initial conditions for each strategy, with $\epsilon = 10^{-4}$, were $x_1(0) = 1$, $x_2(0) = 1$, $u(0) = -2.0$, $v = 0$. We have chosen $\alpha = 100$ and $\beta = 0$. The terminal costs for the SPMCD and ECD algorithms were both 3.311. The v-reachable set for the ECD algorithm, found by introducing the uncertain input, is shown in Figure 6. The ECD algorithm produces a v-reachable set identical to that of the SPMCD algorithm but in a nonstiff (efficient) manner.

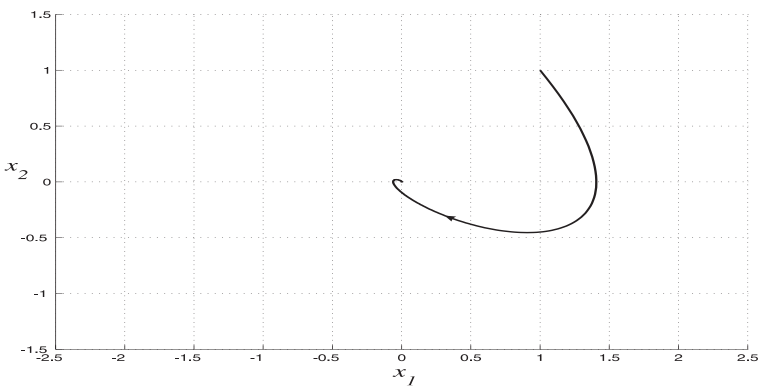


Figure 5: State trajectory for the ECD algorithm, Example 5.1 [$u(0) = -2$].

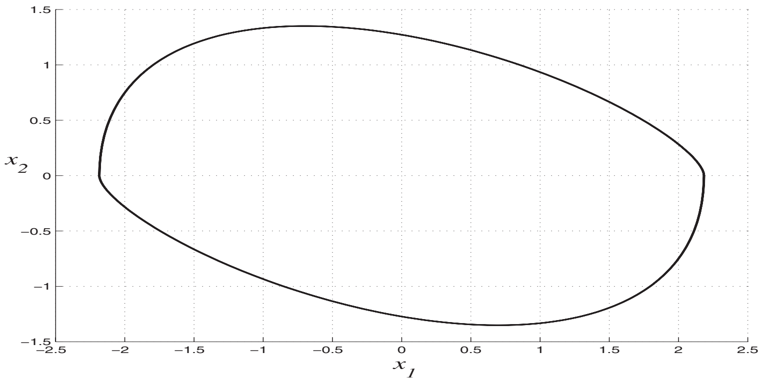


Figure 6: The v-reachable set for the ECD algorithm in Example 5.1.

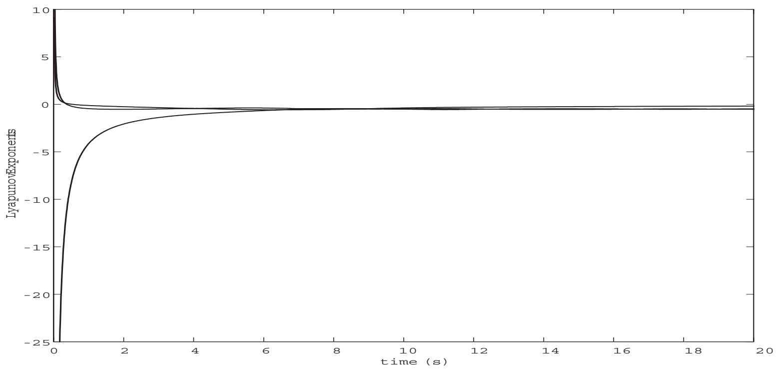


Figure 7: Lyapunov exponents for the ECD algorithm applied to Example 5.1.

The CPU time to carry out the integration was found using Matlab's built-in stopwatch timer commands `tic`, `toc` for the ECD and SPMCD algorithms. We terminated this integration when the distance to the target (origin) fell below $r = 10^{-4}$. With $\epsilon = 10^{-4}$, initial conditions $x_1(0) = 1$, $x_2(0) = 1$, $u(0) = -2.0$, $v = 0$, and using Matlab's `ode45` integrator we found the CPU time for the ECD and SPMCD algorithms to be 0.047 second and 209.23 seconds, respectively. The ECD method was approximately 4450 times as fast as SPMCD. The efficiency of the ECD algorithm can be related to its generally nonstiff behavior. We can study the stiffness of the ECD algorithm by examining its Lyapunov exponent time histories [5, pp. 205–207]. In Figure 7 we see widely separated Lyapunov exponents for a short initial period, followed by closely grouped exponents for the vast majority of the integration. This behavior is in stark contrast to that of the SPMCD algorithm with widely separated eigenvalues.

7.2 Performance Comparison: ECD and Min-Max

The ECD algorithm generates on-line, closed-loop, minimizing player strategies. These strategies are derived from (12), a function that exhibits the optimal structure of the Hamilton–Jacobi–Bellman equation [9, p. 222]

$$0 = \min_{u} \left[f_0(\mathbf{x}, u, v) + \frac{\partial V}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, u, v) \right], \tag{37}$$

where $V(\mathbf{x})$ is called the value function. If we select $W(\mathbf{x}) = V(\mathbf{x})$, we can recover the optimal solution for the particular problem; this task may be quite difficult as we must solve a partial differential equation to obtain $V(\mathbf{x})$. Instead, we have selected a descent function to the target and made trajectories decrease this function over time. Aside from solving (37), the selection of an appropriate descent function is qualitative in nature; the specification of guidelines for determining such a function is not the focus of this paper. However, for a given descent function we can show that this method produces results comparable to a game-theoretic controller. We will accomplish this task by modifying the ECD algorithm and comparing the before and after v-reachable sets with that which results from the use of a Min-Max control algorithm [5, pp. 530–531].

7.2.1 ECD Modification

Having specified an approximation to the value function, the ECD algorithm was developed by essentially neglecting the disturbance v. We can improve performance by adding a term to the ECD algorithm that counters the effect of that disturbance.

Once an approximation to the value function is given, we have

$$\begin{aligned} \dot{W}_0 &= f_0(\mathbf{x}, u, v) + \frac{\partial W(\mathbf{x})}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, u, v) \\ &= \frac{1}{2}(\mathbf{x}^\mathsf{T} \mathbf{Q} \mathbf{x} + u^\mathsf{T} \mathbf{R} u - \rho^2 v^\mathsf{T} v) + \frac{\partial W(\mathbf{x})}{\partial \mathbf{x}} (\mathbf{A}\mathbf{x} + \mathbf{B}u + \mathbf{E}v). \end{aligned} \tag{38}$$

Player 1, controlling u, expects Player 2, controlling v, to maximize (38). The necessary conditions for maximizing $\dot{W}_0$ with respect to v are given by

$$\frac{\partial \dot{W}_0}{\partial v} = -\rho^2 v^\mathsf{T} + \frac{\partial W(\mathbf{x})}{\partial \mathbf{x}} \mathbf{E} = \mathbf{0}^\mathsf{T}, \tag{39}$$

which yields

$$v = \frac{1}{\rho^2} \left(\frac{\partial W(\mathbf{x})}{\partial \mathbf{x}} \mathbf{E} \right)^\mathsf{T}. \tag{40}$$

Consider the time rate of change of $W(\mathbf{x})$ in Example 5.1,

$$\begin{aligned}\dot{W}(\mathbf{x}) &= \frac{\partial W(\mathbf{x})}{\partial \mathbf{x}}\dot{\mathbf{x}} \\ &= x_2\left(x_1 + \frac{1}{2}x_2\right) + u\left(x_2 + \frac{1}{2}x_1\right) + v\left(x_2 + \frac{1}{2}x_1\right),\end{aligned} \tag{41}$$

and from (40) we have

$$v = \frac{1}{1.1}\left(\frac{1}{2}x_1 + x_2\right). \tag{42}$$

Player 1 may counter the increase in $\dot{W}$ caused by the disturbance through the addition of the term

$$\begin{gathered} u_v = \frac{-1}{1.1}\left(\frac{1}{2}x_1 + x_2\right) \\ \text{or} \\ \dot{u}_v = \frac{-1}{1.1}\left(\frac{1}{2}\dot{x}_1 + \dot{x}_2\right). \end{gathered} \tag{43}$$

The final form of the ECD algorithm is now

$$\dot{u} = -\left[\frac{1}{\epsilon}\mathrm{e}^{-\alpha t} + \beta\|\hat{\mathbf{g}}\|\right]\hat{\mathbf{g}} - \left[\frac{\partial \hat{\mathbf{g}}}{\partial u}\right]^{-1}\frac{\partial \hat{\mathbf{g}}}{\partial \mathbf{x}}\dot{\mathbf{x}} + \dot{u}_v. \tag{44}$$

Figure 8 contains the initial and final v-reachable sets for the ECD algorithm, based upon equations (36) and (44), respectively, along with that of the Min-Max controller of [5, p. 531]. The addition of the term $\dot{u}_v$ has greatly reduced the size

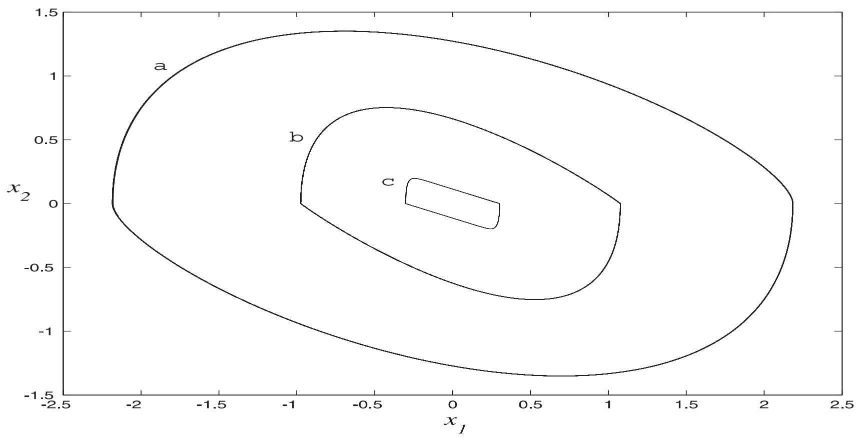


Figure 8: The v-reachable sets for (a) initial ECD, (b) final ECD, and (c) Min-Max.

of the v-reachable set for the ECD algorithm. We observe that the final form of the ECD algorithm produces a v-reachable set that is larger than that given by the Min-Max algorithm; clearly then ECD is not optimal. However, we have specified a closed-loop, on-line controller that does not require us to solve partial differential equations or two-point boundary-value problems.

8 Conclusion

In this chapter we have specified minimizing player strategies through trajectory following methods. These differential equation strategies were derived from a function that represents a trade-off between the current rate of cost accumulation and the penetration rate for the minimizing player's descent function, and eliminates the need to solve min-max problems and two-point boundary-value problems at each state. MCD and TD trajectory following algorithms were applied to a linear-quadratic differential game. We observed that a more robust strategy with respect to the uncertain, or maximizing player's, input could be obtained by introducing a singular perturbation onto the minimizing player's strategy. This yielded the SPMCD algorithm, which we found to be very stiff. Analysis of the SPMCD algorithm using well-known singular perturbation techniques led to the ECD algorithm. ECD was generally nonstiff and provided the same desirable performance characteristics as the SPMCD algorithm. The ECD algorithm is suboptimal, but the performance is comparable with game-theoretic controllers, even with the unknown disturbance v implementing a game-theoretic strategy, as seen in Figure 8.

References

[1] Chow, J. H. and Kokotovic, P. V., A Decomposition of Near Optimum Regulators for Systems with Slow and Fast Modes, *I.E.E.E. Trans. on Automatic Control*, Vol. 21, No. 5, 1976, pp. 701–705.

[2] Grantham, W. J., *A Controllability Minimum Principle*, Ph.D. dissertation, Aerospace and Mechanical Engineering, University of Arizona, Tucson, AZ, 1973.

[3] Grantham, W. J., Trajectory Following Optimization by Gradient Transformation Differential Equations, *Proc. 42nd I.E.E.E. Conf. on Decision and Control,* Maui, HI, December 2003, pp. 5496–5501.

[4] O'Malley, R. E., *Introduction to Singular Perturbations*, Academic Press, New York, 1974.

[5] Vincent, T. L. and Grantham, W. J., *Nonlinear and Optimal Control Systems*, Wiley, New York, 1997.

[6] Vincent, T. L. and Grantham, W. J., Trajectory Following Methods in Control System Design, *J. of Global Optimization*, Vol. 23, 2002, pp. 67–282.

[7] Vincent, T. L., Goh, B. S. and Teo, K. L., Trajectory Following Algorithms for Min-Max Optimization Problems, *J. of Optimization Theory and Applications*, Vol. 75, No. 3, 1992, pp. 501–519.

[8] Vincent, T. L. and Lin, Y. C., Some Performance Measures Useful in the Design of Controllers for Two-Dimensional Dynamical Systems Subject to Uncertain Inputs, *Dynamics and Control*, Vol. 5, 1995, pp. 69–98.

[9] Basar, T. and Olsder, G. J., *Dynamic Noncooperative Game Theory*, Academic Press, London, 1982.

Min-Max Guidance Law Integration

Stéphane Le Ménec
EADS/MBDA-F
78141 Vélizy-Villacoublay, France
stephane.le-menec@mbda.fr

Abstract

This chapter deals with air-to-air missile guidance law design. We model the terminal engagement of a ramjet missile with radar seeker lock on a single target (generic aircraft). We consider a realistic interception simulation with measurement errors and, more particularly, radome aberration errors. We define an extended Kalman filter to compensate in line those errors and to stabilize the guidance loop. Then, to decrease the miss distance against manoeuverable targets we implement an optimized guidance law. The guidance law we propose is based on a linear quadratic differential game linearized around the collision course using the target acceleration estimation provided by the Kalman filter. In this game, the evader control is defined around the Kalman target acceleration estimation to take into account delays and lags due to the filters we apply. The Kalman target acceleration estimation (assumed constant) is a parameter of the differential game kinematics.

1 Introduction

This study deals with guidance laws based on linear quadratic differential games in a realistic simulation model with errors such as thermal noise, radome aberrations and misalignments. Section 2 describes the simulation models and more specially the radome aberration errors in the simulation. Section 3 deals with the Kalman filter we implement to estimate and compensate in line the radome aberrations. In this way, Section 3 describes the radome aberration model we use in the Kalman filter transition matrix. Section 4 explains how we tune the guidance loop gains and the time lag constants for stability reasons. In Section 5, we discuss the way we implement a linear quadratic differential game guidance law in this three-dimensional realistic simulation. On some examples, considering the poor effectiveness of classical proportional navigation against manoeuvring targets and considering the difficulty of estimating precisely the target acceleration (to use an augmented proportional navigation guidance law), we want to take into account the target optimal evasion behaviour (min-max guidance laws) to improve the miss distance. To conclude, we address some possible extensions.

2 Simulation Model

We use a generic three-dimensional (3D) air-to-air interception simulation (Figure 1) in Matlab / Simulink to model a ramjet powered missile (6-DOF missile model) pursuing an aircraft (single target oriented simulation). This simulation is based on several block sets (subsystems). A radar seeker block set implements the seeker inner loop pointing the line of sight (LOS) direction (missile target LOS). The seeker outputs are boresight angle measurements (along y and z dish axes). The boresight angles show the angle gap between the dish axis (controlled by the seeker inner loop) and the LOS.

The guidance processor block set (Figure 2) contains a guidance filter (first-order filter in this case) converting boresight measurements into LOS rotating rate and a guidance law computing acceleration demands. The LOS rotating rate is the main input of proportional navigation (PN)-based guidance laws. The guidance processor also sets the steering guidance mode to Skid To Turn (STT) or Bank To Turn (BTT). The guidance law determines the required missile trajectory demands. The steering law, which is an interface between the guidance function and the autopilot, determines how this is achieved. Old symmetric airframe system configurations only controlled the missile incidence and the side slip angle (STT mode). The BTT missiles roll first around the x missile body axis to control the required incidence. The pure STT and BTT missiles are designed with 2-channel autopilots. Several refinements using 3-channel autopilots allow one to control the roll rate of STT

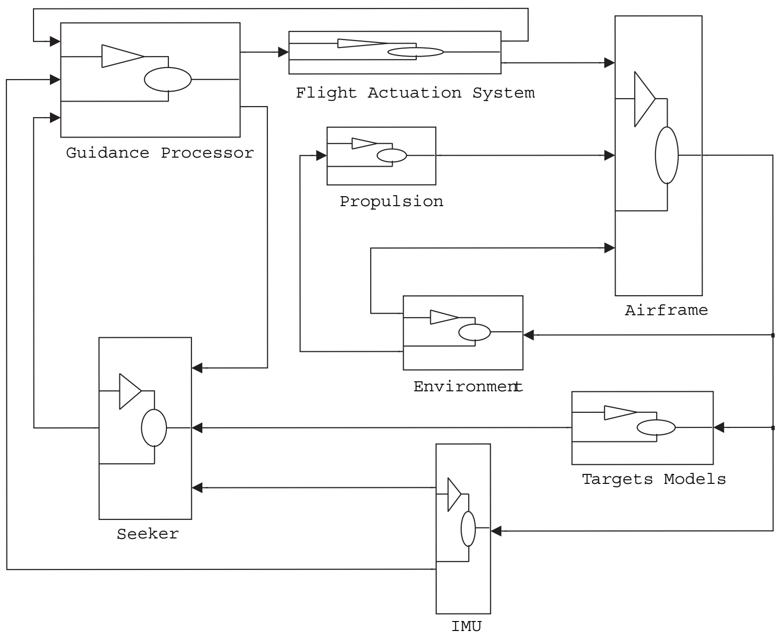


Figure 1: 6-DOF air-to-air interception simulation.

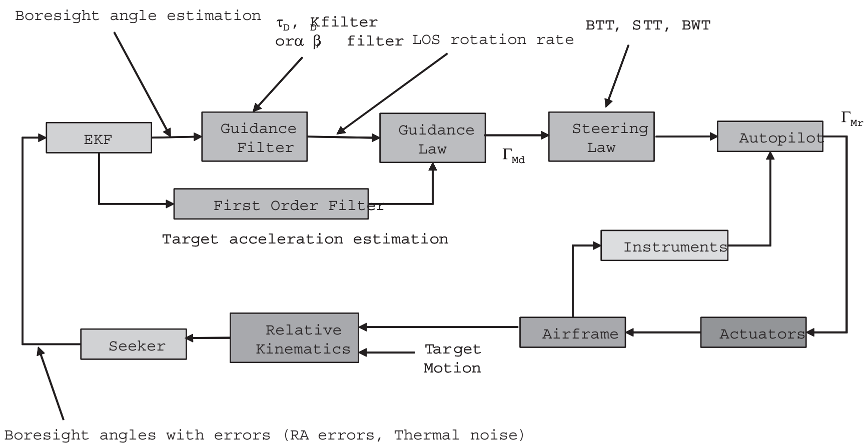


Figure 2: Guidance processor subsystem.

missiles to a constant value and to control small side slip demands in BTT (for low roll rate missiles). We call this steering mode with low roll rate intermediate between the pure STT and BTT mode Banking While Turning (BWT).

In addition, the guidance processor subsystem simulates the Free Flight Control (FFC). The FFC (also called autopilot) controls the missile to follow manoeuvre demands within prescribed limits. The FFC is a 3-channel autopilot able to follow STT, BWT, and BTT modes by applying different roll rate values (G1R). The FFC is equivalent to an STT autopilot when the roll rate is set up to 0 (BWT for intermediate roll rate values and BTT for large roll rate values).

The following subsystems: Actuators (fins controlled by the FFC function), Propulsion (ramjet engine), Atmosphere (environments), Aerodynamic and flight mechanics, and Target model and inertial measurement unit (IMU) form the last part of this simulation. We simulate realistic errors on the boresight measurements: the thermal noise and the radome aberrations (RA). Moreover, we model IMU and seeker misalignments (using biases in transition matrices) that occur in the referential transition computations. At present, the simple target model we use does not allow us to take into account realistic glint errors.

2.1 RA Errors

The thermal noise (due to antenna technology) is linked to the missile target range (large when the target is far and lower when the missile is close to the target) while the RA errors (Nesline et al. [4]) are functions of the gimbal angles and lead to problems when the LOS rotates fast. The gimbal angles (elevation and azimuth angles) describe the dish (antenna) position respect to the missile body axes. Classically, to counter the thermal noise effects, the boresight angle measurements are filtered (guidance filter) when the missile target range is large and less filtered close

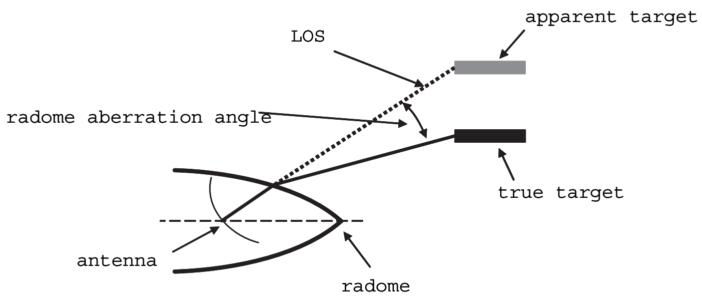


Figure 3: Radome aberration errors.

to the target (to react to target evasions). However, a missile target range variable filter is not well fitted against RA errors.

RA are due to the refraction of electromagnetic waves as they pass through the radome protecting the missile seeker (missile front part, see Figure 3). This refracted signal destabilizes the guidance loop, more significantly at higher altitudes, when the incidence lag time constant is larger. The RA cause instabilities by coupling the LOS angle measurements to the missile body dynamics through the gimbal angles. The dish, which is theoretically perfectly decoupled from the missile attitude, should keep pointing on the LOS despite missile body movements. By changing the missile attitude, the RA errors (directly linked to the dish position) change, involving new acceleration demands.

In this 3D simulation, the RA are errors on the boresight angles proportional to the elevation (angle E) and azimuth (angle A) gimbal angles. We describe the RA errors through 4 constant slopes (see equations below), 2 slopes (one in azimuth and one in elevation) per dish axis (boresight measurements along y (ε_A) and along z (angle ε_E) axes). This simulation is based on the 4 constant slopes k_{aa}, k_{ee}, k_{ae}, k_{ea}; nevertheless, some refinements involve nonconstant slopes $k_{aa}(A,\ E)$, $k_{ee}(A,\ E)$, $k_{ae}(A,\ E)$, $k_{ea}(A,\ E)$.

The RA we discuss are RA errors before precompensation achieved in seeker architectures using look-up tables resulting from seeker laboratory measurements. These tables do not entirely suppress the RA errors, because RA errors also depend on temperatures and altitudes. Moreover, data laboratory measurements involve long and expensive processes.

$$\varepsilon_{Ameasured} = \varepsilon_{Atrue} + k_{aa}.A + k_{ae}.E$$
$$\varepsilon_{Emeasured} = \varepsilon_{Etrue} + k_{ee}.E + k_{ea}.A$$

k_{aa} and k_{ee} are direct RA effects, whilst k_{ae} and k_{ea} are RA slopes describing indirect effects (boresight angle errors along the azimuth axis due to elevation dish movements and boresight angle errors along the elevation axis due to azimuth dish movements).

3 In-Line RA Compensation

Linear analysis on two-dimensional (2D) models allows us to calculate the stability domain in terms of radome slope aberration (a unique parameter in 2D simulations). These theoretical 2D results explain that the guidance loop is stable only for small RA slope values with minimum and maximum limits. The negative stability domain is often smaller than the positive one.

Therefore, the existence of RA requires the careful design of missile radomes and compensation of these errors. To increase the stability domain, we develop an extended Kalman filter (EKF) to estimate and compensate in line those RA errors. The EKF written in dish axes compensates the y and z boresight measurements and provides an estimate of the target acceleration perpendicular to the LOS. The 7-state EKF we implement is defined in dish axes as follows:

$$X = \begin{bmatrix} y \\ s \\ a_{Ty} \\ z \\ p \\ a_{Tz} \\ \delta \end{bmatrix}, \quad A_{M_2}^{\perp} = \begin{bmatrix} a_{My} \\ a_{Mz} \end{bmatrix}, \quad h = \begin{bmatrix} \varepsilon_y = y + \delta.\eta_1(A, E) \\ \varepsilon_z = z + \delta.\eta_2(A, E) \end{bmatrix}$$

$$\vec{S} = \frac{\vec{R}}{R} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \vec{V}^{\perp} = \frac{\vec{V}_{MT}}{R} - \left(\frac{\dot{R}}{R}\right)\vec{S} = \begin{bmatrix} r \\ s \\ p \end{bmatrix}$$

$$\vec{A}_M^{\perp} = \frac{\vec{A}_M}{R} - \left(\frac{\vec{A}_M}{R}.\vec{S}\right)\vec{S} = \begin{bmatrix} a_{Mx} \\ a_{My} \\ a_{Mz} \end{bmatrix}, \quad \vec{A}_T^{\perp} = \begin{bmatrix} a_{Tx} \\ a_{Ty} \\ a_{Tz} \end{bmatrix},$$

where X is the Kalman state vector, $A_{M_2}^{\perp}$ the control, and h the seeker measurements. $\eta_1(A, E)$ and $\eta_2(A, E)$ are functions of azimuth and elevation angles. Both expressions $\delta.\eta_1$ and $\delta.\eta_2$ define the RA errors on y and z dish axes, respectively. $\vec{S}$ is the unit LOS vector and R is the missile target range. $\vec{V}^{\perp}$ is the unit LOS speed vector perpendicular to LOS, $\vec{A}_M^{\perp}$ is the missile acceleration vector perpendicular to LOS and $\vec{A}_T^{\perp}$ is the target acceleration vector perpendicular to LOS. Both acceleration vectors are normalized by range. ε_y and ε_z are angles, and y and z are metric distances ($\vec{S}$ coordinates). As $\vec{S}$ is the LOS vector normalized by range R (unit vector) and as boresight angles are small angles, the tangent function approximation makes y and z similar to angles.

In this Kalman filter, the RA errors are described through one unique constant parameter δ ($\dot{\delta} = 0$ in the EKF transition matrix) to have a low dimension Kalman filter. We use a low dimension Kalman filter to ensure rapid time convergence and for observability reasons. The RA model we implement in the Kalman filter uses the radome symmetry around the x missile body axis. This model is called

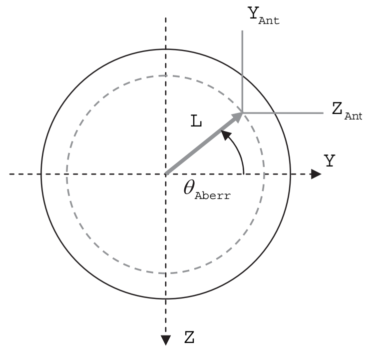


Figure 4: Radome in body axes (rear view), body referential Z down oriented.

the geometric RA model (in simulation, the physical RA errors are computed with the 4-slope model). In the geometric model, the RA errors are proportional to the radial distance L defined by the projection of the dish unit vector (coordinates (1, 0, 0) in dish coordinate frame) in the missile body referential (Y, Z) plan (see Figure 4). The dish unit vector components are $(X_{Ant}, Y_{Ant}, Z_{Ant})$ in the missile body coordinate frame. Notice that in both RA models (geometric and 4 slopes) the boresight errors are defined according to the dish pointing direction and not according to the true target LOS, which is equivalent to the dish position plus the true boresight angles, but the boresight angles are much smaller than the gimbal angles.

$$L = \sqrt{Y_{Ant}^2 + Z_{Ant}^2}$$

$$aberr = \delta.L, \quad \theta_{aberr} = \arctan \frac{Z_{Ant}}{Y_{Ant}}$$

$$\begin{cases} Y_{aberr} = aberr.\cos\theta_{aberr} \\ Z_{aberr} = aberr.\sin\theta_{aberr} \end{cases}.$$

After projection in dish, Y_{aberr} and Z_{aberr} give the RA errors on the EKF boresight measurements. The transition from the body referential to the dish referential is described by two rotations, a first elevation (angle E) rotation and then an azimuth (angle A) one. E is up oriented. A is positive on the right. In the body axis as in the dish axis, the z axes are down oriented. In the body axis as in the dish axis, the y axes are right oriented. In this way, we express the geometric RA boresight errors as follows:

$$L = \sqrt{(\sin A)^2 + (\cos A \sin E)^2}$$

$$\begin{cases} \varepsilon_y = y + \delta.\eta_1(A, E) \\ \varepsilon_z = z + \delta.\eta_2(A, E) \end{cases}$$

$$\begin{cases} \delta.\eta_1(A, E) = Y_{aberr} \cos A + Z_{aberr} \sin E \sin A \\ \delta.\eta_2(A, E) = Z_{aberr} \cos E \end{cases}$$

$$\begin{cases} \delta.\eta_1(A, E) = \delta \frac{\sqrt{1-\cos^2 A \cos^2 E}}{\sqrt{1+(\frac{\sin E}{\tan A})^2}} \cos A \cos^2 E \\ \delta.\eta_2(A, E) = -\delta \frac{\sqrt{1-\cos^2 A \cos^2 E}}{\sqrt{1+(\frac{\sin E}{\tan A})^2}} \frac{\cos E \sin E}{\tan A} \end{cases} .$$

The minus sign in the RA error expression (elevation dish axis) is due to the fact that the z axes are down oriented.

RA errors are described in different ways in the EKF and in the simulation. Figure 5 represents RA errors with the geometric model and with the 4-slope model. $\delta = -0.03$ corresponds almost to 3% of RA errors. The difference between the RA models plus the difficulty in estimating accelerations explain why we obtain imperfect target acceleration estimations.

Although the EKF directly provides LOS rotating rate estimations ($\vec{V}^\perp$ vector), we compute the LOS rotating rate using the compensated boresight angles through the guidance filter. The gain and the lag time constant of the guidance filter are easy to tune. By filtering twice using the EKF and the guidance filter we extend the stability domain for large test scenarios.

4 Guidance Loop Stability

Despite calibration and estimation techniques, there always remain some RA errors which classically require one to increase the seeker inner loop lag time constant and hence cause a larger miss distance. In a similar way, we tune the guidance

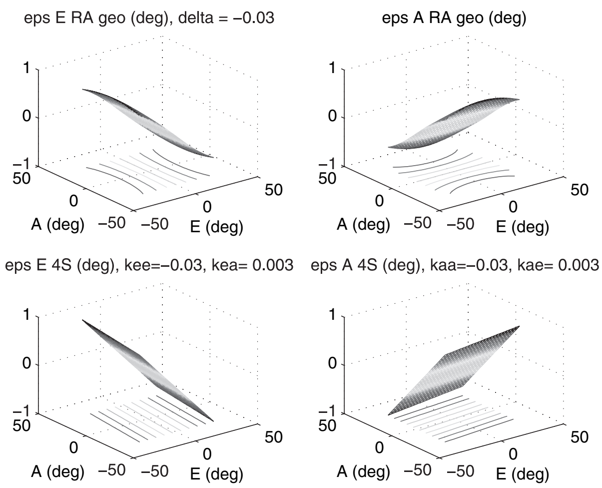


Figure 5: Geometric (EKF) and 4 slopes (simulation) RA errors on boresight measurements.

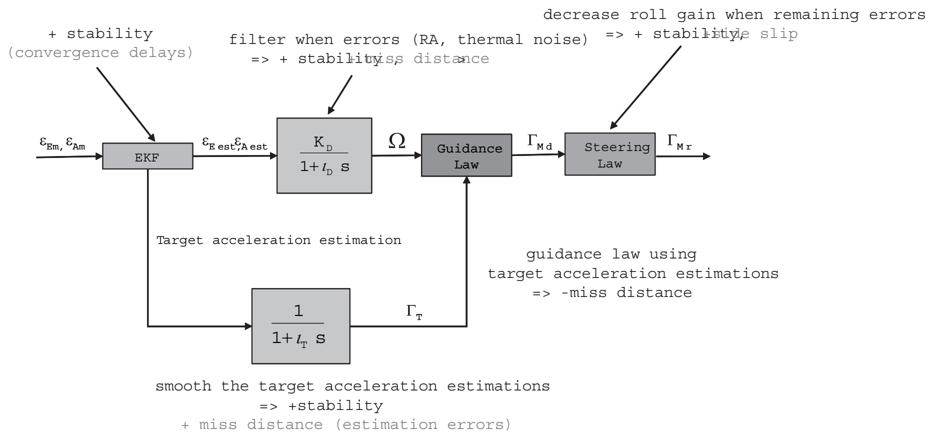


Figure 6: Gains and lag time constants in the guidance loop.

filter to smooth the LOS rotating rate computation and the guidance processor implements a classical PN reacting slowly to LOS movements. The guidance law gain is set to small values between 3 and 4 (according to the target missile range).

Moreover, the ramjet missiles are BWT missiles with high roll gain values (BWT/BTT) designed to avoid large side slip amounts. These missiles with small wings have larger pitch acceleration capabilities (along the z positive body axis) than in yaw (y body axis). Moreover, engine constraints prevent us from controlling side slip most of the time. To achieve accelerations in the right direction in the case of pure BTT autopilots or to decrease the side slip demands in the case of 3-channel rolling autopilots we use large roll gains (compare to pitch and yaw) increasing the error autopilot sensitivity (destabilized first in roll). See Yueh et al. [5].

Figure 6 explains the effects of different gains and lag time constants on the guidance loop stability.

5 Miss Distance Improvement

PN has a poor effectiveness against manoeuvrable targets. The guidance laws using target acceleration estimates as the Augmented Proportional Navigation (APN) require precise estimates to really improve the miss distance. For bad estimates, the results are sometimes worse than those obtained using PN. The EKF target acceleration estimates are usable only after the EKF convergence and require one to apply a first-order filter against residual noise. The EKF initialization values set a trade-off between convergence duration and precision. We design an EKF converging in less than one second accepting some bias on the target estimations and we filter the target acceleration estimates using a first-order filter. However, we take into account the lag and the bias on the target acceleration estimates by using a Linear Quadratic Differential Game guidance law (LQDG) with an evader playing around the estimates (assumed constant). See Anderson [1]. Nevertheless, the EKF estimates quality differs between scenarios with high and low RA errors.

We use the linear quadratic min-max guidance law (LQDG) described for 2D models in Ben-Asher et al. [2] and Gutman [3]. The kinematics is linearized around the initial collision course and takes into account the autopilot first-order time constant (i.e., the pursuer controls its normal acceleration via a first-order lag).

$$\dot{X} = AX + Bu + D(w + w_0)$$
$$X = \begin{bmatrix} y \\ \dot{y} \\ \Gamma_M^{\perp} \end{bmatrix} A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{T} \end{bmatrix} B = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{T} \end{bmatrix} D = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$
$$\min_u \max_w J = \frac{b}{2} y^2(t_f) + \frac{1}{2} \int_{t_0}^{t_f} [u^2(t) - \gamma^2 w^2(t)] dt,$$

where u, the pursuer control is the missile acceleration demand perpendicular to LOS. $\Gamma_M^{\perp}$ is the achieved pursuer acceleration (perpendicular to LOS). T is the autopilot lag time constant. w is the evader control around the estimate w_0, and w_0 is a parameter of the linearized model. The state variable y is the miss distance perpendicular to LOS (see Figure 7). The x and y axes define the 2D collision course coordinate frame corresponding in 3D to the dish referential. b and γ are weights in the cost function, respectively, on the terminal miss distance and on the evader control. γ describes the evader manoeuvrability with respect to the pursuer manoeuvrability that we want to take into account. In this case, we consider γ values representing the evader control which consists in acceleration variations around w_0. We choose b large enough to reduce the miss distance but not too large to avoid large terminal guidance gain values. We accept reasonable $y(t_f)$ values.

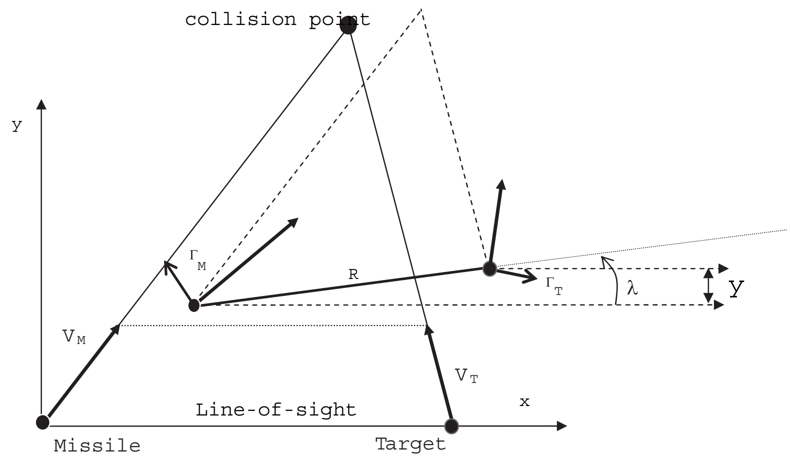


Figure 7: 2D collision course triangle.

Too large guidance gains can destabilize the guidance loop. This type of optimal guidance law has a variable guidance gain. When the time to go is large, the gain is almost equal to the PN guidance gain. When the time to go decreases, the optimal guidance gain is able to grow to infinite values if we require miss distances equal to zero ($b \to \infty$ in the criterion J).

This type of criterion J with a penalty term on controls is a classical way to constrain the maximal admissible control values when the controls are unbounded in the dynamics. This kinematics seems suitable for a BTT air-to-air missile. The BTT steering law is designed to avoid acceleration saturations, and the type of targets we consider are less manoeuvrable than the missile.

We have applied this 2D guidance law in a similar way in the 3D simulation we presented in the first section. In the 3D simulation, the (perpendicular to LOS) missile acceleration demand is computed as follows:

$$\vec{\Gamma}_{M_d} = NV_c\vec{\Omega} \wedge \vec{i}_{MT} + \frac{N}{2}\vec{\Gamma}_T^{\perp} + C\vec{\Gamma}_M^{\perp},$$

where N and C are optimal guidance gains (N and C are functions depending on the parameters, time to go, T, γ, and b), V_c is the closing velocity, $\vec{\Omega}$ is the measured LOS rotating vector, $\vec{i}_{MT}$ is the measured unit LOS vector, $\vec{\Gamma}_T^{\perp}$ is the estimated target acceleration perpendicular to LOS, and $\vec{\Gamma}_M^{\perp}$ is the achieved missile acceleration perpendicular to LOS. $\vec{\Gamma}_T^{\perp}$ is similar to $\vec{A}_T^{\perp}$ computed (in dish) by the Kalman filter, except that $\vec{A}_T^{\perp}$ is normalized by R. The EKF estimates a_{Ty} and a_{Tz}. a_{Tx} is computed as follows:

$$a_{Tx} = \sqrt{1 - a_{Ty}^2 - a_{Tz}^2}.$$

For more convenience, $\vec{\Gamma}_{M_d}$ is computed in missile target referential (LOS referential, $\vec{i}_{MT} = (1, 0, 0)$ and the x component of $\vec{\Gamma}_{M_d}$ is 0). For this reason, all input vectors ($\vec{\Omega}$, $\vec{\Gamma}_T^{\perp}$, and $\vec{\Gamma}_M^{\perp}$) are first computed in missile target referential. Then, $\vec{\Gamma}_{M_d}$ is transformed in missile body referential to control the FFC.

As the simulation model is highly nonlinear, we mainly analyze the min-max guidance law interest by running Monte Carlo simulations. The parameters are defined according to the following process:

- We first define a set of representative test scenarios (with manoeuvrable targets) and the RA stability domain we require (seeker designer constraints).
- Then, we set the Q (state noise matrix), R (measurement noise), and P_0 (initial covariance) EKF matrices according to the convergence duration (less than 1 sec.) that we specify (linked to estimation precision).
- We fix the guidance filter (K_D, τ_D) parameters according to average remaining errors.
- We set γ to take into account the bias on the estimated target acceleration we have.

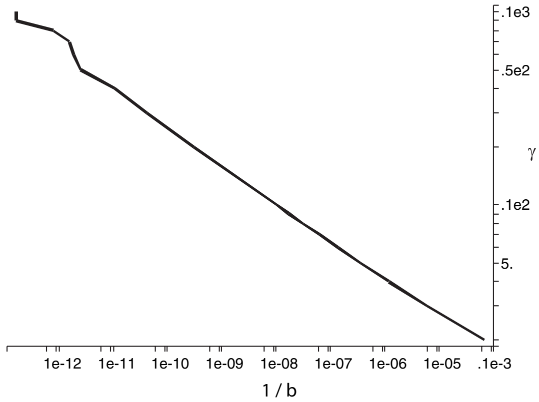


Figure 8: Trade-off between miss distance and optimal evader acceleration capability.

- We tune b as large as possible following the $(\gamma,\, 1/b)$ singularity curve (see Figure 8). This figure describes the trade-off between miss distance and optimal evader acceleration we have to take into account. The possible trades-off are above the $(\gamma,\, 1/b)$ line (for more details see Ben-Asher et al. [2]).
- T is defined by the FFC lag time constant.
- The FFC roll gain is taken as small as possible staying inside the side slip value constraints (or chosen as large as possible until the guidance loop is stable). Some refinements involve variable G1R gains with a steering law policy depending on parameters such as estimated side slip, time to go, and total acceleration demands. In this chapter, we consider constant G1R values.

We compare the maximum miss distances we obtain with the min-max guidance law (LQDG) and with PN. By taking into account the target acceleration and the missile lag time constant, in accordance with classical results the missile guidance law demands are smaller than those for using PN close to the target (see Figure 9). PN, which considers only the LOS rotating rate, controls larger terminal missile accelerations than APN and LQDG in the presence of manoeuvring targets. Therefore, LQDG compared to PN decreases the RA error variations, making the Kalman estimations easier and stabilizing the LOS measurements. Figures 10 and 11 describe results on the same scenario (constant target acceleration $\simeq 4.3$ g) with PN and with LQDG, respectively, for several RA δ parameters between -0.1 and 0.1 (Monte Carlo runs). For this comparison, to be able to explore several RA error amounts, we use the geometric RA model for the EKF and in place of the physical RA simulation model. Of course, as the RA model is the same in the simulation and in the EKF, the Kalman filter runs better and is able to compensate larger RA δ parameter values. We remark on the top right figures (Figures 10 and 11) that the δ parameter is more difficult to estimate when PN is used rather

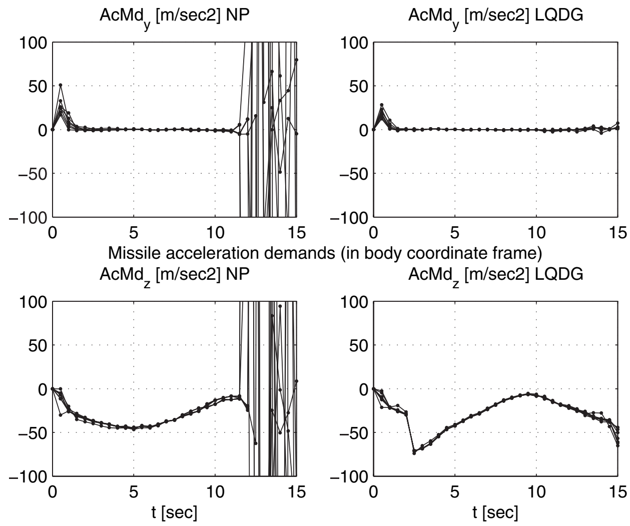


Figure 9: NP and LQDG missile acceleration demands ($-0.1 \leq \delta \leq 0.1$ RA geometric model, $K_D = 10$, $\tau_D = 0.05$ sec., $\tau_T = 0.2$ sec., G1R $= 10$ (BTT).

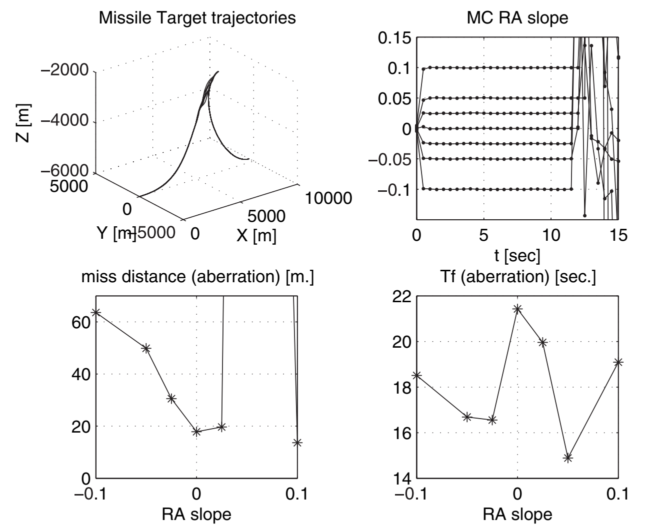


Figure 10: PN miss distance sensitivity to RA (RA geometric model, $K_D = 10$, $\tau_D = 0.05$ sec., $\tau_T = 0.2$ sec., G1R $= 10$ (BTT), PN gain $= 4$).

than LQDG. Therefore, the bad RA error estimates destabilize the LOS and imply large miss distances. This effect increases when the target acceleration increases. With LQDG, the missile acceleration demands are smaller, the EKF runs well, the LOS is stabilized, and the miss distances are smaller than 5 meters. Figure 12 shows

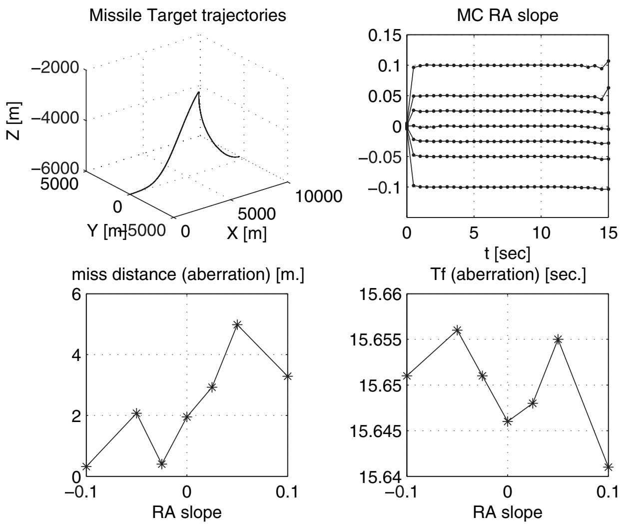


Figure 11: LQDG miss distance sensitivity to RA (RA geometric model, $K_D = 10$, $\tau_D = 0.05$ sec., $\tau_T = 0.2$ sec., G1R = 10 (BTT)).

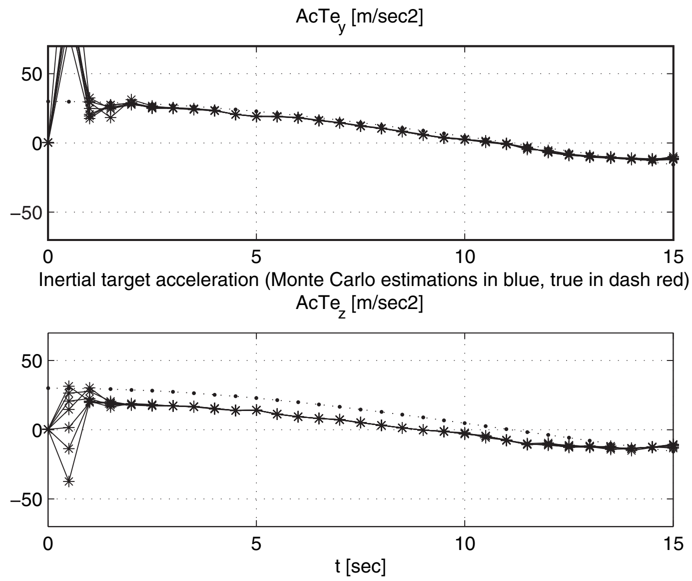


Figure 12: EKF target acceleration estimates (LQDG, $-0.1 \leq \delta \leq 0.1$ RA geometric model, $K_D = 10$, $\tau_D = 0.05$ sec., $\tau_T = 0.2$ sec., G1R = 10 (BTT).

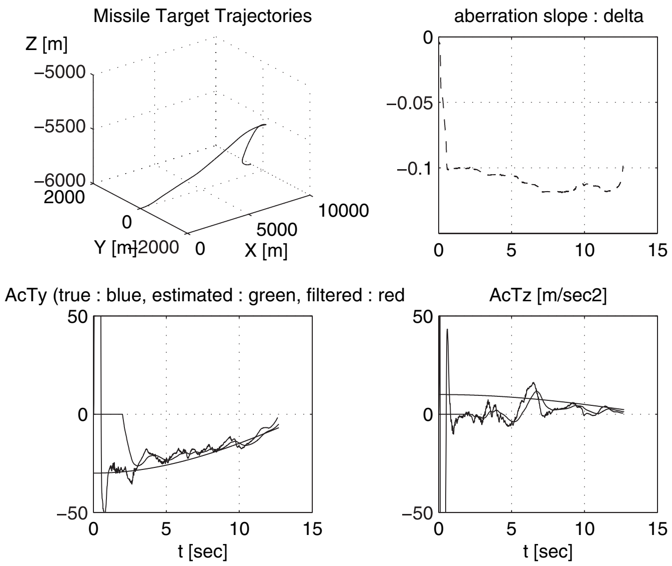


Figure 13: LQDG, $\gamma = 4$, $b = 10e^6$, 4-slope RA model, $k_{ee} = k_{aa} = -0.1$, $k_{ea} = k_{ae} = 0.01$, $K_D = 1$, $\tau_D = 0.5$ sec., $\tau_T = 0.5$ sec., G1R = 1 (BWT).

the target acceleration estimates we obtain on the Monte Carlo simulations with LQDG. We note some biases resulting from EKF settings and misalignment errors. τ_T is the lag time constant of the first-order filter we apply to smooth the EKF target acceleration estimates.

Notice that classically the larger miss distances obtained using PN rather than APN are due to saturations and lag time constants in the autopilot. In the examples here, FFC saturations do not play a major role. The PN miss distances are due to instabilities resulting from large terminal missile acceleration demands in the presence of residual RA errors.

Figure 13 shows an LQGD case with the 4-slope RA model to simulate physical RA errors. We decrease the roll gain to control the missile in BWT and stabilize the guidance loop. The top right figure shows the δ parameter the EKF uses to identify the 4-slope RA errors. We also remark that the δ parameter is no longer constant.

Moreover, by changing the γ parameter to ∞, the optimized guidance law we test is close to an optimal APN considering a missile first-order lag, because the optimal evader control is then equal to zero. We run the same scenario with different γ values (see Figure 14). For these tests, we use the 4-slope RA model in the simulation. We note an optimal value of $\gamma = 2.5$ leading to a miss distance less than 5 meters, compared to 14 meters when not using the game representation ($\gamma = \infty$). Those results show how interesting it is to consider a min-max guidance law.

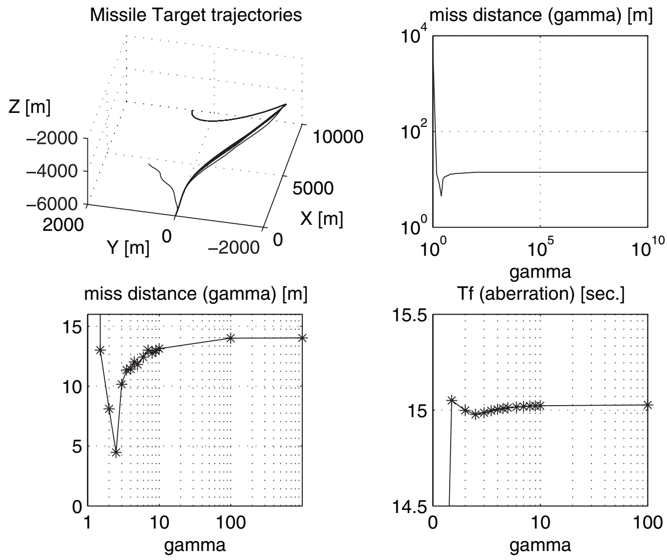


Figure 14: LQDG miss distance sensitivity to γ (4-slope RA model, $k_{ee} = k_{aa} = -0.03$, $k_{ea} = k_{ae} = 0.003$, $K_D = 1$, $\tau_D = 0.5$ sec., $\tau_T = 2$ sec., G1R = 1 (BWT), $b = 10e^6$).

6 Conclusion

This study proposes a guidance law based on a linear quadratic differential game to take into account imperfect estimates we obtain using Kalman filtering techniques (EKF). The separation principle, which consists of optimizing the guidance law apart from the estimation process, leads to difficulties in the presence of measurement errors. Therefore, we address the problem of guidance loop stability in the presence of radome aberration compensation. With a min-max guidance law partially using the target acceleration estimates we try to reduce the miss distance against manoeuvrable targets on a large RA domain. By taking into account target acceleration estimates we decrease the guidance law acceleration demands, especially close to the target (compared to PN). By decreasing the missile terminal acceleration demands, we decrease the RA error variations (remaining errors after compensation) and we stabilize the LOS.

Future works could try to estimate the misalignments errors with the EKF and then increase the measurement accuracy. Other comparisons between PN and min-max guidance laws could address weaving targets.

References

[1] Anderson G.M. (1981) Comparison of Optimal Control and Differential Game Intercept Missile Guidance Laws. In *Journal of Guidance and Control*, Vol. 4, No. 2, March–April 1981.

[2] Ben-Asher J.Z. and Yaesh I. (1998) Advances in Missiles Guidance Theory. In *AIAA, Progress in Astronautics and Aeronautics*, Paul Zarchan Editor in Chief, Reston, Virginia, Vol. 180.

[3] Gutman S. (2005) Applied Min-Max Approach to Missile Guidance and Control. In *AIAA, Progress in Astronautics and Aeronautics*, Paul Zarchan Editor-in-Chief, Reston, Virginia, Vol. 209.

[4] Nesline F.W. and Zarchan P. (1984) Radome Induced Miss Distance. In *AIAA GN&C conference, Aerodynamically Controlled Homing Missiles*, 84–1845.

[5] Yueh W.R. and Ching-Fang L. (1984) Bank-To-Turn Guidance Performance Analysis With In-Flight Radome Error Compensation. In *AIAA GN&C conference, Aerodynamically Controlled Homing Missiles*, 84–1889.

Agent-Based Simulation of the N-Person Chicken Game

Miklos N. Szilagyi
Department of Electrical and Computer Engineering
University of Arizona
Tucson, AZ 85721, USA
szilagyi@ece.arizona.edu

Abstract

We report computer simulation experiments using our agent-based simulation tool to model the multi-person Chicken game. We simplify the agents according to the Pavlovian principle: their probability of taking a certain action changes by an amount proportional to the reward or penalty received from the environment for that action.

The individual agents may cooperate with each other for the collective interest or may defect, i.e., pursue their selfish interests only. Their decisions to cooperate or defect accumulate over time to produce a resulting collective order that determines the success or failure of the public system. After a certain number of iterations, the proportion of cooperators stabilizes to either a constant value or oscillates around such a value.

The payoff (reward/penalty) functions are given as two curves: one for the cooperators and another for the defectors. The payoff curves are functions of the ratio of cooperators to the total number of agents. The actual shapes of the payoff functions depend on many factors. Even if we assume linear payoff functions, there are four parameters that are determined by these factors.

The payoff functions for a multi-agent Chicken game have the following properties. (1) Both payoff functions increase with an increasing number of cooperators. (2) In the region of low cooperation the cooperators have a higher reward than the defectors. (3) When the cooperation rate is high, there is a higher payoff for defecting behavior than for cooperating behavior. (4) As a consequence, the slope of the D function is greater than that of the C function and the two payoff functions intersect. (5) All agents receive a lower payoff if all defect than if all cooperate.

We have investigated the behavior of the agents under a wide range of payoff functions. The results show that it is quite possible to achieve a situation where the enormous majority of the agents prefer cooperation to defection.

The Chicken game of using cars or mass transportation in large cities is considered as a practical application of the simulation.

Key words. Agent-based simulation, cooperation, Chicken game, Prisoner's Dilemma.

1 Introduction

The Chicken game is an interesting social dilemma. When two people agree to drive their cars toward each other, each has two choices: drive straight ahead or swerve. If both swerve (a mutually cooperative behavior), they both receive a certain reward R. If both go ahead (mutual defection), they are both severely punished (P). The dilemma arises from the fact that if one of them swerves (chickens out) but the other does not, then the cooperator receives a sucker's payoff S while the defector gets a high reward for following the temptation T. In this game $P < S < R < T$.

In the Prisoner's Dilemma game (Poundstone [6]), defection dominates cooperation: regardless of what the other player does, each player receives a higher payoff for defecting behavior (D) than for cooperating behavior (C), in spite of the fact that both players are better off if they both cooperate than if both defect ($S < P < R < T$). In the Chicken game there is no domination. The D-choice yields a higher payoff than the C-choice if the other person chooses C but the payoff is less for D than for C if the other player's choice is D. Mutual defection is the worst outcome for both players. As a consequence, cooperation is more likely in the Chicken game than in the Prisoner's Dilemma. Evidently, cooperation grows with the severity of punishment for mutual defection. A good example of the Chicken game is the Cold War when mutual defection would have led to a nuclear catastrophe.

The two-person Chicken game has received some attention in the literature ([1], [4], [6], [8]). However, realistic social situations always have much more than two participants. Therefore, in this chapter, the problem of the multiperson Chicken game is presented by computer simulation.

2 The Model

People who participate in a social system are extremely complicated. Their actions are determined by their personalities as well as by the stimuli they receive from the environment. (We use the term "personality" in the sense of decision heuristics in a repeated "game.") In this study we will simplify them according to a very simple principle: their probability of taking a certain action changes by an amount proportional to the reward or penalty received from the environment for that action. Such a personality is based on Pavlov's experiments and Thorndike's ([13]) law: if an action is followed by a satisfactory state of affairs, then the tendency of the agent to produce that particular action is reinforced.

Such simplified "persons" are usually called *agents* to avoid the temptation to draw too far-reaching conclusions from computer simulations. Nevertheless, the results will show some tendencies at least qualitatively. (An additional benefit is that the agents are neutral; therefore, we can get rid of the awkward usage of his/hers, etc.)

The individual agents may cooperate with each other for the collective interest or may defect, i.e., pursue their selfish interests. Their decisions to cooperate or

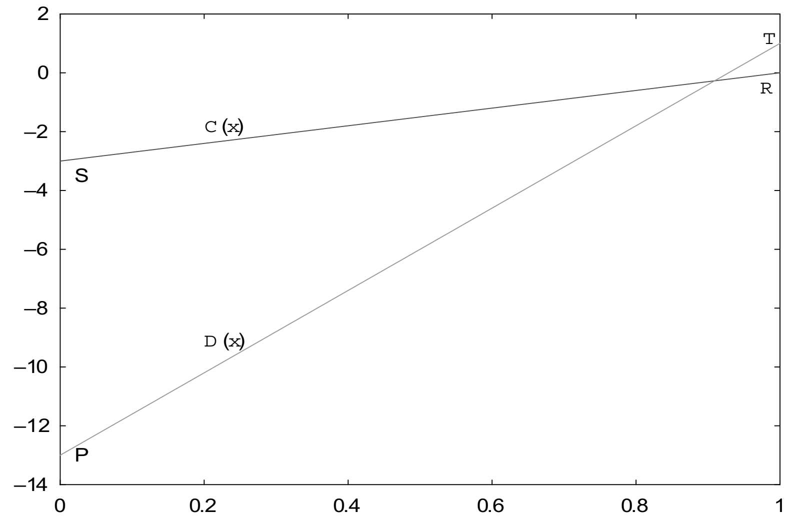


Figure 1: Reward/penalty functions for defectors (D) and cooperators (C). The horizontal axis (x) represents the ratio of the number of cooperators to the total number of agents; the vertical axis is the reward/penalty provided by the environment. In this case, $D(x) = -13 + 14x$ and $C(x) = -3 + 3x$.

defect will accumulate over time to produce a resulting collective order that will determine the success or failure of the society.

Let us first consider the payoffs for both choices (Figure 1). We will extend the notations introduced above. Point P now corresponds to the punishment when all agents defect, point R is the reward when all agents cooperate, T is the temptation to defect when everybody cooperates, and S is the sucker's payoff for cooperating when everyone defects. The horizontal axis represents the number of cooperators related to the total number of agents. We will assume that the payoffs are linear functions of this ratio x. Accordingly, we connect by straight lines point S with point R (cooperators' payoff function C) and point P with point T (defectors' payoff function D). Thus, the payoff to each agent depends on its choice and on the distribution of other players among cooperators and defectors. The payoff function $C(x)$ is the same for all cooperators and $D(x)$ is the same for all defectors.

We immediately note five features of these payoff functions: (1) Both payoff functions increase with an increasing number of cooperators. (2) In the region of low cooperation the cooperators have a higher reward than the defectors. (3) When the cooperation rate is high, there is a higher payoff for defecting behavior than for cooperating behavior. (4) As a consequence, the slope of the D function is greater than that of the C function and the two payoff functions intersect. (5) All agents receive a lower payoff if all defect than if all cooperate. This is a typical case of a multi-agent Chicken game (Schroeder [7]).

The actual values of P, R, S, and T depend on many factors determined by the problem underlying the dilemma. If the parameters are selected appropriately, the simulation will exhibit behavior that is close enough to the behavior of real people. However, even if we do not know the exact values, the simulation will still yield interesting qualitative conclusions.

We will use our agent-based model developed for simulated social and economic experiments with a large number of decision-makers operating in a stochastic environment (Szilagyi and Szilagyi [11]). In this model the participating agents are described as stochastic learning cellular automata, i.e., as combinations of cellular automata (Wolfram [16], Hegselmann and Flache [3]) and stochastic learning automata (Narendra and Thathachar [5], Flache and Macy [2]). Stochastic learning means that behavior is not determined but only shaped by its consequences, i.e., an action of the agent will be more probable but still not certain after a favorable response from the environment.

Szilagyi and Szilagyi ([11]) describe the model in detail. We will only briefly explain its most important features here.

The simulation environment is a two-dimensional array of the participating agents. Its size is limited only by the computer's virtual memory. The behavior of a few million interacting agents can easily be observed on the computer's screen.

There are two actions available to each agent, and each agent must choose between cooperation and defection. Each agent has a probability distribution for the two possible actions. The agents as stochastic learning automata take actions according to their probabilities updated on the basis of the reward/penalty received from the environment for their previous actions. The updating occurs simultaneously for all agents. The cooperators and the defectors are distributed randomly over the array.

A linear updating scheme is used: the change in the probability of choosing the previously chosen action again is proportional to the reward/penalty received from the environment (payoff curves). Of course, the probabilities always remain in the interval between 0 and 1.

The updated probabilities lead to new decisions by the agents that are rewarded/penalized by the environment. With each iteration, the software tool draws the array of agents in a window on the computer's screen, with each agent in the array colored according to its most recent action. In an iterative game, the aggregate cooperation proportion changes in time, i.e., over subsequent iterations. After a certain number of iterations, this proportion stabilizes to either a constant value or oscillates around such a value. The experimenter can view and record the evolution of the society of agents as it changes in time. The outcome of the game depends on the personalities of the agents. In this work we assume a Pavlovian personality for all agents.

We will assume that the agents interact with all other agents simultaneously. It means that in a cellular automaton formulation the neighborhood extends to the entire array of agents.

3 Simulation

Throughout this work we will assume that the total number of agents is 10,000 and the initial ratio of cooperators is 50%.

Even if we use only linear payoff functions, we have four parameters (P, S, R, and T). Naturally, the results of the simulation strongly depend on the values of these parameters. This is a four-dimensional problem that can be handled by a systematic variation of the parameters.

We start with the payoff functions shown in Figure 1 ($S = -3$, $P = -13$, $R = 0$, $T = 1$). The result of the simulation in this case is that after a relatively small number of iterations the ratio of cooperators starts oscillating about the stable value of 76%. The actual number of iterations needed for the oscillation to start depends on the learning rate used in the probability updates.

Let us first vary one parameter at a time. If we increase the value of T, the ratio of cooperators steadily decreases because the defectors receive a higher reward when there are more cooperators.

Increasing the value of R drastically increases the cooperation ratio because a larger and larger part of the C function is above the D function. At $R = 0.4$ we already reach a point when no one defects anymore.

Changing the value of S makes a big difference, too. When $S = P = -13$, the result is that the ratio of cooperators oscillates around 48%. This is the limiting case when the Chicken game becomes a Prisoner's Dilemma. The simulations result in this case that the number of cooperators always decreases with time and can reach the original 50% only in the best case when the two payoff functions are very close to each other. This is in perfect agreement with the analytical study of the multi-agent Prisoner's Dilemma game by Szidarovszky and Szilagyi ([9]).

As we move the value of S up from this limiting case, the ratio of cooperators steadily grows and at $S = 0$ (the unrealistic case of constant payoff for the cooperators) reaches 93%.

If we move point P up, the situation for the defectors steadily improves and the ratio of cooperators decreases to 43% at the limiting case of $S = P = -3$. Compare this result with the previous one when the two coinciding points were much lower.

Let us now keep P and S without change and move both T and R but in such a way that their difference does not change. We start at $R = -1$, $T = 0$ and move both points up. The ratio of cooperators changes from 0.70 to 0.83 when $R = 0.35$ and $T = 1.35$. A further small increase ($R = 0.38$, $T = 1.38$) results in total cooperation.

If we keep R and T without change and move both S and P up so that their difference does not change, the ratio of cooperators increases again and reaches 91% at the limiting case of $S = 0$, $P = -10$.

Finally, if we move P up so that the slope of the D function remains constant (that moves point T up as well), the ratio of cooperators will decrease drastically and reaches 18% in the limiting case of $P = -3$, $T = 11$. Compare this with the result above when the value of T remained 1.

For rational players the intersection point (x^*) of the two payoff functions is a Nash equilibrium. Indeed, when $x < x^*$, the cooperators get a higher payoff than the defectors; therefore, their number increases. When $x > x^*$, the defectors get a higher payoff and the cooperators' number decreases. This is, however, not true for the more realistic Pavlovian agents. In this case the relative situation of the two payoff curves with respect to each other does not determine the outcome of the dilemma. It is equally important to know the actual values of the payoffs. For the payoff functions presented in Figure 1, the solution is 76% cooperation while $x^* = 91\%$ for this case.

If we shift the horizontal axis up and down, then x^* does not change but the solutions do (Szilagyi [10]). For the payoff functions shown in Figure 1, the following cases are possible:

(a) Both curves are positive for any value of x. In this case the cooperators and the defectors are all always rewarded for their previous actions; therefore, they are likely to repeat those actions. As a result, little change occurs in the cooperation ratio, especially when the rewards are large. The number of cooperators remains approximately equal to that of the defectors.
(b) The $C(x)$ curve is entirely positive but $D(x)$ changes sign from negative to positive as the value of x grows. In this case the cooperation ratio changes from 0.5 to almost 1 as we shift the horizontal axis up, i.e., for smaller rewards the cooperation ratio is larger.
(c) When both $C(x)$ and $D(x)$ change signs, the cooperation ratio changes from near 1 to 0.76 as we shift the horizontal axis up, i.e., for smaller rewards the cooperation ratio is smaller. Maximum cooperation occurs when $S = 0$, i.e., the environment is neutral to cooperators during maximum defection.
(d) The $C(x)$ curve is entirely negative but $D(x)$ changes sign from negative to positive as the value of x grows. The solution oscillates around a stable equilibrium that grows with the rewards.
(e) Both $C(x)$ and $D(x)$ are negative for all values of x (punishment for any action). The solution oscillates around a stable equilibrium. As the punishments grow, all actions tend to change at each iteration and the solution approaches 50%.

As we see, the best solutions appear when the $D(x)$ function changes sign with increasing x, i.e., when the defectors are rewarded when there are few of them, and punished when there are many of them, which is quite a realistic situation.

If we change more than one parameter at the same time, different solutions will appear, but the basic trends remain the same.

A perturbation to the initial state may cause the system to evolve into a different future state within a finite period of time or cause no effect at all. For the payoff functions of Figure 1, the result is the same if we start with 10%, 50%, or 90% cooperators.

The reason for this is explained by Szilagyi and Szilagyi ([12]) as follows. When the cooperators receive the same total payoff as the defectors, i.e.,

$$xC(x) = (1 - x)D(x), \tag{1}$$

an equilibrium occurs. This may happen if $C(x)$ and $D(x)$ are both negative or both positive. In the first case, a small number of cooperators are punished severely and a large number of defectors are punished slightly. This leads to a stable equilibrium at this point. In the second case, a large number of cooperators are rewarded slightly and a small number of defectors are rewarded greatly. This point corresponds to an unstable equilibrium.

If $C(x)$ and $D(x)$ are both linear functions of x, then equation (1) is a quadratic equation that has up to two real solutions. When both equilibria are present, the stable equilibrium is the solution at any value of the initial cooperation probability that is smaller than the value of the unstable equilibrium. For Figure 1, 76% cooperation is the stable equilibrium and total cooperation is the unstable equilibrium. As any initial condition falls below the unstable equilibrium, the result is always 76%.

4 Conclusion

The multi-agent Chicken game has nontrivial but remarkably regular solutions. The experiments performed with our simulation tool show interesting new results. A possible application of these results is the problem of public transportation.

Today's economy is based on the automobile. Every day, hundreds of millions of people use their cars to visit a remote place or just to go to a supermarket nearby. Trillions of dollars have been spent on building highways. This alone is a tremendous waste, before we even consider air pollution and dependence on foreign oil. In most cars one person uses hundreds of horsepowers. Can't we do better?

The answer to this question is public transportation. If there were no cars but there were reliable trains and buses, everyone could get anywhere quickly, without traffic jams (Watson, Bates and Kennedy [14], Willis [15]).

Let us see what happens in a more realistic situation when people do have cars, but buses are also available. If everyone is using a car, the roads are clogged and no one can move. If some people still decide to choose the bus, they will be at an advantage because of the special lanes for buses, but both the car drivers (defectors) and the bus riders (cooperators) are punished for the irresponsible behavior of the majority. The defectors are punished more severely than the cooperators. On the other hand, if everyone uses the bus, the buses will be crowded but they can freely move on the empty streets. All passengers can get to their destinations relatively quickly, which corresponds to a reward. If some defectors choose to use their cars anyway, they get an even larger reward because they can move even faster and also avoid the crowd in the bus. This situation is exactly the same as what we have investigated in this chapter.

We have provided some insight into the conditions of decentralized cooperation in spatially distributed populations of agents. However, many questions remain open. The most important of them is a reliable method of choosing the parameters of the payoff functions. Future research may find answers to these questions. As a result, the study of N-person Chicken games may lead us to a better understanding of some basic social dynamics, the factors stimulating or inhibiting cooperative behavior, and the emergence of social norms. We may even obtain some insight into the possibility of changing human behavior.

References

[1] Brams S. (1975), *Game Theory and Politics*, Free Press, New York.

[2] Flache A. and Macy M.W. (1996), Weakness of strong ties: Collective action failure in a highly cohesive group, *J. Math. Sociology*, **21**, 3–28.

[3] Hegselmann R. and Flache A. (1998), Understanding complex social dynamics: A plea for cellular automata based modelling, *J. Artificial Soc. Social Simul.*, **1**, 3.

[4] Komorita S.S. and Parks C.D. (1994), *Social Dilemmas*, Brown and Benchmark, Madison, WI.

[5] Narendra K.S. and Thathachar M.A.L. (1989), *Learning Automata* (*An Introduction*), Prentice–Hall, Englewood Cliffs, NJ.

[6] Poundstone W. (1992), *Prisoner's Dilemma*, Doubleday, New York, 197–201.

[7] Schroeder, D.A. ed. (1995), *Social Dilemmas: Perspectives on Individuals and Groups*, Praeger, Westport, CT.

[8] Snidal D. (1991), Relative gains and the pattern of international cooperation, *Amer. Political Sci. Rev.*, **85**, 701–726.

[9] Szidarovszky F. and Szilagyi M.N. (2002), An analytical study of the N-person prisoners' dilemma, *Southwest J. Pure Appl. Math.*

[10] Szilagyi M.N. (2003), An investigation of N-person Prisoners' Dilemmas, *Complex Systems*, **14**, 155–174.

[11] Szilagyi M.N. and Szilagyi Z.C. (2000), A tool for simulated social experiments. *Simulation*, **74**:1, 4–10.

[12] Szilagyi M.N. and Szilagyi Z.C. (2002), Nontrivial Solutions to the N-Person Prisoners' Dilemma, *Systems Res. Behavioral Sci.*, **19**:3, 281–290.

[13] Thorndike E.L. (1911), *Animal Intelligence*, Hafner, Darien, CT.

[14] Watson A.Y., Bates R.R. and Kennedy D. (1988), *Air Pollution, the Automobile, and Public Health*. National Academy Press, Washington, D.C.

[15] Willis T. (1992), *Cars: an Environmental Challenge*, Chicago Press, Chicago.

[16] Wolfram S. (1994), *Cellular Automata and Complexity*, Addison–Wesley, Reading, MA.

The Optimal Trajectory in the Partial-Cooperative Game

Onik Mikaelyan
Department of Mechanics
Yerevan State University
1 Alex Manoukyan Street
375025 Yerevan, Republic of Armenia
mikons51@yahoo.com

Rafik Khachaturyan
Department of Mechanics
Yerevan State University
1 Alex Manoukyan Street
375025 Yerevan, Republic of Armenia

Abstract

This chapter investigates a partially cooperative game in an extended form. The method of finding the optimal behavior of players and the value for such games are presented. During the course of the game, auxiliary games and Nash's equilibrium situations are considered to define the distribution between players of a coalition. An example is also presented.

Introduction

The optimal trajectory is constructed on the basis of the finite positional game of n players with the terminal payoff in [1] for a partial-cooperative game. Shapley's vector is used to define the distribution between the players of the summative payoff of a coalition in the auxiliary cooperative game. The condition of partial cooperation increases the players' possibilities in comparison with the condition of perfect cooperation for multi-choice and multistage [2, 3] games. In the discussed partially cooperative game we followed the steps from [1] and used the algorithm there which we called Petrosyan–Ayoshin's algorithm. To find the portion of a player from the coalitional payoff we used players' payoffs which corresponded to a Nash equilibrium [4]. In the auxiliary game, a coalition of players may appear as one player and consequently we will find the distribution between the players in this game.

1 The basic model

Let Γ be the dynamic positional n-person game with complete information. We denote the set of players via $N = \{1, \ldots, n\}$. Let $K(x_0)$ be a game tree with initial position x_0. According to the definition of a positional game, on $K(x_0)$ a decomposition of the position's set into $n+1$ sets $P_1, \dots, P_n, P_{n+1}$ is given, where P_i is the set of the personal positions of the player i, and P_{n+1} is the set of the terminal positions. In the game Γ the payoffs of the players are defined by the real functions $h_i : P_{n+1} \longrightarrow R^1_+, i \in N$.

Definition 1.1. Partial-cooperative behavior is the behavior when a player can be either in the cooperation or can play individually. The player who is accepting a decision in the position x (who is doing a move, i.e., removing the "moving point" from the position x to an alternative position), we'll denote by $i(x)$, $(x \in P_{i(x)})$. Let's change the positional game Γ, assuming that players can cooperate for certain conditions. The transformed game we will call the partial-cooperative game and we'll denote it again by Γ.

In this section the rules of a partial cooperation are described.

Definition 1.2. Let $i \in N$. The function $f_i : P_i \to \{0, 1\}$ is called i player's cooperative function, if for every move $\{x_0, \dots, x', x'', \dots, \bar{x}\}$, where $x' \in P_i$ and $\bar{x} \in P_{n+1}$, from the condition $f_i(x') = 1$ follows that for every position $y \in P_i \cap \{x'', \dots, \bar{x}\}$, if it exists, $f_i(y) = 1$.

The vector function $f = (f_1, \dots, f_n)$ is called the cooperative function of the game.

The cooperative function makes it possible to show the coalition $C(x)$ in every position x:

$$C(x) = \begin{cases} S_f(x), & \text{if } f_i(x) = 1, \\ \{i\}, & \text{if } f_i(x) = 0, \end{cases}$$

where $S_f(x) = S^1_f(x) \cup S^2_f(x)$,

$$S^1_f(x) = \{j \in N \mid P_j \cap K(x) = \emptyset \quad \text{and} \quad \exists y \in P_j \cap \{x_0, \dots, x\} : f_j(y) = 1\},$$

$$S^2_f(x) = \{j \in N \setminus S^1_f(x) \mid f_j(y) = 1, \quad \forall y \in P_j \cap K(x)\}.$$

The set $S^1_f(x)$ consists of players who have displayed cooperative activity along the way, leading to the x and don't make decision during the further development of the game on the subtree $K(x)$ with a final position x. From the definition of the cooperative function we'll count players belonging to $S^1_f(x)$. The set $S^2_f(x)$ includes players who cooperate in all their personal positions of the subtree $K(x)$. By saying that the player $i \in N$ keeps a cooperative behavior in the position $x \in K(x_0)$, we mean that in the position x the player i makes a decision which proceeds from interests of the coalition (i.e., it is striving to reach the maximal summative payoff of players $i \in C(x)$). The players of the set $N \setminus S_f(x)$ in the position x are considered individual players. Since $S_f(x)$ is defined by the cooperative function f, the coalition structure in the position x, $S_f(x), \{j_1\}, \{j_2\}, \dots, \{j_{|N \setminus S_f(x)|}\}$, also forms depending on f.

Now let's define the positional game $\Gamma_f(x_0)$. The game $\Gamma_f(x_0)$is created on the basis of the partial-cooperative game Γ by means of the cooperative function f.

The tree of the game $\Gamma_f(x_0)$ coincides with the tree $K(x_0)$ of the game Γ. The set N_f of participants of the game $\Gamma_f(x_0)$ is formed by means of the cooperative function f and consists of subsets of the set N.

$$N_f = \{C \subset N : \exists x \in K(x_0), C(x) = C\}.$$

The set $C \in N_f$ will be considered as a participant of the game in the positional game $\Gamma_f(x_0)$, which makes a decision in the positions $x \in W(C)$, where

$$W(C) = \{x \in K(x_0) : C(x) = C\}.$$

The payoff of the player $C \in N_f$ of the game $\Gamma_f(x_0)$ is defined on the set of final positions of the game tree $K(x_0)$ as a sum of payoffs of players $i \in C$ of the game Γ:

$$h_S(x) = \sum_{i \in C} h_i(x), \quad x \in P_{n+1}, \quad h_i(x) \geq 0, \quad i \in N.$$

It's possible that every set $C \in N_f$ will include only one player. It takes place when for every position x of a tree of the game $K(x_0)$ the set $S_f(x)$ includes one element or it is empty. $\Gamma_f(x_0)$ may be a one-player game (the set N_f includes only the player N) as an opposite case. This will happen if $C(x) = N$ for every position $x \in P_i$, $i \in N$. In the more complicated cases, the set N_f may consist of all the subsets of the set N.

Particularly from the definition of the partial-cooperative game we can form the following conclusion: It's possible that a player, in spite of the fact that he is in his cooperative behavior position, is playing individually.

One of the ways to get the solution of the positional game $\Gamma_f(x_0)$ is the construction of Nash equilibrium situations. The payoff of each player $C \in N_f$ in the game $\Gamma_f(x_0)$ will be defined in accordance with Nash equilibrium situations (which, of course, can be not unique, but it always exists because Γ is a game with complete information). So we can construct a function $\Phi(C)$, $C \in N_f$, which defines the payoff of every coalition arising during the game Γ in accordance with the cooperative function f.

However, it is still incomprehensible how we can define the distribution of the payoff of a coalition between players by means of the function $\Phi(C)$. Now we use another approach to find a convenient solution. That approach is based on applying the Nash equilibrium in the sense of a coalition. To construct the solution we'll use the method of backwards induction, step by step finding the optimal behavior of players.

2 Algorithm for Construction of Optimal Way

In this section we offer the method of construction of a solution of the game $\Gamma_f(x_0)$ that also leads to the construction of the corresponding optimal way.

We construct the solution of the game $\Gamma_f(x_0)$ by the method of backwards induction, moving from the final positions to the primary position. The procedure for seeking the solution is still the scheme of the construction of the Nash equilibrium in the usual positional game. The existing differences are as follows. Let the subtree $K(x)$ belong to the domain of the cooperative behavior of the player i. As directed, the cooperative function f defines some coalitional structure for every position. Then in the set of final positions of the subtree $K(x)$ instead of payoffs of the player i it's necessary to consider the payoffs of coalitions which include the player i. On the subtree $K(x)$, the solution of the player i, which maximizes the payoff of the coalition to which it (the player) belongs, can be defined easily by using the scheme of Nash. But, since the payoff of the player i does not differ from the payoff of the coalition, in the definition of the player i in it's personal positions which are between position x and primary position x_0, where the player i is playing individually, difficulties occur. If the share of the player i in the payoff of the coalition is known, then, using the scheme of Nash again, we can find the solution of the player i on it's (player's) personal positions along the way $\{x_0, \ldots, x\}$. So, the definition of changes of the payoff, for a transition of the player from cooperative behavior to individual behavior, pretends the main problem which is considered in our algorithm.

Hereafter, we'll use the following assignments. Let x be some position. Let us denote by $Z(x)$ the set of positions which are just next to x. We denote the player who is making a decision in the position x, $x \in P_i$ in the game Γ by $i(x) \in N$. We'll say that the decision of the player $i(x)$ in the position x will move to the position $\bar{x} \in Z(x)$. Let us introduce the auxiliary function c_f which is defined by means of cooperative function $f = (f_1, \ldots, f_n)$:

$$c_f(x) = \begin{cases} 1, & \text{if } f_{i(x)}(x) = 1, \\ 0, & \text{if } f_{i(x)}(x) = 0. \end{cases}$$

Let the length of the game $\Gamma_f(x_0)$ consist of $T + 1$ positions. Consider the decomposition of the set of all positions of the tree of the game $K(x_0)$ into $T + 1$ sets $X_0, X_1, \ldots, X_t, \ldots, X_T = \{x_0\}$, where the set X_t consists of positions which are reachable from initial position x_0 after $T - t$ steps. We will denote by x_t, $t = 1, \ldots, T$ the set of positions which belong to the set X_t.

Initial Step. Let's consider the set of final positions P_{n+1}. The coalition structure in the position $x \in P_{n+1}$ coincides with the coalition structure in the position x_1, $x \in Z(x_1)$. According to the cooperative function f, in the position x_1 coalitions $S_f(x_1), \{j_1\}, \ldots, \{j_{|N \setminus S_f(x_1)|}\}$ are formed. The payoff of the player $i_f = S_f(x_1)$ in the position $x \in Z(x_1)$ is equal to

$$\sum_{i \in S_f(x_1)} h_i(x).$$

The payoff of the player $i_f = \{j_k\}, k = 1, \ldots, |N \setminus S_f(x_1)|$ in the terminal position x is $h_{jk}(x)$.

Step 1. Now move $Z(x_1)$, $x_1 \in X_1$ from the final positions to the preceding positions. Consider the position x_1. Let $c_f(x_1) = 1$. Then the player $i(x_1) \in N$ is cooperated and in the position x_1 is moving the player $i_f(x_1) = S_f(x_1)$, $i_f(x_1) \in N_f$. We ascribe to the player $i_f(x_1)$ to choose the position $\bar{x}_1 \in Z(x_1)$ from the condition

$$\max_{x \in Z(x_1)} \sum_{i \in S_f(x_1)} h_i(x) = \sum_{i \in S_f(x_1)} h_i(\bar{x}_1).$$

If $c_f(x_1) = 0$ then the player $i(x_1)$ is not cooperating. Hence, $i_f(x_1) = \{i(x_1)\}$. In this case we ascribe to the player $i_f(x_1)$ to choose the position $\bar{x}_1$ from the condition

$$\max_{x \in Z(x_1)} h_{i(x_1)}(x) = h_{i(x_1)}(\bar{x}_1).$$

Applying analogous reasonings we can create a way with an origin in $x_1 \in X_1$ for every position x_1 of the set X_1. So, in every subtree $K(x_1)$, $x_1 \in X(x_1)$, the position $\bar{x}_1$ is fixed which is pretending to be the supposed final position of constructing way of the game $\Gamma_f(x_0)$. Therefore, instead of considering the terminal functions h_i, $i \in N$ in the set P_{n+1} of final positions, we can use functions $r_i^1 : X_1 \to R_+^1$, $i \in N$ which are given in the set X_1:

$$r_i^1(x_1) = \begin{cases} h_i(\bar{x}_1), & \text{if } x_1 \notin P_{n+1}; \\ h_i(x_1), & \text{if } x_1 \in P_{n+1}. \end{cases}$$

Step 2. We continue moving to the core of the game tree and find decisions of the players $i_f \in N_f$ in the positions $x_2 \in X_2$. If in the set X_1 the payoff of each player is known $i_f(x_2) \in N_f$, $x_2 \in X_2$, then we can construct the trajectory of the game on the subtrees $K(x_2)$, $x_2 \in X_2$.

Consider the set $Y(x_2) = Y_1(x_2) \cup Y_2(x_2)$, where

$$Y_1(x_2) = \{x \in Z(x_2) \mid c_f(x_2) = 0, \qquad i(x_2) \in S_f(x)\}$$

and

$$Y_2(x_2) = \{x \in Z(x_2) \mid c_f(x_2) = 1 \quad \text{and} \quad S_f(x) \setminus S_f(x_2) \neq \emptyset\}.$$

In each position of the set $Y(x_2) \subset Z(x_2)$ the payoff of the coalition $i_f(x_2) = \{i(x_2)\}$ when $c_f(x_2) = 0$, or of the coalition $i_f(x_2) = S_f(x_2)$ when $c_f(x_2) = 1$ is not separated from the payoff $\sum_{i \in S_f(x_1)} r_i^1(x_1)$ of coalition $i_f(x_1) = S_f(x_1)$, $x_1 \in Z(x_2)$. Since the payoff of the player $i(x_2)$ is not separated from the payoff $\sum_{i \in S_f(y_1)} r_i^1(\bar{y}_1)$ of the coalition $S_f(y_1)$, then his payoff in the position $y_1 \in Z(x_2)$ is not known.

In the common case there is an absence of information about payoffs in the positions where the coalition structure changes with regard to the current player which is making the decision. For each position $x_2 \in X_2$ we'll consider two basic cases.

1) Let $Y(x_2) = \emptyset, c_f(x_2) = 0$. Hence, in the position x_2 the player $i_f(x_2) = \{i(x_2)\}$ is making the decision. We ascribe to the player $i_f(x_2)$ to choose the position $\bar{x}_2 \in Z(x_2)$ from the condition

$$\max_{x \in Z(x_2)} r^1_{i(x_2)}(x) = r^1_{i(x_2)}(\bar{x}_2).$$

Now let $c_f(x_2) = 1$. Since $Y(x_2) = \emptyset$, coalitions $S_f(x_2)$ and $S_f(x_1)$ coincide. In this case we'll ascribe to the player $i_f(x_2) = \{i(x_2)\}$ to choose the position $\bar{x}_2 \in Z(x_2)$ from the condition

$$\max_{x \in Z(x_2)} \sum_{i \in S_f(x_2)} r^1_i(x) = \sum_{i \in S_f(x_2)} r^1_i(\bar{x}_2).$$

2) Let $Y(x_2) \neq \emptyset$. As shown above, uncertainty arises connected with the payoffs of the coalition $i_f(x_2) = \{i(x_2)\}$ when $c_f(x_2) = 0$, and of coalition $i_f(x_2) = S_f(x_2)$ when $c_f(x_2) = 1$. In order to construct the way of the game on the subtree $K(x_2)$, it's necessary to find some distribution of the payoff of the coalition $S_f(y_1)$ for each position $y_1 \in Y(x_2)$. In order to find the required distribution we'll describe the following procedure. Let's consider an auxiliary game $\Gamma_f(y_1)$ on the subtree $K(Y_1)$ with a set of players $S_f(y_1), \{j_1\}, \dots, \{j_{|N \setminus S_f(y_1)|}\}$. Out of coalition $S_f(y_1)$ players' strategies are fixed and coincide with Nash equilibria strategies in the basic game. In this game $\Gamma_f(y_1)$ the Nash equilibrium situation exists, because the considered game is a game with complete information. Let's denote it by x.

Now we'll define the coefficient of cooperation α_i for the player $i \in S_f(y_1)$:

$$\alpha_i = \frac{h_i(x)}{\sum_{k \in S_f(y_1)} h_k(x)}.$$

Then we ascribe to the player $i \in S_f(y_1)$ in the game $\Gamma_f(y_1)$ a payoff which is equal to

$$\alpha_i \nu_f(y_1, S_f(y_1), \{j_1\}, \{j_2\}, \dots, \{j_{|N \setminus S_f(y_1)|}\}) = Nh^f_i(y_1),$$

where

$$\nu_f(y_1, S_f(y_1), \{j_1\}, \{j_2\}, \dots, \{j_{|N \setminus S_f(y_1)|}\}) = \sum_{i \in S_f(y_1)} r^1_i(y_1)$$

is the payoff of the coalition $S_f(y_1)$. The distribution of the summative payoff of coalition $S_f(y_1)$ will be the following vector:

$$Nh^f(y_1) = \left(Nh^f_{k_1}(y_1), \dots, Nh_{k_{|S_f(y_1)|}}(y_1)\right),$$

where

$$\sum_{j=1}^{|S_f(y_1)|} Nh^f_{k_j}(y_1) = \nu_f\left(y_1, S_f(y_1), \{j_1\}, \{j_2\}, \dots, \left\{j_{|N \setminus S_f(y_1)|}\right\}\right)$$

is an optimal distribution in the game $\Gamma_f(y_1)$. If the player $\{i(x_2)\} \in N_f$ in the position x_2 selects the position $y_1 \in Y(x_2)$, then his payoff is defined by the vector $Nh^f(y_1)$, which is equal to $Nh^f_{i(x_2)}(y_1)$. So, on the set X_1 a new function of payoffs is given $\bar{r}_i^1 : X_1 \to R^1_+$, $i \in N$, so that for $x_1 \in Z(x_2)$

$$\bar{r}_i^1(x_1) = \begin{cases} Nh_i^f(x_1) & \text{if } x_1 \in Y(x_2),\ i \in S_f(x_1); \\ r_i^1(x_1), & \text{in the opposite case.} \end{cases}$$

Let $c_f(x_2) = 0$. Then for the player $i_f(x_2) = \{i(x_2)\}$ realization of the way, which is passing through the position $\bar{x}_2 \in Z(x_2)$ is optimal. For $\bar{x}_2 \in Z(x_2)$, $\max_{x \in Z(x_2)} \bar{r}^1_{i(x_2)}(x) = \bar{r}^1_{i(x_2)}(\bar{x}_2)$. Now suppose $c_f(x_2) = 1$. Since the player $i(x_2)$ is cooperating, then in the position x_2 the coalition $i_f(x_2) = S_f(x_2)$ is moving. We ascribe to it (the coalition) to choose the position $\bar{x}_2$ for which

$$\max_{x \in Z(x_2)} \sum_{i \in S_f(x_2)} \bar{r}^1_{i(x_2)}(x) = \sum_{i \in S_f(x_2)} \bar{r}^1_{i(x_2)}(\bar{x}_2).$$

So, on each subtree $K(x_2)$, $x_2 \in X_2$, the way is constructed. Hence, to construct the way of the game on subtrees $K(x_3)$, $x_3 \in X_3$, it's enough to define the decisions of the players $i_f(x_3) \in N_f$, $x_3 \in X_3$. Now on the set X_2 we construct the functions $r_i^2 : X_2 \to R^1_+$, $i \in N$, so that for $x_2 \in X_2$ and $i \in N$ we have

$$r_i^2(x_2) = \begin{cases} r_i^1(\bar{x}_2) & \text{if } Y(x_2) = \emptyset; \\ \bar{r}_i^1(\bar{x}_2) & \text{if } Y(x_2) \neq \emptyset; \\ h_i(x_2) & \text{if } x_2 \in P_{n+1}. \end{cases}$$

The following steps of the procedure are similar to Steps 1 and 2. We'll omit the description of each step, and now consider Step t. By proceeding the movement to the core x_0 of the game tree, we reach the positions $x_t \in X_t$. Let the functions $r_i^{t-1} : X_{t-1} \to R^1_+$, $i \in N$, define $i \in N$ players' payoffs after the players $i_f(x_{t-1}) \in N_f$, $x_{t-1} \in X_{t-1}$ perform the prescribed decisions.

Step *t*. We will not discuss the final positions of the set $X_t \cap P_{n+1}$ because they are just like those in section 3.2. Let's define decisions of the players $i_f \in N_f$ in the intermediate positions $X_t \setminus P_{n+1}$ and $Y(x_t) = Y_1(x_t) \cup Y_2(x_t)$, where

$$Y_1(x_t) = \{x \in Z(x_t) | c_f(x_t) = 0 \quad \text{and } i(x_t) \in S_f(x)\},$$

and

$$Y_2(x_t) = \{x \in Z(x_t) | c_f(x_t) = 1 \quad \text{and } S_f(x) \setminus S_f(x_t) \neq \emptyset\}.$$

First let's discuss the case when construction of the new functions of the payoff isn't necessary.

1) Now suppose $Y(x_t) = \emptyset$ for all positions $x_t \in X_t \setminus P_{n+1}$. According to the functions r_i^{t-1}, $i \in N$, if the decision of the player $i_f(x_t)$ leads the game to the position $\bar{x}_t \in Z(x_t)$, then the payoffs which the players $i_f \in N_f$ receive at the end of the game will be equal to $\sum_{i \in S_f(x_t)} r_i^{t-1}(\bar{x}_t)$ for the player $i_f = S_f(x_t)$, and to $r_{jk}^{t-1}(\bar{x}_t)$ for the players $i_f = \{j_k\}$, $k = 1, \ldots, |S_f(x_t)|$ correspondingly. If $c_f(x_t) = 0$, then in the position x_t the player $i_f = \{i(x_t)\}$ is moving and we ascribe to him to choose the position $\bar{x}_t$ from the condition

$$\max_{x \in Z(x_t)} r_{i(x_t)}^{t-1}(x) = r_{i(x_t)}^{t-1}(\bar{x}_t).$$

If $c_f(x_t) = 1$, then in the position x_t the player $i_f = S_f(x_t)$ is moving. We ascribe to him to choose the position $\bar{x}_t$ from the condition

$$\max_{x \in Z(x_t)} \sum_{i \in S_f(x_t)} r_{i(x_t)}^{t-1}(x) = \sum_{i \in S_f(x_t)} r_{i(x_t)}^{t-1}(\bar{x}_t).$$

2) For some position x_t, let the set $Y(x_t) \subset Z(x_t)$ not be empty. It consists of positions for which the payoff of the player $i_f(x_t) \in N_f$ is not defined. To find the decision of the player $i_f(x_t)$ in the position x_t, for every position $y_{t-1} \in Y(x_t)$ we consider an auxiliary game $\Gamma_f(y_{t-1})$. Players' strategies are fixed out of coalition $S_f(y_{t-1})$ and coincide with Nash equilibria strategies in the basic game. The payoff of the coalition $S_f(y_{t-1})$ in that game is equal to

$$\nu_f\left(y_{t-1}, S_f(y_{t-1}),\ \{j_1\}, \{j_2\}, \ldots, \left\{j_{|N \setminus S_f(y_{t-1})|}\right\}\right) = \sum_{i \in S_f(y_{t-1})} r_i^{t-1}(y_{t-1}),$$

Let's define the distribution of the summative payoff of the coalition $S_f(y_{t-1})$ in the game $\Gamma_f(y_{t-1})$ in the following way: in this game a Nash equilibrium situation with complete information exists. We denote the corresponding issue by x. We'll define the coefficient of a cooperation α_i for the player $i \in S_f(y_{t-1})$ in the following way:

$$\alpha_i = \frac{h_i(x)}{\sum_{k \in S_f(y_{t-1})} h_k(x)}.$$

Then we ascribe to the player $i \in S_f(y_{t-1})$ in the game $\Gamma_f(y_{t-1})$ a payoff is equal to

$$\alpha_i \nu_f\left(y_1, S_f(y_{t-1}), \{j_1\}, \{j_2\}, \ldots, \left\{j_{|N \setminus S_f(y_{t-1})|}\right\}\right) = Nh_i^f(y_{t-1}).$$

The distribution of the summative payoff of the coalition $S_f(y_{t-1})$ will be the vector

$$Nh^f(y_{t-1}) = \left(Nh_{k_1}^f(y_{t-1}), \ldots, Nh_{k_{|S_f(y_{t-1})|}}^f(y_{t-1})\right),$$

as

$$\sum_{j=1}^{|S_f(y_{t-1})|} Nh_{k_j}^f(y_{t-1}) = \nu_f(y_{t-1}, S_f(y_{t-1}), \{j_1\}, \{j_2\}, \ldots, \{j_{|N \setminus S_f(y_{t-1})|}\}).$$

Distribution (2.24) we'll name the optimal distribution in the auxiliary game.

$$\bar{r}_i^{t-1}(x_{t-1}) = \begin{cases} Nh_i^f(x_{t-1}), & \text{if } x_{t-1 \in Y(x_t)}, i \in S_f(x_{t-1}); \\ r_i^{t-1}(x_{t-1}), & \text{in the opposite case.} \end{cases}$$

If $c_f(x_t) = 0$ then we ascribe to the player $i_f(x_t) = \{i(x_t)\}$ to choose the position $\bar{x}_t \in Z(x_t)$ from the condition

$$\max_{x \in Z(x_t)} \bar{r}^t_{i(x_t)}(x) = \bar{r}^t_{i(x_t)}(\bar{x}_t).$$

If $c_f(x_t) = 1$ then in the position x_t the player $i_f(x_t) = S_f(x_t)$ is making a decision. We ascribe to him to choose the position $\bar{x}_t$ from the condition

$$\max_{x \in Z(x_t)} \sum_{i \in S_f(x_t)} \bar{r}^t_{i(x_t)}(x) = \sum_{i \in S_f(x_t)} \bar{r}^t_{i(x_t)}(\bar{x}_t).$$

For $i \in N$ and $x_t \in X_t$ let's define the functions r_i^t

$$r_i^t(x_t) = \begin{cases} r_i^{t-1}(\bar{x}_t), & \text{if } Y(x_t) = \emptyset; \\ \bar{r}_i^{t-1}(\bar{x}_t), & \text{if } Y(x_t) \neq \emptyset; \\ h_i(x_t), & \text{if } x_t \in P_{n+1}. \end{cases}$$

Continuing to go down along the tree of the game $K(x_0)$ to the initial position x_0 and defining by succession the decisions of the players $i_f \in N_f$ on the remaining sets X_τ, $\tau = t+1, \ldots, T$, we'll construct the way realized in the game Γ if the cooperative function $f = (f_1, \ldots, f_n)$ is given. This procedure we will call *the optimal way of the partial-cooperative game* $\Gamma_f(x_0)$ and denote it by $x(f)$.

3 The Value of the Game $\Gamma_f(x_0)$

According to the procedure proposed above, we found the payoffs of the players $i \in N$, which correspond to the optimal way $x(f)$. This payoffs are defined as $r_i^T(x_0)$, $i \in N$. We'll call the vector $r(x_0) = (r_1^T(x_0), \ldots, r_n^T(x_0))$ the value of the partial-cooperative game.

Example 3.1. Let's consider the positional game Γ with the game tree $K(x_0)$ represented in Figure 1. Here $N = \{1, 2, 3, 4, 5\}$. Individual positions of player 1 are x_0, x_{41}, and x_{46}, of player 2 are x_{11}, x_{23}, and x_{24}, of player 3 are x_{12} and x_{42}, of player 4 are x_{22}, x_{31}, and x_{55}, and of player 5 are x_{32} and x_{34}. The payoffs are written in the final positions; moreover, in each column the upper number is the payoff of player 1 and so on. Let the cooperative function $f = (f_1, \ldots, f_5)$ have the following form:

$$f_1(x_0) = f_1(x_{41}) = F_1(x_{46}) = 0, \quad f_2(x_{11}) = 1 = f_2(x_{23}),$$
$$f_2(x_{24}) = 0, \quad f_3(x_{12}) = f_3(x_{42}) = 1, \quad f_4(x_{22}) = 0,$$
$$f_4(x_{31}) = f_4(x_{55}) = 1, \quad f_5(x_{32}) = f_5(x_{34}) = 0.$$

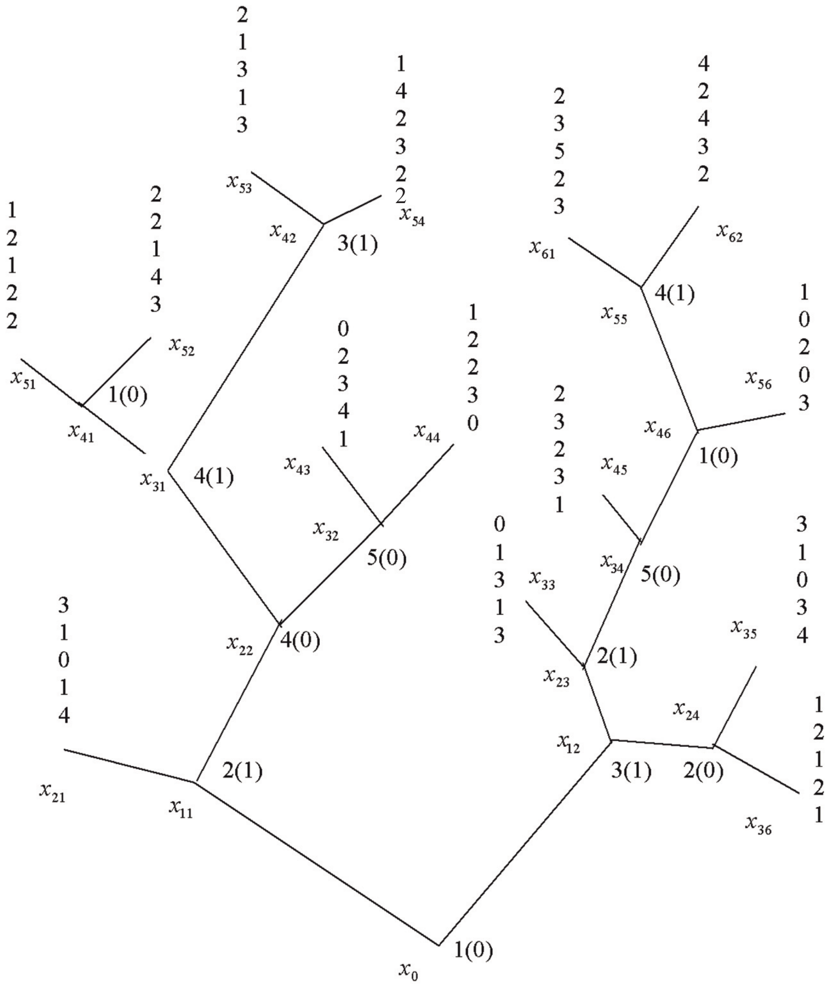


Figure 1.

Let's define coalitions in some positions where they can arise according to the cooperative function f:

$$S_f(x_{42}) = \{2; 3; 4\}, \quad S_f(x_{31}) = \{2; 3; 4\}, \quad S_f(x_{11}) = \{2; 3\},$$
$$S_f(x_{55}) = \{2; 3; 4\}, \quad S_f(x_{23}) = \{2; 3; 4\}, \quad S_f(x_{12}) = \{3; 4\},$$
$$S_f(x_{41}) = \{2; 4\}.$$

Coalitional structures can also be found in other positions. In the result, the set of the players in the game $\Gamma_f(x_0)$ is

$$N_f = \{1; 2; 3; 4; 5; \{2; 3\}; \{2; 3\}; \{3; 4\}; \{2; 3; 4\}\}.$$

The payoffs of the players {2; 3}, {2; 4} and {2; 3; 4} $\in N_f$ in $\Gamma_f(x_0)$ are given by the sum of the corresponding terminal payoffs of players which are in cooperation in the game Γ.

Now we'll construct the optimal trajectory of the partial-cooperative game $\Gamma_f(x_0)$. We began the procedure of construction of the optimal trajectory from the set X_0 which consists of final positions x_{61}, x_{62}. The coalitional structure in positions where the player is performing the last step in the game Γ_f is the same as it is in the corresponding terminal positions. We begin the procedure of finding the optimal way from the farthest from x_0 by number of steps terminal positions, i.e., from the consideration of one-step subtrees $K(x)$. According to the procedure described above, we'll decrease the number of steps in the game $\Gamma_f(x_0)$ stage by stage.

Only $K(x_{55})$ corresponds to the first stage. Coalition structures in x_{61}, x_{62}, and x_{55} are the same, i.e., $S_f(x_{55}) = \{2; 3; 4\}, \{1\}, \{5\}$. In x_{55} the cooperating player 4 is moving. The payoffs of the players from $S_f(x_{55})$ in the x_{61} and x_{62} are given in the form of triplets of numbers (10; 2; 3) and (9; 4; 2) correspondingly, where the first component is the payoff of players {2; 3; 4}, the second is the payoff of player 1, and the third is the payoff of player 5. Although for player 4 personally it is profitable to choose the position x_{62}, proceeding from the coalition opinions, he chooses x_{61}, where the coalition payoff, which is equal to $3 + 5 + 2 = 10$, is more than it is in x_{62}, where it's equal to $2 + 4 + 3 = 9$. Thus $r^1(x_{55}) = (2; 3; 5; 2; 3)^*$ (where * means transposition).

The subtrees $K(x_{41})$, $K(x_{42})$, and $K(x_{46})$ are disposed to the second stage. A coalition structure in the position x_{41} is the same as it is in the final positions x_{51} and x_{52}, i.e., $S_f(x_{41}) = \{2; 4\}, \{1\}, \{3\}, \{5\}$. Then payoffs for the players from N_f in the positions x_{51} and x_{52} are given in the form of four numbers (4; 1; 1; 5) and (6; 2; 1; 3), where the first component is the payoff of player {2; 4}, the second of player 1, the third of player 3, and the fourth of player 5. As in x_{41} the uncooperating player 1 is moving. Then it's profitable for him to choose the position x_{52}, where his payoff is equal to 2. Then $r^2(x_{41}) = (2; 2; 1; 4; 3)^*$. In position x_{42} the cooperating player 3 is moving as a member of coalition {2; 3; 4}, he is choosing x_{54}, where the coalition payoff, equal to $4 + 2 + 3 = 7$, is more than it is in x_{53}, where it's equal to $1 + 3 + 1 = 5$. So $r^2(x_{42}) = (1; 4; 2; 3; 2)^*$. In the case of transition from x_{55} to the individual position of player 1, x_{46}, we'll get $r^2(x_{46}) = (2; 3; 5; 2; 3)^*$ by the above-mentioned method.

The positions x_{31}, x_{32}, and x_{34} are in the third stage. In this stage also there aren't any coalition changes, hence $r^3(x_{31}) = (1; 4; 2; 3; 2)^*$, $r^3(x_{32}) = (0; 2; 3; 4; 1)^*$, $r^3(x_{34}) = (2; 3; 5; 2; 3)^*$.

The positions x_{22}, x_{23}, and x_{24} form the fourth stage of construction of the optimal way. Since in x_{22} player 4 is leaving the coalition {2; 3; 4} and $Y(x_{22}) = \{x_{31}\}$, it's necessary to consider the game $\Gamma(x_{31})$. In this game x_{52} is the Nash equilibrium point,

$$\sum_{i \in \{2;3;4\}} h_i(x_{52}) = 2 + 1 + 4 = 7,$$

the coefficients of the cooperation are equal to $\alpha_2 = \frac{2}{7}$, $\alpha_3 = \frac{1}{7}$, $\alpha_4 = \frac{4}{7}$. The payoff of coalition {2; 3; 4} in x_{31} is equal to

$$V_f(x_{31}, \{2; 3; 4\}, \{1\}, \{5\}) = 4 + 2 + 3 = 9.$$

Then the distribution of the summative payoff in x_{31} will be

$$Nh_2^f(x_{31}) = \alpha_2 \cdot V_f = \frac{18}{7}, \quad Nh_3^f(x_{31}) = \frac{9}{7}, \quad Nh_4^f(x_{31}) = \frac{36}{7}.$$

So, by means of the new function of payoff, in x_{31} we'll have

$$r^3(x_{31}) = (r_1^3(x_{31}), \bar{r}_2^3(x_{31}), \bar{r}_3^3(x_{31}), \bar{r}_4^3(x_{31}), r_5^3(x_{31}))^* = \left(1; \frac{18}{7}; \frac{9}{7}; \frac{36}{7}; 2\right)^*.$$

As an individual player, player 4 in x_{22} chooses the payoff in x_{31} which is $r^4(x_{22}) = (1; \frac{18}{7}; \frac{9}{7}; \frac{36}{7}; 2)^*$. In positions x_{23} and x_{24} there are no coalition changes, thus $r^4(x_{23}) = (2; 3; 5; 2; 3)^*$ and $r^4(x_{24}) = (1; 2; 1; 2; 1)^*$.

The positions x_{11} and x_{12} form the last but one stage. In x_{11} there are no coalition changes. Player 2 from the coalition {2; 3} is choosing the payoff in x_{22}, where the payoff of the coalition {2; 3} is more than in x_{21}. So $r^5(x_{11}) = (1; \frac{18}{7}; \frac{9}{7}; \frac{36}{7}; 2)^*$.

In x_{12} player 2 is leaving the coalition {2; 3; 4}. It's necessary to consider the auxiliary game $\Gamma(x_{23})$, because $Y(x_{12}) = \{x_{23}\}$. Here $NE(x_{23}) = \{x_{61}\}$,

$$\sum_{i \in \{2;3;4\}} h_i(x_{61}) = 3 + 5 + 2 = 10,$$

and the coefficients of cooperation are equal to $\alpha_2 = \frac{3}{10}$, $\alpha_3 = \frac{1}{2}$, $\alpha_4 = \frac{1}{5}$.

We have $V_f(x_{23}, \{2; 3; 4\}, \{1\}, \{5\}) = 3 + 5 + 2 = 10$. Then the summative payoff of the coalition {2; 3; 4} in x_{23} is divided between players of the coalition $S_f(x_{23}) = \{2; 3; 4\}$ in the following way:

$$Nh_2^f(x_{23}) = \alpha_2 \cdot V_f = \frac{3}{10} \cdot 10 = 3, Nh_3^f(x_{23}) = \frac{1}{2} \cdot 10 = 5, \quad \text{and}$$
$$Nh_4^f(x_{23}) = \frac{1}{5} \cdot 10 = 2.$$

So, the new payoff in x_{23} will be

$$r^4(x_{23}) = (r_1^4(x_{23}), \bar{r}_2^4(x_{23}), \bar{r}_3^4(x_{23}), \bar{r}_4^4(x_{23}), r_5^4(x_{23}))^* = (2; 3; 5; 2; 3)^*.$$

In x_{12} the individual player 3 chooses the position x_{23}, where $r^5(x_{12}) = (2; 3; 5; 2; 3)^*$. In the core of the game tree x_0, $Y(x_0) = \emptyset$; therefore, it will be better if player 1 chooses the payoff in x_{12}, where $r_1^5(x_0) = 2$, than in x_{11}, where $r_1^6(x_0) = 1$.

In the result, in the game $\Gamma_f(x_0)$ the $x(f) = \{x_0, x_{12}, x_{23}, x_{34}, x_{46}, x_{55}, x_{61}\}$ pretends to be the optimal way, and the value of the game equals $r(x_0) = (2; 3; 5; 2; 3)$.

By changing the cooperative function f, we'll have the class of all partial-cooperative games $\Gamma_f(x_0)$, which can be defined in the tree $K(x_0)$. But this, in the case of other privileges, also gives the player the chance to find the best cooperative function f for him.

References

[1] Petrosyan L., Ayoshin D. The Value of Dynamic Games with Partial-Cooperation (Rus.). Proceedings of the Institute of Mathematics And Mechanics Ur. D. of RAS. Vol. 6, pp. 160–172, 2000.

[2] Ayoshin D., Tanaka T. The Core And the Dominance Core in Multi-Choice Multistage Games with Coalitions in a Matrix Form. Submitted to the Proceedings of NACA98 (International Conference on Nonlinear Analysis and Convex Analysis), 1998.

[3] Chih-Ru Hsiao, Raghavan T. The Shapley Value for Multi-Choice Cooperative Games (I). Games and Economic Behavior, 1993. Vol. 5, pp. 240–256.

[4] Nash J. F. Equilibrium points in n-person games. Proc. Nat. Acad. Sci. U.S.A., 36, 48–49, 1950.